KENTECH
Korea Institute of Energy Technology

Energy Awakening

2026년 에너지공학 탐구 동아리 탐구과제 자료집

(한국에너지공과대학교 입학기획운영팀)

The steps
you are about to take
will be recorded forever
in world energy history

Contents

I. 베이지안 최적화 : 확률로 여는 최적화의 세계
[국문] 008
[영문] 018

II. 바이오에너지와 CCUS의 만남 : BECCS를 활용한 탄소 제거 원리와 미래 전망
[국문] 030
[영문] 039

III. 배터리 안정성의 핵심, 배터리 관리 시스템 (BMS)
[국문] 050
[영문] 061

IV. 비귀금속 촉매 기반 수전해 기술 - 분리막과 나노 관점에서 이해하는 그린수소 생산
[국문] 074
[영문] 084

V. 전력반도체 : 친환경 에너지 시대로 향하는 길
[국문] 098
[영문] 111

VI. 이미지에서 패턴을 학습하다 - CNN의 개념과 활용
[국문] 126
[영문] 136

VII. 수소 경제 시대를 대비하기 위한 수소 저장 및 운송 기술
[국문] 148
[영문] 157

VIII. 다성분계 고용체에서의 격자 왜곡이 전위 이동 억제 및 기계적 물성에 미치는 영향 : 고엔트로피 합금(HEA)의 강화 기전 분석
[국문] 168
[영문] 183

IX. 분산 에너지 자원의 미래, 마이크로 그리드
[국문] 202
[영문] 211

X. 우주용 전력원으로서의 페로브스카이트 태양전지 : 소재 특성과 방사선 내구성
[국문] 222
[영문] 245



탐구자료는 고등학생들이 에너지 분야에 대한 흥미를 느끼고, 주제에 대해 깊이 탐구할 수 있도록 구성된 학습 자료입니다.

탐구자료를 활용하여 탐구하는 법

1. 학생은 엔진 (enzine.co.kr) - 리서치 - 탐구자료에 업로드된 자료를 참고하여 자신이 탐구하고 싶은 혹은, 관심있는 자료를 자유롭게 선택합니다.
2. 탐구자료의 내용을 이해하고, Open Question을 참고하여 탐구 주제를 선정합니다.
3. 탐구를 진행하는 과정에서 궁금한 점이 생겼을 경우, 엔진 (enzine.co.kr) - 커뮤니티 - 자유게시판을 통해 탐구자료 제작자인 켄들러와 질의응답을 할 수 있습니다.
4. 탐구가 끝난 후에는 엔진 (enzine.co.kr) - 리서치 - 탐구보고서에 업로드되어 있는 탐구보고서 양식에 맞추어 탐구 보고서를 작성하여 제출합니다.

FAQ (자주 하는 질문)

1. **Open Question 모두에 대해 답변해야 하나요?**
   → 아니요. 하나만 선택해도 되고, 새로운 주제를 자유롭게 정해도 됩니다. 하지만 탐구자료와 연관된 주제를 선정하여야 합니다.

2. **추가자료가 있는 탐구자료의 경우, 예시 답변을 탐구보고서에 활용해도 되나요?**
   → 네. 다만 예시 답변 그대로가 아닌, 자신의 탐구 활동이 포함된 탐구보고서를 제출하셔야 합니다.

Tip 7~8월에는 멘토링 활동이 진행됩니다. 탐구활동에 대한 조언을 구하고 싶거나 직접 질의응답을 희망한다면 멘토링에 참여하는 것을 추천합니다.

I

베이지안 최적화 :

확률로 여는 최적화의 세계

p. 008 - 027

베이지안 최적화: 확률로 여는 최적화의 세계

Abstract
베이지안 최적화(Bayesian Optimization, BO)는 현실의 문제를 효율적으로 탐색할 수 있는 방법이다. 본 자료는 BO의 기본 원리를 기반으로, 탐험과 활용의 균형을 이루며 최적점을 찾아가는 과정을 설명한다. 또한 다양한 응용 사례를 통해 실제 활용 가능성과 앞으로의 과제를 살펴본다. 이를 통해 베이지안 최적화가 단순한 기법을 넘어, 불확실한 현실 속에서 확률적 사고를 기반으로 최적화 문제를 해결하는 통계학적 방법임을 고찰한다.

Introduction

현대의 많은 문제들은 방대한 탐색 공간 속에서 최적의 해답을 찾아야 한다는 공통점을 지닌다. 이러한 문제들의 복잡성은 무한에 가까운 가능성 중 단 하나의 최적 해답을 찾아야 한다는 데서 비롯된다. 예를 들어, 새로운 촉매를 발견한다고 해보자. 이론적으로 수많은 원소와 조합이 가능하지만, 실제로 모든 경우를 실험하는 것은 불가능하다. 가능한 선택지는 수없이 많고, 모든 선택지를 시도하기에는 자원이 턱없이 부족하다. 이는 단순히 실험에 국한되지 않고, 로봇의 동작 방식이나 온라인 서비스의 최적화와 같은 다양한 영역에서 반복되는 공통된 고민이다. 우리는 제한된 자원 안에서 적은 시도로 최적에 답에 가까워지는 방법이 필요하다.

이러한 한계를 극복하기 위해 등장한 것이 '베이지안 최적화(Bayesian Optimization, BO)'이다. 베이지안 최적화는 확률적 추론을 통해 실험 데이터를 바탕으로 목적 함수의 형태와 불확실성을 추정하고, 이를 토대로 새로운 탐색 지점을 선택하여 빠르고 효율적으로 최적점을 찾을 수 있다. 그 핵심에는 목적 함수를 근사하는 확률적 대리 모델(Probabilistic Surrogate Model)과, 탐험(exploration)과 활용(exploitation) 사이의 균형을 맞추는 획득 함수(Acquisition Function)가 자리 잡고 있다.

이번 탐구 자료에서는 베이지안 최적화의 기본 원리와 핵심 구성요소를 이해하고, 다양한 응용 사례와 최신 연구 동향을 살펴본다. 이를 통해 단순히 알고리즘의 원리 이해를 넘어, 많은 불확실성과 제약이 있는 현실 속에서 어떻게 적은 비용으로 좋은 해답을 찾을 수 있을지에 대해 고찰한다. 이 과정에서 확률적 사고의 중요성에 대해 생각해 본다.

Fundamentals of Bayesian Optimization

Gaussian Process

베이지안 최적화에서 실제 목적 함수를 근사하는 데 가장 널리 사용되는 확률적 대리 모델은 Gaussian Process(GP)이다. GP는 함수 값에 대한 사전적인 기댓값을 나타내는 사전 평균 함수 $\mu(\cdot)$와 두 입력값 사이의 함수 값 유사성 또는 상관관계를 나타내는 공분산 함수 $\kappa(\cdot,\cdot)$로 특징지어진다.(Rasmussen, 2004) GP 모델은 n개의 데이터 쌍 $D_n = (X, y)$이 주어졌을 때, 관측된 출력이 가우시안 분포에서 추출된다고 가정한다. 관측된 출력 y는 실제 목적 함수 $f(X)$에 노이즈 ϵ이 더해진 상태, 즉 $y = f(X) + \epsilon$으로 모델링 되는데, 이때 노이즈 ϵ은 평균이 0이고 분산이 σ_ϵ^2인 정규분포 $N(0, \sigma_\epsilon^2)$를 따른다고 가정한다.

새로운 입력 x에 대한 예측된 출력과 관측된 출력 y를 동시에 고려한 결합 분포는 다음과 같이 다변수

가우시안 분포를 따른다.(Wang et al., 2023; Rasmussen, 2004)

$$\begin{bmatrix} y \\ \bar{y} \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X,X) + \sigma_\epsilon^2 I & K(X,x) \\ K(X,x)^T & \kappa(x,x) \end{bmatrix}\right) \tag{1}$$

다음은 각 항에 대한 자세한 설명이다.

- $\begin{bmatrix} y \\ \bar{y} \end{bmatrix}$: 관측된 출력값들과 예측할 출력값들을 합쳐서 전체 시스템의 출력 분포를 나타낸다.
- y: 이미 관측된 n개의 데이터 포인트에 대한 출력값들의 벡터이다.
- $\bar{y}$: 예측하고자 하는 새로운 입력 x에 해당하는 출력값이다.
- $\sim \mathcal{N}(0, \cdot)$: 해당 벡터가 정규 분포를 따름을 의미한다.
- $\begin{bmatrix} K(X,X) + \sigma_\epsilon^2 I & K(X,x) \\ K(X,x)^T & \kappa(x,x) \end{bmatrix}$: 공분산 행렬로, 모든 출력값들 간의 상관관계와 불확실성을 나타낸다.
- $K(X,X)$: 이미 관측된 n개의 입력 데이터에 대한 공분산 행렬로, 각 요소 $\kappa(x_i, x_j)$는 두 입력 x_i와 x_j 사이의 유사도를 나타내며, 일반적으로 커널 함수 $\kappa(\cdot,\cdot)$를 통해 계산된다.
- $\sigma_\epsilon^2 I$: 관측 노이즈 항으로 σ_ϵ^2는 관측 노이즈의 분산, I는 $n \times n$크기의 행렬을 나타낸다.
- $K(X,x)$: 관측된 입력데이터 X와 새로운 입력 데이터 x 사이의 공분산 벡터이다. X의 각 데이터 포인트와 x 사이의 유사도를 나타낸다.
- $K(X,x)^T$: $K(X,x)$의 행과 열이 바뀐 전치 행렬이다.
- $\kappa(x,x)$: 새로운 입력 데이터 x 자체의 공분산, 즉 새로운 입력 x에서 예측될 함수값 y의 분산이다.

이 분포는 베이즈 정리를 사용하여 관측된 데이터가 주어졌을 때, 새로운 입력 x에서의 함수값 y에 대한 조건부 분포를 계산하는 데 사용된다. 이 조건부 분포에서 얻어진 평균 $\mu(x)$와 분산 $\sigma^2(x)$은 각각 예측값과 예측의 불확실성을 나타내며, 다음 탐색할 지점을 결정하는 데 결정적인 역할을 한다.

이 분포에서 예측 평균과 분산은 다음과 같이 계산된다.(Wang et al., 2023)

$$\mu(x) = K(x,X)\big(K(X,X) + \sigma_\epsilon^2 I\big)^{-1} y \tag{2}$$

$$\sigma^2(x) = \kappa(x,x) - K(X,x)^T\big(K(X,X) + \sigma_\epsilon^2 I\big)^{-1} K(X,x) \tag{3}$$

GP 모델의 핵심은 공분산 함수 $\kappa(\cdot,\cdot)$이며, 이 함수는 두 입력 지점이 얼마나 유사한지를 정의하며, 커널 함수라고도 불린다. 일반적으로 사용되는 커널 함수로는 Squared Exponential (Gaussian) Kernel과 Matérn Kernel이 있다.(Frazier, 2018) Squared Exponential Kernel의 예시는 다음과 같다.(Wang et al., 2023)

$$\kappa_{\text{SE}}(x,x') = \sigma_f^2 exp(-\frac{(x-x')^2}{2l^2}) \tag{4}$$

이 예시에서 σ_f^2는 신호 분산을, l은 길이 스케일을 나타내는 변수이다. $l, \sigma_f, \sigma_\epsilon$와 같은 파라미터들을 하이퍼파라미터라고 부르고, $\theta = (l, \sigma_f, \sigma_\epsilon)$라고 표시한다. 일반적으로 최적의 하이퍼파라미터는 log marginal likelihood를 최대화하여 추정한다.(Wang et al., 2023)

$$\log p(y \mid X, \theta) = -\tfrac{1}{2} y^T K_{\bar{y}}^{-1} y - \tfrac{1}{2} \log|K_{\bar{y}}| - \tfrac{n}{2} \log 2\pi \tag{5}$$

여기서 $K_{\bar{y}} = K(X, X) + \sigma_f^2 I$이다.

Acquisition Function

Acquisition Function(AF)은 다음에 샘플링할 위치를 결정하여 목적 함수의 최적점에 도달하도록 검색을 안내하는 유틸리티 함수이다. AF의 기본 원리는 GP에서 얻은 불확실성(분산)과 최적성(예측 평균) 정보를 사용하여 탐험(exploration)과 활용(exploitation) 사이의 적절한 균형을 맞추는 것이다. 탐험은 아직 충분히 평가되지 않았거나 불확실성이 큰 영역을 탐색하고, 활용은 이미 높은 성능을 보였거나 최적일 가능성이 높은 영역을 탐색한다. 주요 AF들은 다음과 같다.(Wang et al., 2023)

1. Probability of Improvement(PI): 현재까지의 최고 값 f^*보다 더 나은 값을 찾을 확률을 최대화한다.

$$PI(x; D_n) = P(f(x) \geq f^*) = \Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) \tag{6}$$

2. Expected Improvement(EI): f^*에 대한 예상 개선값을 계산한다.

$$EI(x; D_n) = \mathbb{E}[\max(0, f(x) - f^*)] = \big(\mu(x) - f^*\big)\Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) + \sigma(x)\,\phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) \tag{7}$$

3. Upper Confidence Bound(UCB): 불확실성과 예상 보상을 결합하여 최적의 회귀를 달성한다.

$$UCB(x; D_n) = \mu(x) + \beta\sigma(x) \tag{8}$$

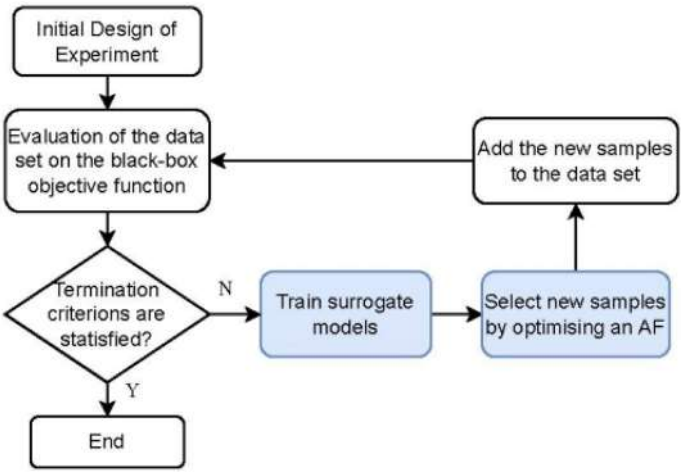


Figure 1. 베이지안 최적화의 흐름 (Wang et al., 2023)

정리하면, 베이지안 최적화는 목적 함수에 대한 사전 분포를 설정하고, 새로운 데이터를 관측한 뒤 이를 반영해 업데이트된 분포를 얻는다. 이렇게 해서 함수에 대한 불확실한 신념을 점차 정교하게 바꿔 나간다. 이때 GP와 같은 확률적 대리 모델이 목적 함수의 분포를 근사하고, 획득 함수(Acquisition Function) 가 이 분포를 활용해 다음 탐색 지점을 제안한다. 새로 선택된 지점은 실제 목적 함수를 통해 평가되며, 이 결과는 다시 데이터셋에 추가되어 모델을 재학습한다. 이러한 "모델링 → 탐색 → 평가 → 갱신"의 과정이 반복되면서, BO는 점점 더 효율적으로 최적점에 다가간다. Figure 1은 이 과정을 직관적으로 보여주는 흐름도로, 베이지안 최적화의 전형적인 흐름을 한눈에 제시한다.

Applications of Bayesian Optimization

앞서 살펴본 가우시안 프로세스 기반의 확률적 대리 모델과 탐험–활용 균형을 맞추는 획득 함수는 다양한 실제 문제에 적용될 수 있다. 아래에서는 대표적인 응용 분야들을 통해 BO가 현실 문제 해결에 어떻게 기여하는지를 살펴본다.

A/B Testing

온라인 광고, 웹사이트 디자인, 앱 기능 비교와 같은 A/B 테스팅에서는 사용자의 반응을 관찰하여 어느 버전이 더 효과적인지를 판단한다. 그러나 무한정 실험을 진행하기에는 시간과 돈(쿼리 예산)이 한정적이다. 베이지안 최적화는 GP를 이용해 사용자 반응의 불확실성까지 고려하여 모델링하고, 획득 함수를 통해 다음에 실험할 버전을 전략적으로 선택한다(Shahriari et al., 2016). 이를 통해 제한된 자원 내에서도 기회 비용을 최소화하여 최적화된 버전을 예측할 수 있다.

Robotics and Reinforcement Learning

로봇 공학에서 베이지안 최적화는 로봇 동작 방식을 결정하는 '정책'을 탐색하고 최적화하는 데 사용된다. 예를 들어, Sony AIBO ERS-7 로봇의 보행 매개변수를 베이지안 최적화를 통해 조정하여 로봇의 걷는 속도나 움직임의 부드러움을 최적화할 수 있다. 또한, 로봇이 길을 찾을 때 자신의 위치에 대한 불확실성을 최소화할 수 있다(Lizotte, n.d.). 심지어 복잡한 AI 시스템에서는 신경망의 성능을 자동으로 높이거나, 이미지에서 중요한 부분에 주목하는 정책을 배우도록 돕는다(Shahriari et al., 2016).

Environmental Monitoring and Sensor Network

환경 모니터링에서는 제한된 센서를 활용하여 온도, 대기, 토양, 해양 등에서의 오염 물질 농도를 추정해야 하는 경우가 많다. 이러한 센서들은 자기 주변의 정보(부분적이고, 노이즈가 있는 정보)를 제공하는데, 모든 센서를 동시에 가동하기에는 비용이 크기 때문에 (Shahriari et al., 2016), BO는 글로벌 모델을 모델링하고, 어느 센서를 활성화할지를 결정한다. 그 결과 소수의 센서만 사용해도 건물 내 최저/최고 온도 지점이나 교통 혼잡 지역과 같은 중요한 정보를 효과적으로 추출할 수 있다(Srinivas et al., 2012).

Automatic Machine Learning and Hyperparameter Tuning

머신러닝 분야에서 베이즈 최적화는 인공지능 모델을 만드는 과정을 자동화하고 최적화하는 데 사용된다. 인공지능 학습에는 어떤 모델을 쓸지(예: Neural Network, Random Forests 등)와 그 모델의 세부 설정 값(하이퍼파라미터)을 정하는 것이 매우 중요한데, 이 모든 조합을 하나씩 테스트하는 것은 많은 시간과 비용이 필요하다. 베이지안 최적화는 주어진 데이터 셋에서 작업을 해결하기 위해 최상의 모델과 관련 하이퍼파라미터를 자동으로 선택하는 것을 목표로 사용될 수 있다. 이 방식은 deep belief networks, convolutional neural networks, MCMC 등 다양한 모델의 하이퍼파라미터 튜닝에서 성공적으로 사용되어(Shahriari et al., 2016), 적은 자원으로도 높은 성능의 모델을 찾을 수 있게 한다.

Recent Advances of Bayesian Optimization

베이지안 최적화는 '불확실성이 크고 실험 비용이 높은 상황에서 효율적으로 탐색한다'는 장점을 바탕으로, 온라인 실험, 로봇 공학, 환경 모니터링, 머신러닝 등 다양한 분야에서 응용되고 있다. 그러나 실제 응용 과정에서는 차원이 커지거나 데이터에 잡음이 섞이는 등 새로운 도전 과제가 등장한다. 이러한 어려움을 해결하기 위해 다양한 연구가 진행되고 있다. 이번 파트에서는 (Wang et al., 2023)의 종합 리뷰를 참고하여 주요 연구 동향을 살펴본다.

High-Dimensional Optimization

베이지안 최적화는 저차원 문제에서는 성공적으로 적용되지만, 탐색 공간의 차원이 10–20을 넘어가면 성능이 저하된다. GP와 같이 데이터가 특정한 분포를 따른다는 가정을 하지 않는 모델은 차원이 증가함에 따라 탐색 공간이 기하급수적으로 커지기 때문에 모델 학습이 매우 어려워진다. 고차원 공간에서 신뢰할 수 있는 GP 모델을 구축하기 위해 더 많은 관측 데이터가 필요하지만, GP는 데이터 크기에 대해 $\mathcal{O}(N^3)$의 계산 복잡성을 가지므로 확장성 문제가 발생한다. 이러한 어려움을 극복하기 위해 고차원 목적 함수가 실제로는 저차원 부분 공간 내에서만 크게 변화하며, 실제 목적 함수가 몇 개의 저차원 함수의 합으로 표현될 수 있다는 가정에 기반한 방법들이 연구되고 있다. 대규모 데이터의 경우 GP 대신 Bayesian Neural Networks, Random Forests, Dropout Deep Neural Networks와 같이 더 확장 가능하고 유연한 모델을 사용하고, 지역 모델링(Local Modeling) 및 배치 선택(Batch Selection) 방식이 사용된다.

Combinational Optimization

정수, 집합, 시퀀스, 범주형, 그래프 등의 조합 공간에서의 최적화는 커널 및 거리 측정 정의의 어려움과 기울기 기반 AF 최적화 방법 적용의 어려움, 가능한 해의 수가 변수의 수에 따라 기하급수적으로 증가하여 탐색 공간이 매우 커지는 조합 폭발 현상에 의해 어려움을 겪는다. 이 어려움을 극복하기 위해 GP 대신 이산 데이터에 더 적합한 모델(Neural Network, Random Forests, Tree-structured Parzen estimator 등)을 사용하거나, 커널 계산의 거리 측정 방식을 수정하거나, 연속 변수와 이산 변수가 모두 포함된 혼합 탐색 공간을 위한 커널 조합이 연구되고 있다.

Nosiy and Robust Optimization

베이지안 최적화에서는 GP를 구성하기 위해 입력 지점의 측정값에는 노이즈가 없으며, 관측값의 노이즈는 등분산 정규분포를 따른다고 가정한다. 하지만 실제 문제에서는 이러한 가정이 유지되지 않아 최적화 성능이 저하될 수 있다. 출력 노이즈를 다루기 위해 노이즈를 고려한 새로운 획득 함수를 설계하는 데 중심을 두는데, EI를 노이즈 관측에 맞게 수정하거나 정보 기반 AF(Predictive Entropy Search, Thompson Sampling 등)를 이용한다. 이상치에는 가우시안 노이즈 대신 Student-t 분포와 같이 꼬리가 두꺼운 노이즈 모델을 사용하는 강건 GP를 사용한다. 입력 노이즈는 이분산성 GP(heteroscedastic GP)나 NIGP(noisy input GP), unscented EI와 같은 기법으로 처리된다.

Multi-Fidelity Optimization

목적 함수를 다양한 비용으로 여러 수준의 충실도에서 평가할 수 있을 때, 모든 충실도 모델에서 최대량의 정보를 공동으로 학습하여 최적화 속도를 높이고 최적화 비용을 줄일 수 있다. MFO는 여러 충실도 데이터를 활용하기 위해 다중 충실도 모델을 구축하고 다중 충실도 최적화를 위한 AF를 설계한다. 다중 충실도 모델에는 각 충실도 수준에 대해 독립적인 GP를 학습하는 방법인 개별 GP 모델링, 여러 충실도 데이터를 공동으로 모델링하여 다른 충실도 데이터 간의 상관관계를 포착하는 Co-Kriging 모델이 있다. 최근 다중 충실도 최적화를 위한 AF에서는 ES(Entropy Search), PES(Predictive Entropy Search), MES(Max-value Entropy Search)와 같은 정보 이론 기반 접근 방식이 인기를 얻고 있다.

Challenges and Future Directions

베이지안 최적화는 비용이 많이 드는 블랙박스 최적화 문제를 효율적으로 해결하는 강력한 최적화 방법이며, 다양한 실제 문제에 성공적으로 적용되어 왔다. 그러나 이러한 발전에도 불구하고 아직 수많은 과제가 남아있는데, 이는 과학 기술의 새로운 응용 분야들이 새로운 과제와 요구 사항을 제기하고 있기 때문이다.

Distributed BO

분산 BO는 탐색 공간, 샘플링 과정, 비용이 많이 드는 평가 및 GP를 분산하여 처리하기 위해 등장했다. 예를 들어, HyperSpace는 넓은 탐색 공간을 일정 부분 중첩되도록 분할하고, 이러한 HyperSpace의 가능한 모든 조합에 대해 GP 모델을 갖추게 하여 최적화 루프를 병렬로 실행할 수 있게 했다.

분산 BO 설계에는 세 가지 주요 과제가 있다. 첫째, 수렴 속도와 통신 비용 간의 균형을 달성하는 것이 중요하다. 둘째, 시간 가변적인 통신 비용, 각기 다른 계산 능력, 이기종적인 평가 시간 등으로 인해 발생하는 비동기 설정을 처리하는 연구는 아직 활발히 연구되지 않았다. 마지막으로, 복잡한 통신 네트워크나 통신 제약 조건 등 실제 환경에서 발생할 수 있는 실용적인 시나리오를 고려한 연구가 필요하다(Wang et al., 2023).

Dynamic Optimization

네트워크 자원 할당, 추천 시스템, 객체 추적 등과 같은 실제 사례에서는 최적화해야 할 목적 함수가 시간이 지남에 따라 변경될 수 있는데, 이러한 시나리오를 '동적 최적화'또는 '시간 의존적 문제'라고 한다. 동적 최적화를 위한 대부분의 BO 방법은 부분적인 정보를 이용한 순차적 의사 결정을 모델링하여 누적 보상을 최대화하는 것을 목표로 하는 MAB(Multi-Armed Bandit) 설정을 기반으로 한다(Zhao et al., 2021).

그러나, 시간 의존적 목적 함수를 위한 효과적인 대리 모델을 구축하고 유망한 해답을 식별하기 위한 AF 설계 및 최적점을 추적은 여전히 어려운 문제이다. 또한, 전이 학습과 같은 머신러닝의 발전을 활용하여 이전 실행에서 얻은 유용한 정보를 활용하는 것도 흥미로운 연구 방향 중 하나이다(Wang et al., 2023).

Algorithmic Fairness

머신러닝 알고리즘이 다양한 분야에 광범위하게 사용되면서, 알고리즘의 공정성에 대한 우려가 커지고 있다. 일반적인 머신러닝 분야에서 사용하는 불공정성 완화 기법은 모델을 훈련하기 전 데이터 분포의 균형을 다시 맞추거나, 공정성 제약 조건 하에 모델을 훈련을 하거나 정확도 지표와 공정성을 결합하여 모델을 만들거나, 훈련이 끝난 후 모델을 조정하는 방법을 사용한다.

베이지안 최적화에서의 공정성을 고려한 기존의 연구 사례로는 FairBO와 협업 BO가 있다(Perrone et al., 2021). 하지만, 베이지안 프레임워크에서 공정성을 고려하는 것은 아직 거의 탐구되지 않은 영역이다. 베이지안 최적화 맥락에서 공정성 측정 및 수학적 정의가 아직 명확하지 않아 공정성 요구 사항을 베이지안 최적화에 더 정확하게 통합하기 위한 명확한 정의가 필요하다. 그리고 대리 모델이 얼마나 고정한지, 선택된 새로운 샘플이 얼마나 공정한지에 대한 조사가 부족하다. 마지막으로, 현재의 편향 감소 전략은 단일 공정성 정의에만 적용될 수 있는 경우가 많으므로(Wang et al., 2023), 실제 환경에서 다양한 공정성 정의를 포괄하는 실용적인 공정성 인식 방법 설계가 중요하다.

Conclusion

베이지안 최적화는 불확실성과 자원 제약 속에서도 최적의 해답에 다가가기 위한 강력한 방법이다. 전통적인 실험 설계나 단순한 무작위 탐색은 한계가 있다. 베이지안 최적화는 확률적 대리 모델과 획득

함수를 통해 매번의 실험에서 얻은 정보를 최대한 활용하며 탐색 경로를 끊임없이 갱신한다.

실제 사례에서도 베이지안 최적화의 유용성은 뚜렷하게 드러난다. 인터넷 기업의 A/B 테스팅에서는 사용자 반응을 효율적으로 비교하고, 로보틱스에서는 로봇의 보행과 제어 전략을 빠르게 최적화하며, 환경 모니터링에서는 제한된 센서만으로도 넓은 영역의 상태를 추정할 수 있다. 또한 머신러닝 분야에서는 하이퍼파라미터 최적화를 통해 복잡한 모델의 성능을 크게 끌어올리는 데 활용되고 있다. 이러한 예시는 베이지안 최적화가 단순한 이론적 도구를 넘어, 실제 세계의 복잡한 문제들을 해결하는 데 중요한 역할을 하고 있음을 보여준다.

그러나 동시에, 베이지안 최적화가 앞으로 극복해야 할 과제도 많다. 탐색 공간의 차원이 커지면서 발생하는 고차원 문제, 이산적이고 조합적인 변수들로 인해 탐색 공간이 폭발적으로 커지는 조합 최적화 문제, 관측 노이즈나 이상치로 인해 불확실성이 커지는 강건 최적화 문제, 그리고 비용과 정확도가 다른 데이터들을 함께 활용해야 하는 다중 충실도 최적화 문제는 여전히 활발히 연구되는 영역이다. 더 나아가, 분산 시스템에서의 효율적 탐색, 시간에 따라 변하는 동적 환경에서의 적응, 그리고 알고리즘 공정성과 같은 사회적 요구까지 고려한다면, 베이지안 최적화의 잠재적 확장성은 무궁무진하다.

결국 베이지안 최적화는 단순한 최적화 기법을 넘어, 불확실성과 제약 속에서도 효율적으로 선택하고 학습하는 방법을 제시한다는 점에서 큰 의미가 있다. 이번 탐구에서 우리는 BO의 원리와 응용, 그리고 미래 과제를 살펴보았다. 앞으로도 이 방법은 다양한 분야에서 계속 확장되며 활용될 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

Q.베이지안 최적화는 가우시안 프로세스(GP)와 같은 확률적 대리 모델로 사용하여, 불확실성이 큰 영역을 탐험(exploration)하면서 동시에 좋은 성능을 보이는 영역을 활용(exploitation)한다. 추가자료로 제시된 파이썬 예제 코드를 실행해 결과를 관찰한 뒤, 다음 단계를 따라 분석하고 응용해 보자.

1. **주어진 함수** $f(x)$**의 형태를 확인하고, 어떤 구간이 예측이 어려울지 예상한 뒤 코드를 실행해 결과를 확인하고, 실제 함수 그래프와 비교했을 때 얼마나 정확히 최적점을 찾았는지 분석한다.**

2. **초기 샘플링 개수를** ***10, 20, 30*****으로 변경해 실행한 뒤, 신뢰 구간의 폭과 예측 함수가 어떻게 변화하는지 분석한다.**

3. **커널 함수의** ***length scale*****을** ***0.05, 1.0*** **등으로 변경해 실행한 뒤, 신뢰 구간의 폭과 예측 함수가 어떻게 변화하는지 분석한다.**

4. Gaussian Process**의 커널을** ***RBF, WhiteKernel, CompundKernel*** **등으로 변경해 실행한 뒤, 신뢰 구간의 폭과 예측 함수가 어떻게 변화하는지 분석한다.**

5. **본문에 제시된 다양한 적용 사례*****(A/B Testing, Robotics and Reinforcement Learning, Environmental Monitoring and Sensor Network, Automatic Machine Learning and Hyperparameter Tuning)*** **중 하나를 선택하고, 어떤 최적화 목표와 하이퍼파라미터를 설정할지 설계한다. 이때 베이지안 최적화를 적용했을 때 기대되는 효율성(예: 탐색 방식, 샘플 효율성 등)과 기술적·계산적 측면에서 나타날 수 있는 한계점(예.고차원 공간에서의 성능 저하, 범주형 변수 처리의 어려움 등)을 이유와 함께 분석한다.**

References

Frazier, P. I. 2018. A Tutorial on Bayesian Optimization. arXiv preprint arXiv:1807.02811 URL http://arxiv.org/abs/1807.02811.

Lizotte, D. n.d. Automatic Gait Optimization with Gaussian Process Regression. Unpublished manuscript .

Perrone, V., M. Donini, M. B. Zafar, R. Schmucker, K. Kenthapadi, and C. Archambeau, 2021. Fair Bayesian Optimization. Pages 854–863 *in* Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21). ACM.

Rasmussen, C. E., 2004. Gaussian Processes in Machine Learning. Pages 63–71 *in* O. Bousquet, U. V. Luxburg, and G. Rätsch, editors. Advanced Lectures on Machine Learning. Springer.

Shahriari, B., K. Swersky, Z. Wang, R. P. Adams, and N. D. Freitas. 2016. Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE **104**:148–175.

Srinivas, N., A. Krause, S. M. Kakade, and M. Seeger. 2012. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. IEEE Transactions on Information Theory **58**:3250–3265.

Wang, X., Y. Jin, S. Schmitt, and M. Olhofer. 2023. Recent Advances in Bayesian Optimization. ACM Computing Surveys **55**:1–36.

Zhao, P., L. Zhang, Y. Jiang, and Z.-H. Zhou. 2021. A Simple Approach for Non-stationary Linear Bandits. arXiv preprint arXiv:2103.05324 URL http://arxiv.org/abs/2103.05324.

Contributors



Taegyeong Kim

M. ktg0829@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Energy AI and Environmental Climate Technology. Belonging Meta Lab, as an undergraduate researcher, I am working on catalyst optimization and novel catalyst development using machine learning. I aim to save energy and contribute to the development of a catalyst for removing harmful substances by greatly reducing the experimental cost and time required to develop a catalyst.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Sumyeong Ahn

M. sumyeongahn@kentech.ac.kr

Sumyeong Ahn is an assistant professor in the Energy AI track within the Department of Energy Engineering at the Korea Institute of Energy Technology. His research interests include robust AI, trustworthy AI, and foundation models, such as large language models and large multimodal models. Recently, his focus has been on developing robust fine-tuning and inference algorithms. Beyond AI itself, he also focuses on AI for X -- leveraging AI for real-world applications, including energy system, medical system and other critical applications.

KENTECH
Korea Institute of Energy Technology

Bayesian Optimization: Unlocking the World of Optimization Through Probability

Abstract

Bayesian Optimization (BO) is a method for efficiently exploring real-world problems. This document explains the process of finding the optimal point by balancing exploration and exploitation, based on the fundamental principles of BO. It also examines its practical applicability and future challenges through various application cases. Through this, we consider that Bayesian Optimization is not merely a technique, but a statistical method that solves optimization problems based on probabilistic thinking within uncertain realities.

Introduction

Many modern problems share the common challenge of finding the optimal solution within an enormous search space. The complexity of these problems stems from the need to identify a single optimal solution among nearly infinite possibilities. For example, consider discovering a new catalyst. While countless elements and combinations are theoretically possible, experimentally testing every single case is impossible. The number of possible choices is astronomical, and resources are too scarce to attempt every single one. This challenge isn't limited to experiments; it's a recurring concern across diverse fields, such as determining how a robot should move or optimizing online services. We need a method to get closer to the optimal solution with fewer attempts, within our limited resources.

To overcome these limitations, 'Bayesian Optimization (BO)' was developed. Bayesian Optimization uses probabilistic inference to estimate the form and uncertainty of the objective function based on experimental data. It then selects new exploration points based on this estimation to find the optimal point quickly and efficiently. At its core are the Probabilistic Surrogate Model, which approximates the objective function, and the Acquisition Function, which balances exploration and exploitation.

This research material explores the fundamental principles and core components of Bayesian optimization, examining various application cases and the latest research trends. It goes beyond merely understanding the algorithm's principles to contemplate how to find good solutions at low cost within the reality of significant uncertainty and constraints. This process prompts reflection on the importance of probabilistic thinking.

Fundamentals of Bayesian Optimization

Gaussian Process

In Bayesian optimization, the most widely used probabilistic surrogate model for approximating the true objective function is the Gaussian Process (GP). A GP is characterized by a prior mean function $\mu(\cdot)$, representing the prior expectation of the function value, and a covariance function $\kappa(\cdot,\cdot)$, representing the similarity or correlation of function values between two inputs.(Rasmussen, 2004) Given n data pairs $D_n = (X, y)$, the GP

model assumes that the observed output is drawn from a Gaussian distribution. The observed output y is modeled as the actual objective function $f(X)$ plus noise ϵ, i.e., $y = f(X) + \epsilon$. Here, the noise ϵ is assumed to follow a normal distribution $N(0, \sigma_\epsilon^2)$ with mean 0 and variance σ_ϵ^2.

The joint distribution considering both the predicted output for a new input x and the observed output y simultaneously follows a multivariate Gaussian distribution as follows.(Wang et al., 2023; Rasmussen, 2004)

$$\begin{bmatrix} y \\ \bar{y} \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} K(X,X) + \sigma_\epsilon^2 I & K(X,x) \\ K(X,x)^T & \kappa(x,x) \end{bmatrix}\right) \tag{1}$$

The following is a detailed explanation of each item.

- $\begin{bmatrix} y \\ \bar{y} \end{bmatrix}$: Represents the output distribution of the entire system by combining the observed output values and the predicted output values.
- y: The vector of output values for the n data points already observed.
- $\bar{y}$: The output value corresponding to the new input x to be predicted.
- $\sim \mathcal{N}(0, \cdot)$: Indicates that the vector follows a normal distribution.
- $\begin{bmatrix} K(X,X) + \sigma_\epsilon^2 I & K(X,x) \\ K(X,x)^T & \kappa(x,x) \end{bmatrix}$: The covariance matrix, representing the correlation and uncertainty between all output values.
- $K(X,X)$: The covariance matrix for the n observed input data points. Each element $\kappa(x_i, x_j)$ represents the similarity between two inputs x_i and x_j and is typically computed via a kernel function $\kappa(\cdot,\cdot)$.
- $\sigma_\epsilon^2 I$: The observation noise term, where σ_ϵ^2 is the variance of the observation noise and I denotes an $n \times n$ matrix.
- $K(X,x)$: The covariance vector between the observed input data X and the new input data x. It represents the similarity between each data point in X and x.
- $K(X,x)^T$: The transpose of $K(X,x)$, with rows and columns swapped.
- $\kappa(x,x)$: The covariance of the new input data x itself, i.e., the variance of the predicted function value y from the new input x.

This distribution is used to calculate the conditional distribution of the function value y at a new input x given the observed data, using Bayes' theorem. The mean $\mu(x)$ and variance $\sigma^2(x)$ obtained from this conditional distribution represent the prediction and the uncertainty of the prediction, respectively, and play a crucial role in determining the next point to explore.

The predicted mean and variance from this distribution are calculated as follows.(Wang et al., 2023)

$$\mu(x) = K(x,X)\big(K(X,X) + \sigma_\epsilon^2 I\big)^{-1} y \tag{2}$$

$$\sigma^2(x) = \kappa(x,x) - K(X,x)^T\left(K(X,X) + \sigma_\epsilon^2 I\right)^{-1} K(X,x) \tag{3}$$

The core of the GP model is the covariance function $\kappa(\cdot,\cdot)$, which defines how similar two input points are and is also called the kernel function. Commonly used kernel functions include the Squared Exponential (Gaussian) Kernel and the Matérn Kernel.(Frazier, 2018) An example of the Squared Exponential Kernel is as follows.(Wang et al., 2023)

$$\kappa_{\mathrm{SE}}(x,x') = \sigma_f^2 exp(-\frac{(x-x')^2}{2l^2}) \tag{4}$$

In this example, σ_f^2 represents the signal variance, and l is a variable indicating the length scale. Parameters such as l, σ_f, and σ_ϵ are called hyperparameters and denoted as $\theta = (l, \sigma_f, \sigma_\epsilon)$. Typically, optimal hyperparameters are estimated by maximizing the log marginal likelihood.(Wang et al., 2023)

$$\log p(y \mid X, \theta) = -\tfrac{1}{2} y^T K_{\bar{y}}^{-1} y - \tfrac{1}{2}\log|K_{\bar{y}}| - \tfrac{n}{2}\log 2\pi \tag{5}$$

Here, $K_{\bar{y}} = K(X,X) + \sigma_f^2 I$.

Acquisition Function

The Acquisition Function (AF) is a utility function that guides the search toward the optimal point of the objective function by determining the next location to sample. The fundamental principle of the AF is to balance exploration and exploitation by utilizing uncertainty (variance) and optimality (predicted mean) information obtained from the GP. Exploration searches areas that have not yet been sufficiently evaluated or where uncertainty is high, while exploitation searches areas that have already shown high performance or are likely to be optimal. The main AFs are as follows.(Wang et al., 2023)

1. Probability of Improvement(PI): Maximizes the probability of finding a value better than the current best value f^*.
$$PI(x; D_n) = P(f(x) \geq f^*) = \Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) \tag{6}$$
2. Expected Improvement(EI): Calculate the expected improvement value for f^*.
$$EI(x; D_n) = \mathbb{E}[\max(0, f(x) - f^*)] = \left(\mu(x) - f^*\right)\Phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) + \sigma(x)\,\phi\left(\frac{\mu(x) - f^*}{\sigma(x)}\right) \tag{7}$$
3. Upper Confidence Bound(UCB): Combining uncertainty and expected rewards achieves optimal regression.
$$UCB(x; D_n) = \mu(x) + \beta\sigma(x) \tag{8}$$

To summarize, Bayesian optimization sets a prior distribution for the objective function, observes new data, and then updates this distribution to reflect the new information. This process gradually refines the uncertain beliefs about the function. Here, a probabilistic surrogate model like a GP approximates the distribution of the objective function, and an acquisition function utilizes this distribution to suggest the next exploration point.

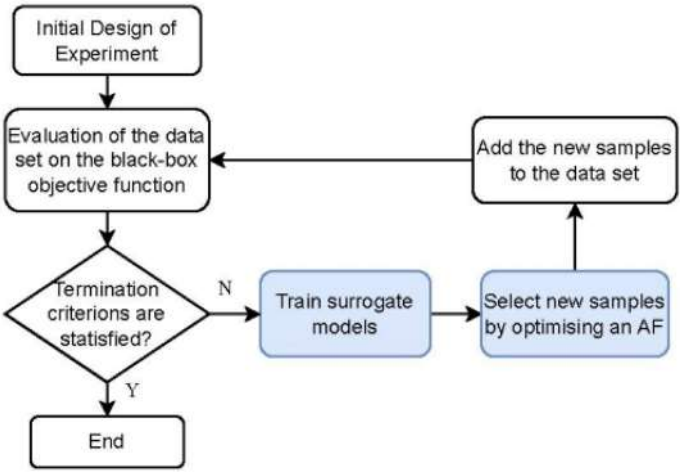


Figure 1. Flowchart of Bayesian Optimization (Wang et al., 2023)

The newly selected point is evaluated using the actual objective function, and this result is added back to the dataset to retrain the model. Through the iterative process of "Modeling → Exploration → Evaluation → Update," BO progressively approaches the optimal point more efficiently. Figure 1 is a flowchart that intuitively illustrates this process, presenting the typical flow of Bayesian optimization at a glance.

Applications of Bayesian Optimization

The Gaussian process-based probabilistic surrogate model and the acquisition function that balances exploration and exploitation, as discussed earlier, can be applied to various real-world problems. Below, we examine how BO contributes to solving real-world problems through representative application areas.

A/B Testing

In A/B testing for online ads, website design, and app feature comparisons, user responses are observed to determine which version is more effective. However, time and money (query budget) are limited for conducting experiments indefinitely. Bayesian optimization models user response uncertainty using GP and strategically selects the next version to test via an acquisition function(Shahriari et al., 2016). This enables predicting the optimized version while minimizing opportunity costs within constrained resources.

Robotics and Reinforcement Learning

In robotics, Bayesian optimization is used to explore and optimize the 'policy' that determines a robot's behavior. For example, the walking parameters of the Sony AIBO ERS-7 robot can be adjusted through Bayesian optimization to optimize the robot's walking speed or the smoothness of its movements. Furthermore, it can minimize uncertainty about a robot's position when navigating(Lizotte, n.d.). Even in complex AI systems, it helps automatically improve neural network performance or learn policies that focus on important parts of images(Shahriari et al., 2016).

Environmental Monitoring and Sensor Network

In environmental monitoring, it is often necessary to estimate pollutant concentrations in temperature, air, soil, and marine environments using a limited number of sensors. These sensors provide information about their

local surroundings (partial and noisy data). Since operating all sensors simultaneously is costly (Shahriari et al., 2016), BO models the global distribution and decides which sensors to activate. As a result, even with only a few sensors, critical information such as the minimum/maximum temperature points within a building or areas of heavy traffic congestion can be effectively extracted(Srinivas et al., 2012).

Automatic Machine Learning and Hyperparameter Tuning

In the field of machine learning, Bayesian optimization is used to automate and optimize the process of creating artificial intelligence models. In AI training, determining which model to use (e.g., Neural Network, Random Forests, etc.) and its detailed configuration values (hyperparameters) is crucial. Testing all possible combinations one by one requires significant time and cost. Bayesian optimization can be employed to automatically select the optimal model and its associated hyperparameters for solving a task given a dataset. This approach has been successfully applied to hyperparameter tuning for various models, including deep belief networks, convolutional neural networks, and MCMC(Shahriari et al., 2016), enabling the discovery of high-performance models even with limited resources.

Recent Advances of Bayesian Optimization

Bayesian optimization is applied in various fields such as online experiments, robotics, environmental monitoring, and machine learning, based on its advantage of 'efficiently exploring in situations with high uncertainty and high experimental costs'. However, new challenges arise in actual application processes, such as increasing dimensionality or noise in the data. Various studies are being conducted to overcome these difficulties. This part reviews key research trends by referencing the comprehensive review in (Wang et al., 2023).

High-Dimensional Optimization

Bayesian optimization performs well for low-dimensional problems, but its performance degrades when the search space dimension exceeds 10–20. Models like GPs, which do not assume data follows a specific distribution, become extremely difficult to train as the search space grows exponentially with increasing dimension. Building reliable GP models in high-dimensional spaces requires more observational data, but GPs exhibit computational complexity of $\mathcal{O}(N^3)$ with respect to data size, leading to scalability issues. To overcome these difficulties, methods are being researched based on the assumption that the high-dimensional objective function actually varies significantly only within low-dimensional subspaces, and that the true objective function can be expressed as the sum of several low-dimensional functions. For large-scale data, more scalable and flexible models such as Bayesian Neural Networks, Random Forests, and Dropout Deep Neural Networks are used instead of GPs, employing techniques like Local Modeling and Batch Selection.

Combinational Optimization

Optimization in combinatorial spaces such as integers, sets, sequences, categories, and graphs faces challenges due to the difficulty in defining kernels and distance measures, the difficulty in applying gradient-based AF optimization methods, and combinatorial explosion, where the number of possible solutions increases exponentially with the number of variables, leading to an extremely large search space. To overcome these challenges, research is being conducted on using models more suitable for discrete data (such as Neural Networks, Random Forests, Tree-structured Parzen estimators) instead of GP, modifying the distance metric used in kernel computation, or developing kernel combinations for mixed search spaces containing both continuous and discrete variables.

Nosiy and Robust Optimization

In Bayesian optimization, it is assumed that measurements at input points are noise-free and that observation noise follows an isosvarian normal distribution when constructing the GP. However, in real-world problems, these assumptions may not hold, potentially degrading optimization performance. The focus is on designing new acquisition functions that account for noise to handle output noise, either by modifying EI to fit noisy observations or by using information-based AF (Predictive Entropy Search, Thompson Sampling, etc.). For outliers, robust GPs are used that employ fat-tailed noise models, such as the Student-t distribution, instead of Gaussian noise. Input noise is handled using techniques like heteroscedastic GPs (heteroscedastic GP), noisy input GPs (NIGP), or unscented EI.

Multi-Fidelity Optimization

When the objective function can be evaluated at multiple fidelity levels with varying costs, the maximum amount of information can be jointly learned across all fidelity models to accelerate optimization speed and reduce optimization costs. MFO constructs multi-fidelity models to utilize multiple fidelity data and designs an AF for multi-fidelity optimization. Multi-fidelity models include individual GP modeling, which trains independent GPs for each fidelity level, and Co-Kriging models, which jointly model multiple fidelity datasets to capture correlations between different fidelity data. Recently, information theory-based approaches such as Entropy Search (ES), Predictive Entropy Search (PES), and Max-value Entropy Search (MES) have gained popularity for AF in multi-fidelity optimization.

Challenges and Future Directions

Bayesian optimization is a powerful optimization method that efficiently solves costly black-box optimization problems and has been successfully applied to various real-world problems. However, despite these advances, numerous challenges remain, as new applications in science and technology continue to present fresh challenges and requirements.

Distributed BO

Distributed BO emerged to distribute processing across the search space, sampling process, costly evaluations, and GP. For example, HyperSpace partitions a large search space into overlapping segments and equips GP models for all possible combinations of these HyperSpaces, enabling parallel optimization loops.

Distributed BO design faces three main challenges. First, achieving a balance between convergence speed and communication cost is crucial. Second, research addressing asynchronous settings arising from time-varying communication costs, differing computational capabilities, and heterogeneous evaluation times has not yet been extensively explored. Finally, studies considering practical scenarios that may occur in real-world environments, such as complex communication networks or communication constraints, are needed(Wang et al., 2023).

Dynamic Optimization

In practical applications such as network resource allocation, recommendation systems, and object tracking, the objective function to be optimized may change over time. Such scenarios are referred to as ‘dynamic optimization’ or ‘time-dependent problems’. Most BO methods for dynamic optimization are based on the Multi-Armed Bandit (MAB) setting, which models sequential decision-making with partial information and aims to maximize cumulative reward(Zhao et al., 2021).

However, constructing effective surrogate models for time-dependent objective functions, designing effective action-frequency (AF) schedules to identify promising solutions, and tracking optimal points remain challenging problems. Furthermore, leveraging advances in machine learning, such as transfer learning, to utilize useful information gained from previous runs is also an interesting research direction(Wang et al., 2023).

Algorithmic Fairness

As machine learning algorithms are widely used across various fields, concerns about algorithmic fairness are growing. Common fairness mitigation techniques in general machine learning involve rebalancing the data distribution before training the model, training the model under fairness constraints, combining accuracy metrics with fairness to build the model, or adjusting the model after training.

Existing research examples considering fairness in Bayesian optimization include FairBO and collaborative BO(Perrone et al., 2021). However, incorporating fairness within the Bayesian framework remains largely unexplored. Fairness metrics and mathematical definitions in the Bayesian optimization context are still unclear, necessitating precise definitions to integrate fairness requirements more accurately into Bayesian optimization. Furthermore, there is insufficient investigation into how fixed the surrogate model is and how fair the selected new samples are. Finally, current bias reduction strategies are often only applicable to a single fairness definition(Wang et al., 2023), making it crucial to design practical fairness-aware methods that encompass diverse fairness definitions in real-world settings.

Conclusion

Bayesian optimization is a powerful method for approaching optimal solutions even amidst uncertainty and resource constraints. Traditional experimental design or simple random exploration has limitations. Bayesian optimization maximizes the use of information gained from each experiment through probabilistic surrogate models and acquisition functions, continuously updating the exploration path.

The utility of Bayesian optimization is clearly evident in real-world applications. In internet companies' A/B testing, it efficiently compares user responses; in robotics, it rapidly optimizes robot walking and control strategies; and in environmental monitoring, it estimates the state of large areas even with limited sensors. Furthermore, in machine learning, it is used to significantly boost the performance of complex models through hyperparameter optimization. These examples demonstrate that Bayesian optimization plays a crucial role in solving complex real-world problems, extending beyond being a simple theoretical tool.

However, Bayesian optimization also faces numerous challenges that must be overcome. High-dimensional problems arising from large search spaces, combinatorial optimization problems where search spaces explode due to discrete and combinatorial variables, robust optimization problems where uncertainty increases due to observation noise or outliers, and multi-fidelity optimization problems requiring the joint use of data with differing costs and accuracies remain active areas of research. Furthermore, considering efficient search in distributed systems, adaptation in time-varying dynamic environments, and even social demands like algorithmic fairness, the potential scalability of Bayesian optimization is virtually limitless.

Ultimately, Bayesian optimization holds significant meaning beyond being a mere optimization technique, as it offers a method for efficient selection and learning even amidst uncertainty and constraints. In this exploration, we examined the principles, applications, and future challenges of BO. This method will continue to expand and be utilized across diverse fields.

Example Questions

You may also conduct research on a topic of your choice other than the following question.

Q. Bayesian optimization uses probabilistic surrogate models like Gaussian Processes (GP) to simultaneously explore areas of high uncertainty and exploit regions demonstrating strong performance. After running the Python example code provided in the supplementary materials and observing the results, follow these steps to analyze and apply the approach.

1. Determine the form of the given function $f(x)$, anticipate which intervals may be difficult to predict, execute the code to verify the results, and analyze how accurately the optimal point was found when compared to the actual function graph.

2. After running the analysis with the initial sample size changed to 10, 20, and 30, examine how the width of the confidence interval and the prediction function change.

3. After running the kernel function with different length scales such as 0.05 and 1.0, analyze how the width of the confidence interval and the prediction function change.

4. After running Gaussian Processes with different kernels such as RBF, WhiteKernel, and Compound-Kernel, analyze how the width of the confidence interval and the prediction function change.

5. Select one of the various application cases presented in the text (A/B Testing, Robotics and Reinforcement Learning, Environmental Monitoring and Sensor Network, Automatic Machine Learning and Hyperparameter Tuning) and design the optimization objective and hyperparameters to be set. Analyze the expected efficiency gains (e.g., search strategy, sample efficiency) when applying Bayesian optimization, along with potential technical and computational limitations (e.g., performance degradation in high-dimensional spaces, difficulties in handling categorical variables), providing reasons for each.

References

Frazier, P. I. 2018. A Tutorial on Bayesian Optimization. arXiv preprint arXiv:1807.02811 URL http://arxiv.org/abs/1807.02811.

Lizotte, D. n.d. Automatic Gait Optimization with Gaussian Process Regression. Unpublished manuscript .

Perrone, V., M. Donini, M. B. Zafar, R. Schmucker, K. Kenthapadi, and C. Archambeau, 2021. Fair Bayesian Optimization. Pages 854–863 *in* Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21). ACM.

Rasmussen, C. E., 2004. Gaussian Processes in Machine Learning. Pages 63–71 *in* O. Bousquet, U. V. Luxburg, and G. Rätsch, editors. Advanced Lectures on Machine Learning. Springer.

Shahriari, B., K. Swersky, Z. Wang, R. P. Adams, and N. D. Freitas. 2016. Taking the Human Out of the Loop: A Review of Bayesian Optimization. Proceedings of the IEEE **104**:148–175.

Srinivas, N., A. Krause, S. M. Kakade, and M. Seeger. 2012. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. IEEE Transactions on Information Theory **58**:3250–3265.

Wang, X., Y. Jin, S. Schmitt, and M. Olhofer. 2023. Recent Advances in Bayesian Optimization. ACM Computing Surveys **55**:1–36.

Zhao, P., L. Zhang, Y. Jiang, and Z.-H. Zhou. 2021. A Simple Approach for Non-stationary Linear Bandits. arXiv preprint arXiv:2103.05324 URL http://arxiv.org/abs/2103.05324.

Contributors



Taegyeong Kim

M. ktg0829@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Energy AI and Environmental Climate Technology. Belonging Meta Lab, as an undergraduate researcher, I am working on catalyst optimization and novel catalyst development using machine learning. I aim to save energy and contribute to the development of a catalyst for removing harmful substances by greatly reducing the experimental cost and time required to develop a catalyst.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Sumyeong Ahn

M. sumyeongahn@kentech.ac.kr

Sumyeong Ahn is an assistant professor in the Energy AI track within the Department of Energy Engineering at the Korea Institute of Energy Technology. His research interests include robust AI, trustworthy AI, and foundation models, such as large language models and large multimodal models. Recently, his focus has been on developing robust fine-tuning and inference algorithms. Beyond AI itself, he also focuses on AI for X -- leveraging AI for real-world applications, including energy system, medical system and other critical applications.

KENTECH
Korea Institute of Energy Technology

II

바이오에너지와 CCUS의 만남 :

BECCS를 활용한 탄소 제거 원리와 미래 전망

p. 030 - 047

Vol. 9 (2025)

바이오에너지와 CCUS의 만남: BECCS를 활용한 탄소 제거 원리와 미래 전망

Abstract

BECCS(Bioenergy with Carbon Capture and Storage)는 자연의 탄소흡수원인 바이오매스를 활용하여 에너지를 생산하고, 이 과정에서 발생하는 이산화탄소를 포집·저장함으로써 대기 중 탄소의 순감을 목표로 하는 핵심적인 탄소 제거 기술이다. 본 탐구자료는 BECCS가 식물의 광합성이라는 생화학적 원리를 기반으로, 연소 및 가스화라는 열공학적 에너지 전환 과정을 거쳐, 최종적으로 화학적 흡수 및 지질학적 저장을 통해 탄소를 영구 격리하는 전 과정을 심층적으로 분석한다. 또한, 기술의 잠재력과 함께 토지 이용 경쟁, 에너지 효율 저하, 장기 저장 안정성, 사회·경제적 비용과 같은 현실적인 한계와 딜레마를 균형 있게 고찰한다. 이를 통해 BECCS가 단순한 공학적 해결책을 넘어, 그 기술이 생태계, 식량 시스템, 그리고 사회 전체와 맺는 복잡한 관계 속에서 지속가능성을 확보해야 하는 다차원적 과제임을 제시하고자 한다.

Introduction

지구의 기후 시스템은 자연적인 탄소 순환(Carbon cycle)을 기반으로 운영된다. 대기와 해양, 생물권 및 지각 속 탄소가 서로 교환되며 지구의 기온과 생명을 유지한다. 이와같이 견고하고 복잡한 탄소순환에서 식물은 자연 광합성을 통해 대기 중 탄소를 흡수하고 고정하여, 무기탄소를 유기물로 전환하는 핵심 매개체 역할을 수행한다. 그러나 산업화 이후, 지구의 자연적 탄소 순환 체계의 균형이 흔들리고 있다. 인간의 산업활동에 따른 화석연료 연소는 지층에 격리되어있던 탄소 에너지 자원을 단기간에 대기로 방출하였고, 대기중에 과하게 축적된 탄소는 자연의 흡수 능력을 한계에 다다르게 했다. 이는 전 지구적 기후 변화의 핵심 원인이 되었다 (Intergovernmental Panel on Climate Change (IPCC), 2021). 이러한 불균형을 해결하고자 등장한 개념이 BECCS(Bioenergy with Carbon Capture and Storage)이다. BECCS는 자연의 탄소 흡수 과정을 모방하고 공학적으로 대용량화 시켜 대기 중 CO_2를 실질적으로 순감소시키는 탄소 넷제로를 목표로 한다. 이는 생명체의 탄소 고정 능력을 인간의 에너지 시스템에 통합하여 지구적 수준의 탄소 순환을 능동적으로 재구성하려는 시도로도 볼 수 있다. IPCC(Intergovernmental Panel on Climate Change)는 1.5도 지구 온난화 제한 목표를 달성하기 위한 대부분의 시나리오에서 BECCS와 같은 탄소 제거 기술의 중요성을 강조하고 있다 (Intergovernmental Panel on Climate Change (IPCC), 2018). 따라서 본 탐구자료에서는 BECCS의 근간을 이루는 과학적 원리와 이를 구현하는 공학적 시스템을 심층적으로 소개하고자 한다.

1. 배경지식 : BECCS의 과학

BECCS는 화학, 열공학, 화학공학, 지질학 등 다양한 학문이 융합된 복합 기술 시스템이다. 식물의 자연 광합성 기작을 시작으로 에너지를 생산하고 탄소를 격리하는 연속적인 공학적 과정에 기반한다.

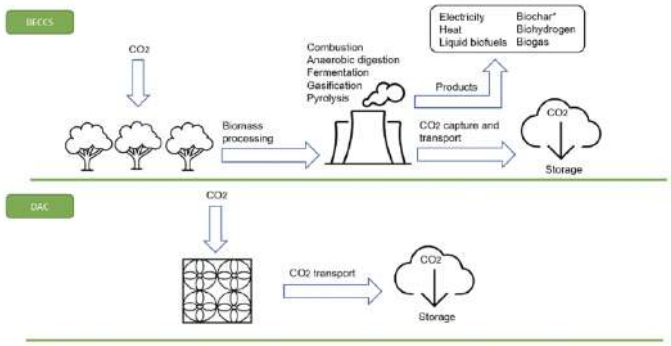


Figure 1. BECCS의 전체 모식도

1-1. 탄소 넷제로의 메커니즘

기존의 바이오매스 생성은 식물이 광합성으로 흡수한(-CO_2) 만큼의 CO_2를 연소 시 다시 배출(+CO_2)하므로, 전체적으로는 탄소 증감이 없는 탄소 중립(Carbon-Neutral)에 가깝다. 반면 BECCS는 연소 시 발생하는 CO_2를 포집하여 지하에 영구히 격리한다. 따라서 전체 과정은

$$\{\text{식물의 } CO_2 \text{ 흡수}(CO_2)\} - \{\text{연소 시 } CO_2 \text{ 배출 후 포집 및 격리}(0)\}$$

로 요약될 수 있으며, 결과적으로 대기 중 CO_2를 순수하게 감소시켜서 배출되는 CO_2와의 균형을 이루어서 '탄소 넷제로'를 구현하게 된다. 하지만 이는 바이오매스의 재배(비료 사용), 가공, 운송 과정에서 발생하는 '숨겨진 탄소 발자국'을 최소화하고, 90% 이상의 높은 포집 효율을 달성했을 때 유효하다 (Fajardy and Mac Dowell, 2017). 이것이 BECCS가 기후 변화 대응의 게임 체인저로 평가받는 이유이자, 동시에 기술의 복잡성이 요구되는 지점이다.

1-2. BECCS의 3대 과학적 과정

BECCS의 전체 시스템은 다음과 같은 연속적 반응으로 이루어진다.

1. 생화학적 탄소 고정 (Biochemical Carbon Fixation)
식물은 태양에너지를 촉매로 사용하여 대기 중 CO_2와 물을 포도당(탄수화물)과 산소로 전환한다. 이 과정은 흡열반응 (Endothermic reaction, $\Delta H > 0$)이며 결론적으로 태양에너지를 화학에너지 형태로 바이오매스 내에 저장하게 된다.

$$6\,CO_2(g) + 6\,H_2O(l) \rightarrow C_6H_{12}O_6(s) + 6\,O_2(g) \quad (\Delta H \approx 2830\,\text{kJ/mol per glucose unit})$$

2. 열공학적 에너지 전환 (Thermo-engineered Energy Conversion)
수확한 바이오매스(주 성분은 셀룰로스, (CH_1O))를 연소시켜 저장된 화학에너지를 열에너지로 방출하는 과정이다. 이 과정은 발열 반응으로 BECCS 전체 시스템에 필요한 에너지를 공급하는 핵심 단계이다.

$$(C_6H_{10}O_5)_n(s) + 6n\,O_2(g) \rightarrow 6n\,CO_2(g) + 5n\,H_2O(g) + \text{Heat} \quad (\Delta H \approx -2830\,\text{kJ/mol per glucose unit})$$

3. 화학/지질공학적 격리 (Chem-Geological Sequestration)
에너지 전환 과정에서 발생한 CO_2를 화학적 흡수제를 이용해 선택적 분리/포집하고, 이를 영구적으로

격리하는 과정이다. 이 과정은 가역적인 산-염기 반응을 이용하며 발열반응인 포집과 흡열반응인 재생, 즉 상반된 열화학적 특성을 가진다.

2. 구조 공학적 분석 : BECCS 발전소의 작동 원리

BECCS 발전소는 기본적으로 화석연료를 연료로 사용하는 화력발전소에, 연소로 생성된 배기가스(Flue Gas)에서 CO_2를 선택적으로 분리하는 연소 후 포집 (Post-combustion Capture) 설비가 결합된 통합 플랜트 형태를 가진다. 현재 이런 형태의 BECCS 플랜트의 가장 대표적인 사례는 영국 Drax 발전소의 BECCS 프로젝트로, 목재 펠릿을 연료로 사용하여 탄소 네거티브를 실현하고 있다 (Drax Group, 2024*a*).

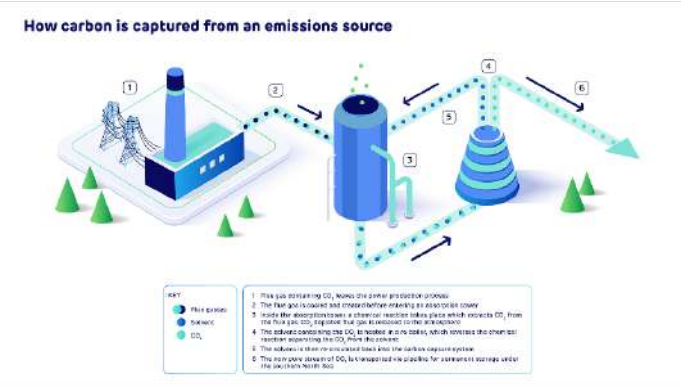


Figure 2. Drax 발전소의 BECCS 모식도

2-1. 지속가능한 바이오매스 확보 및 전처리

BECCS 시스템의 성공적인 운영을 위한 핵심 전제는 지속가능한 연료 확보이다. 만약 에너지 작물 재배를 위해 식량 경작지를 침범하거나 자연림을 훼손한다면, 기술의 근본적인 의미가 퇴색될 수 있다. 따라서 BECCS는 식량 안보 및 생태계와 경쟁하지 않는 농업 부산물(볏짚), 임업 폐기물(폐목재), 또는 척박한 땅에서도 잘 자라는 비식용 에너지 작물(억새 등)을 주요 연료로 활용하는 것을 지향해야한다 (Intergovernmental Panel on Climate Change (IPCC), 2019). 바이오매스는 수확되어 발전소로 이송된 후, 연소 효율 증대를 위해 건조, 분쇄, 또는 압축하여 pellet, chip, briquette, 반탄화 바이오매스, oil, gas 등 다양한 형태의 연료로 만드는 전처리 과정을 거친다.

2-2. CO_2 포집에서 저장까지의 상세 공정 흐름

BECCS 발전소에서 CO_2는 다음과 같은 정밀하게 제어되는 연속 공정을 통해 포집되고 저장된다.

연소 및 에너지 생산 → 이산화탄소 포집 → 흡수제 재생 및 이산화탄소 분리 → 압축/수송 → 지질 저장

연소 및 에너지 생산 (Combustion & Power Generation)
펠릿 형태로 전처리된 바이오매스는 유동층 보일러(Fluidized Bed Boiler)에서 공기와 함께 연소된다. 이 방식은 연료를 고르게 연소시켜 효율을 높이고, 질소산화물(NO)과 같은 오염물질 배출을 줄이는 장점이 있다. 연소 시 발생하는 고온·고압의 증기는 증기 터빈을 회전시켜 발전기를 구동하고 전기를 생산한다.

이산화탄소 포집 (CO_2 Capture)
연소 후 발생한 약 40 60℃의 배기가스는 흡수탑(Absorber)의 하부로 유입된다. 상부에서는 냉각된 아민(Amine) 계열의 액체 흡수제가 분사된다. 대표적인 흡수제인 모노에탄올아민(MEA, $HO(CH_2)_2NH_2$)은 약염기성을 띠며, 산성 기체인 CO_2와 만나 발열 반응을 통해 Carbamate 이온을 형성하며 화학적으로

결합한다. 이 과정을 통해 배기가스 내 CO_2의 90% 이상을 제거할 수 있다. CO_2가 제거된 나머지 가스(주로 질소)는 탑 상부로 배출된다.

$$2\,RNH_2\,(MEA) + CO_2 \;\rightleftharpoons\; RNH_3^+ + RNHCOO^- \quad (R = HO(CH_2)_2)$$

흡수제 재생 및 CO_2 분리 (Solvent Regeneration & CO_2 Stripping)
CO_2를 머금어 농도가 높아진 흡수제(Rich Solvent)는 열 교환기를 거쳐 예열된 후 재생탑(Stripper)으로 이송된다. 재생탑 하부에서는 재비기(Reboiler)를 통해 약 100 120℃의 증기가 공급된다. 이 열에너지는 르 샤틀리에 원리(Le Chatelier's principle)에 따라 위 포집 반응의 역반응을 촉진한다. 즉, 흡열 반응(Endothermic)을 통해 CO_2와 흡수제의 화학 결합이 끊어지면서, 순도 99% 이상의 고농도 CO_2 기체가 분리되어 탑 상부로 빠져나온다. CO_2가 분리되어 농도가 낮아진 흡수제(Lean Solvent)는 다시 냉각되어 흡수탑으로 돌아가 재활용된다.

$$RNH_3^+ \leftrightarrow RNHCOO^- + Heat \;\rightarrow\; 2\,RNH_2\,(MEA) + CO_2$$

이 재생 과정은 외부에서 열에너지를 지속적으로 공급해야만 일어나는 흡열 과정이므로, 전체 공정에서 가장 많은 에너지를 소모하는 단계이며 BECCS의 효율과 경제성을 결정하는 핵심적인 부분이다 (Rochelle, 2009).

압축 및 수송 (Compression Transport)
분리된 고순도 CO_2는 냉각 및 탈수 과정을 거친 후, 다단 압축기(Multi-stage Compressor)를 통해 약 7.4 MPa, 31℃ 이상의 임계점(Critical Point)을 넘어선 초임계유체(Supercritical Fluid) 상태로 압축된다. 초임계 CO_2는 기체와 액체의 성질을 모두 가져 점성은 낮고 밀도는 높아 파이프라인 수송에 매우 효율적이다. 압축된 CO_2는 파이프라인이나 특수 액화 CO_2 운반선박을 통해 최종 저장 부지로 이송된다.

지질 저장 (Geological Storage)

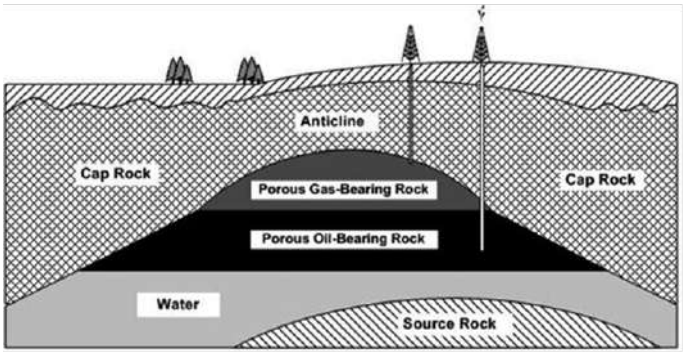


Figure 3. 이상적인 CO_2 저장 환경인 배사(anticline)의 구조

최종적으로 CO_2는 육상 또는 해저의 800m 이상 깊이에 있는 염수층(Saline Aquifer, Fig 3.에선 Water로 표현됨)이나 고갈된 유·가스전에 주입되어 저장된다. 실제로 석유와 가스 등 화석 연료는 Fig.3과 같은 배사 트랩 구조에 안정적으로 존재하는데, 인간이 인위적으로 화석 연료를 채굴하기 전까지는 수백만년부터 수십억년 동안 새어나오지 않고 보존되었다. BECCS에서도 마찬가지로, 배사 트랩의 가스나 오일이 있던 공간(Fig 3.에서의 Porous Gas / Oil Bearing Rock)에 CO_2를 주입해넣어 검증된 자연적 구조로 보존하는 것이다.(석유를 뽑아올리는 유정(oil well)의 반대 과정이라고 생각하면 이해하기 쉽다.) 염수층은 인간이 활용하지 않는, 염분이 높은 물로 채워진 다공성 암석층으로 CO_2 저장 잠재량이 가장 큰 것으로 평가된다. 주입된 CO_2는 (1) 암석층의 구조에 갇히거나(구조적 포획) (2) 미세한 공극에 포획되고(잔류 포획) (3) 지하수에 용해되거나(용해 포획) (4) 수천 년에 걸쳐 주변 광물과 반응하여 탄산염 광물로 변하는

(광물 포획) 메커니즘을 통해 영구적으로 격리된다 (Global CCS Institute, 2023). 상부의 치밀하고 비투과성인 덮개암(Cap Rock)이 덮개 역할을 하여 대기로의 누출을 방지하며, 실제로 노르웨이의 Sleipner 저장 프로젝트는 1996년부터 수천만 톤의 CO_2를 안정적으로 저장하며 지중 저장의 안전성을 입증하고 있다.

이처럼 BECCS는 식물의 광합성부터 초임계유체 수송, 지중 저장에 이르기까지 복잡하고 정교한 시스템이다. 이상적인 상태에서의 설계된 플랜트의 이론적인 설명만 보고있자면 BECCS는 기후 위기를 해결할 완벽한 대안처럼 보인다. 하지만 이 BECCS 기술이 과연 현실 세계 규모에서 지속가능하게 작동할 수 있을지가 관건이다. 따라서, 이번 탐구에서는 기술적 구현 가능성 뿐만 아니라 이 기술이 우리가 살아가는 생태계, 식량 시스템, 그리고 사회 전체에 미칠 영향을 고려하여 그 잠재력과 딜레마에 대한 심도있는 고찰이 필요하다.

3. 생태계를 고려한 재료과학적 탐구: BECCS 기술의 잠재력과 딜레마

BECCS는 에너지 생산과 탄소 제거를 동시에 실현할 수 있는 혁신적인 기술이다. 앞서 서술하였듯이, BECCS는 식물의 광합성을 통해 대기 중 CO_2를 흡수하고, 바이오에너지를 연료로 사용한 뒤, 연소 과정에서 발생하는 CO_2를 포집 및 저장함으로써 순탄소배출량을 0 으로 만드는 '탄소 넷제로' 시스템을 구축하는 기술이다. 이처럼 BECCS는 자연의 탄소 순환 과정을 공학적으로 강화하여, 인류의 산업 활동이 대기 중 탄소 농도를 줄이는 방향으로 전환될 수 있도록 설계된 기술이다 (Intergovernmental Panel on Climate Change (IPCC), 2021) (Intergovernmental Panel on Climate Change (IPCC), 2018).

3-1. BECCS의 잠재력

BECCS는 크게 세 가지 부분에서 잠재력을 지니고 있다. 먼저 가장 큰 장점은 기후 변화 완화 효과가 직접적이라는 점이다. IPCC(2018)(Intergovernmental Panel on Climate Change (IPCC), 2018)는 지구 평균 기온 상승은 1.5℃ 이내로 제한하는 시나리오 대부분에서 BECCS를 주요 기술로 제시하고 있으며, 전 세계적으로 매년 최대 10기가톤(Gt) 수준의 CO_2 제거 잠재력이 있다고 분석한다 (Fajardy and Mac Dowell, 2017). 즉, 이 기술은 단순히 배출을 줄이는 기술을 넘어 이미 대기 중에 존재하는 탄소를 실제로 회수하는 적극적인 기후 대응 기술이다.

또한 전력 안정성 측면에서도 강점을 가진다. 태양광이나 풍력과 달리 날씨나 시간에 영향을 받지 않기 때문에, BECCS 발전은 재생에너지 중심의 전력망에서 기저부하 역할을 수행할 수 있다 (Drax Group, 2024*a*). 따라서 이 기술은 탄소 중립 전력원을 넘어 재생에너지의 간헐성으로 인한 불안정성을 보완하는 보완책으로 활용될 가능성이 있다.

더불어 BECCS는 산업 전반으로 확장 가능한 기술 플랫폼이다. 바이오 에탄올 공장, 제지 산업, 식품 가공 공정 등 다양한 분야에서 발생하는 CO_2를 포집 및 저장할 수 있으며, 이 과정에서 얻은 CO_2를 탄산염 광물화나 친환경 소재 생산에 활용하면 탄소 자원화로 확장할 수 있다 (Global CCS Institute, 2023). 이처럼 BECCS는 에너지, 산업, 소재 분야가 연결되는 탄소 순환형 경제의 기반이 될 수 있다.

3-2. BECCS의 한계점

하지만 이와 같은 기술에 환경적, 기술적, 지질적/안전적 한계, 사회/경제적 한계가 존재한다.

3-2-1. 환경적 한계

BECCS의 가장 큰 환경적 문제는 토지 이용 경쟁이다. 대규모 에너지 작물 재배가 필요하므로 식량 생산지와 경쟁하게 되고, 이로 인해 식량 가격 상승, 산림 파괴, 생물다양성 감소가 발생할 수 있다 (Intergovernmental Panel on Climate Change (IPCC), 2019). 이처럼 BECCS가 오히려 환경적 부담을 옮겨놓는 결과를 초래할 가능성이 있어, 지속가능한 원료 관리가 필수적이다 (**?**).

3-2-2 기술적 한계

BECCS의 기술적 과제는 어떻게 효율적으로 포집 및 저장할 것인가에 있다. 앞서 살펴본 공정은 이미 상용화된 기술이지만, 전체 시스템 효율과 비용이 여전히 큰 문제다. CO_2 포집 및 재생 단계에서 상당한 에너지가 필요해 발전 효율이 약 10-15 % 감소하며, 포집, 압축, 수송, 주입까지 이어지는 연속 공정의 에너지 손실과 열관리 부담이 크다 (Rochelle, 2009). 따라서 현재 연구는 흡수제의 열적 안정성 개선, 저온재생 기술 개발, 포집 공정 통합 설계 및 포집-전환 연속 공정 기술 개발 등에 집중되고 있다. 또한 포집된 이산화탄소를 저장하기 위한 인프라 구축 비용도 높아, 기술적 완성도와 경제성이 함께 요구된다는 한계를 가지고 있다 (Scottish Government, 2021).

3-2-3 지질적/안전적 한계

BECCS의 마지막 단계인 지중 저장은 지질 구조의 안정성에 전적으로 의존한다. CO_2는 800m 이상의 염수층이나 고갈된 유전, 가스전에 주입되기 때문에, 장기간에 걸친 누출 가능성을 완전히 배제할 수는 없다 (Global CCS Institute, 2023). 또한 주입 압력이 높을 경우 지층 내에서 미세한 지진 활동이 유발될 수 있다. 따라서 BECCS의 신뢰성을 확보하기 위해서는 정밀한 지질조사, 실시간 모니터링, 장기 안전성 데이터 구축이 필요하다.

3-2-4 사회/경제적 한계

마지막으로 BECCS는 막대한 초기 투자 비용과 낮은 수익성 때문에 경제적 장벽이 크다. 현재 CO_2 1톤을 포집하고 저장하는데 약 100달러 이상이 필요하며, 이는 일반 화력발전보다 매우 비싼 수준이다 (Global CCS Institute, 2023). 또한 대규모 작물 재배나 지중 저장 시설에 대한 주민반대와 환경 단체의 우려로 인해 사회적 수용성이 낮을 수 있다. 따라서 BECCS의 확산을 위해서는 정부의 정책 지원, 국제 협력, 지역 사회와의 신뢰 형성이 함께 이루어져야 한다.

결국 BECCS는 지속가능한 미래 에너지 체계의 핵심 후보이지만, 기술적 효율성, 환경적 부담, 사회적 수용성 간의 균형이 필수적이다. 이 기술이 진정한 탄소 넷제로를 실현하기 위해서는, 과학기술의 발전뿐 아니라 정책적, 윤리적 접근이 병행되어야 할 것이다.

4. Conclusion

BECCS는 자연의 탄소 순환 원리를 인류의 에너지 시스템 안에 통합한 기술로, 단순히 탄소 배출을 줄이는 것이 아니라 대기 중 탄소를 실제로 제거하는 역할을 한다. 식물의 광합성, 연소과정, 이산화탄소의 포집과 저장을 고부가 연료 혹은 화합물로의 전환 공정 기술과 결합하며 산업 활동을 지속하면서 기후변화 완화에 기여할 수 있는 새로운 가능성을 제시한다. 이러한 이유로 BECCS는 탄소중립을 넘어 탄소 네거티브 사회로 나아가기 위한 핵심 기술로 평가된다.

하지만 BECCS가 이상적인 목표를 현실적으로 달성하기 위해서는 여러 과제가 존재한다. 대규모 바이오매스 재배로 인한 토지 이용, 수자원 문제, 높은 포집 비용과 에너지 손실, 저장 과정의 안정성 등은 해결되어야 할 중요한 한계들이다. 특히 기술적 효율성과 생태적 지속가능성의 균형을 유지하지 못한다면, BECCS의 탄소 감축 효과는 제한될 수 밖에 없다.

따라서 앞으로의 BECCS 박전은 기술향상을 넘어 정책적 지원과 사회적 수용성, 그리고 환경적 책임이 고려되어야 한다. 정부, 기업, 연구기관이 협력하여 지속가능한 바이오매스 관리 체계와 안전한 저장 인프라를 구축한다면 BECCS는 인류가 직면한 기후 위기를 해결하기 위한 실질적인 해답이 될 수 있을 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

1. 100MW*급 *BECCS* 발전소에서 에너지 작물인 '억새'를 사용할 때, 연간 CO_2 순감축량을 계산해보자. (단, 억새의 재배–운송 과정 탄소배출량, 발전소의 포집 효율 *90%* 가정, 저장 과정의 에너지 소모량 등을 모두 고려하여 계산)

- 연간 연료소모량: $F = \dfrac{\text{발전량 (MWh)}}{n \times H}$
- 연소 CO_2 발생량: $E_1 = F \times$ 탄소함량 계수
- 포집 CO_2: $C = 0.9 \times E_1$ (90% 가정)
- 전과정 배출: $E_{\text{전체}} = E_{\text{재배}} + E_{\text{운송}} + E_{\text{전처리}} + E_{\text{압축/수송/주입}}$
- 순감축량: $C - E_{\text{전체}}$
- 데이터 체크리스트: 억새 재배, 비료 사용의 배출 계수, 포집, 재생 단계 열소모와 전력소모, 압출, 수송, 저장 에너지

이후, 각 변수의 입력값(예: 포집 효율 85~95%, 운송·전처리 배출량 ± 50%)을 변경하면서 연간 CO_2 순감축량이 어떻게 달라지는지 확인하고, 각 변수의 민감도를 분석해보자.

***2. BECCS* 기술을 전 세계적으로 확대할 때, 식량 안보와 생물다양성을 해치지 않기 위해 어떤 국제적 규범이나 정책*(예: 지속가능성 인증제, 토지 이용 가이드라인)*이 필요할까?**

BECCS 확산 시 식량, 생물 다양성, 수자원을 동시에 지키는 최소 기준을 만들어보자. 아래의 권고프레임을 고려하여 현장 적용 표준안을 만들고, 실제 사례를 평가해보자.

권고 프레임:

- 토지: 산림 전환 금지, 농업부산물 및 임업 폐기물 우선, 에너지 작물은 비옥하지 않은 토지에 한정

- 생물 다양성: 보호 구역, 완충 구역(예: 하천, 습지 주변 50-100m) 제외, 토양 유기 탄소(SOC) 비감소 원칙

- 물: 지역 물 스트레스 지수 고려, 관개 금지/최소화, 물 재이용 계획 제출

- 탄소 회계: 재배-가공-운송 전과정 배출 보고 의무화, 외부 검증(MRV) 포함

- 정책 연계: 국가 배출권제와 지속 가능성 인증을 연동

***3.* 발전*(전기 생산)* 분야 외에, 바이오에탄올을 생산하는 정유 공장이나 제지 공장 등 다른 산업 분야에 *BECCS* 원리를 적용하여 탄소 네거티브 산업으로 전환할 수 있는 구체적인 방법은 무엇일까?**

"전력용 BECCS"를 산업 부문으로 확장해 배출원 CO_2 농도/유량, 포집열 통합, 압축/수송 경로, 저장 후보지, MRV 체계 등을 고려한 탄소 네거티브 산업 시나리오를 만들어보자.

References

Drax Group, 2024*a*. BECCS at Drax. https://www.drax.com/beccs/.

Drax Group, 2024*b*. Bioenergy with carbon capture, use, and storage (BECCS). https://www.drax.com/about-us/bioenergy-carbon-capture-use-and-storage-beccs/.

Fajardy, M., and N. Mac Dowell. 2017. Can BECCS deliver sustainable and resource-efficient negative emissions? Energy & Environmental Science **10**:1389–1426.

Global CCS Institute, 2023. Global status of CCS 2023. https://www.globalccsinstitute.com/resources/global-status-of-ccs-2023/.

Intergovernmental Panel on Climate Change (IPCC). 2018. Global warming of 1.5℃: an IPCC special report on the impacts of global warming of 1.5℃ above pre-industrial levels and related global greenhouse gas emission pathways. Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC). 2019. Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC). 2021. Climate change 2021: the physical science basis. Cambridge University Press.

Rochelle, G. T. 2009. Amine scrubbing for CO_2 capture. Science **325**:1652–1654.

Scottish Government, 2021. Negative emissions technologies (NETS): feasibility study. https://www.gov.scot/publications/negative-emissions-technologies-feasibility-study/.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Youngsoo Kang

M. yskang@kentech.ac.kr

Prof. YoungSoo Kang at School of Energy Engineering of KENTECH. His research area is carbon neutrality climate technology; Solar CO2 conversion into liquid fuels, Low temperature, microwave absorption and supercritical fluid degradation of waste plastics into fuels like gasoline, diesel, hydrogen gas and high value carbon based materials. He is director of three big research projects “Leaders Project on Artificial Photosynthesis of CO2 Conversion into Liquid Fuels” (one million US dollars per year for 5 years), “Industrial Alchimist Project on Continuous Flow System for Mass Production of Liquid Fuels by CO2 Conversion” (4 million US dollars per year for 5years) and “Low Temperature Degradation of Waste Plastics into Fuels” (0.3 million US dollars, 3 years). He has worked as President of Korean Chemical Society, Physical Chemistry Division during 2023 and has worked as editorial board members of 5 international journals ; Materials Today Sustainability, Nanomaterials, Rare Metals, Molecules, Energies.

KENTECH
Korea Institute of Energy Technology

Vol. 9 (2025)

Bioenergy Meets CCUS: Principles and Future Prospects of Carbon Removal through BECCS

Abstract

Bioenergy with Carbon Capture and Storage (BECCS) is a core carbon-removal technology that utilizes biomass (nature's carbon sponge) to produce energy while capturing and storing the final result, carbon dioxide. The aim is to achieve a reduction of atmospheric CO_2. This research material provides an in-depth analysis of the full BECCS process, from the biochemical principles of photosynthesis to thermo-engineered energy conversion (combustion and gasification), followed by chemical absorption and geological capturing and storage for permanent storage.

It also examines the realistic challenges of BECCS—including competition of land-use, energy efficiency losses, long-term storage stability, and socio-economic costs—through a balanced lens.

Ultimately, this study presents BECCS not merely as an engineering solution, but as a multidimensional task that must ensure sustainability within complex relationships among the ecosystem, food system, and human society.

Introduction

Earth's climate system is governed by the natural carbon cycle, in which carbon is exchanged among the atmosphere-the air, hydrosphere-the water, biosphere-the places where life lives in, and lithosphere-the land itself, maintaining global temperature and life. Within this intricate cycle, plants serve as essential mediators by absorbing atmospheric carbon dioxide through photosynthesis and converting inorganic carbon into organic matter.

However, since industrialization, this natural balance has been disrupted. The combustion of fossil fuels has released vast amounts of geologically stored carbon into the atmosphere, exceeding Earth's natural absorption capacity and in conclusion became a key driver of global climate change (Intergovernmental Panel on Climate Change (IPCC), 2021).

In response to this imbalance, the concept of BECCS emerged. BECCS replicates and scales up nature's carbon-absorption process through engineering, aiming carbon net-zero by achieving a genuine reduction in atmospheric CO_2. It integrates the carbon-fixing ability of living organisms into human energy systems, enabling proactive reconstruction of the global carbon cycle. The IPCC emphasizes that technologies such as BECCS are essential in most pathways limiting global warming to 1.5℃ (Intergovernmental Panel on Climate Change (IPCC), 2018). This research therefore explores both the scientific foundations and the engineering mechanisms underpinning BECCS.

1. Scientific Background: The Science of BECCS

BECCS represents a multidisciplinary system combining chemistry, thermal engineering, chemical engineering, and geology. It begins with the photosynthesis of plants and proceeds through a series of engineered processes that generate energy while capturing and isolating carbon.

1-1. The Mechanism of Carbon Net-Zero

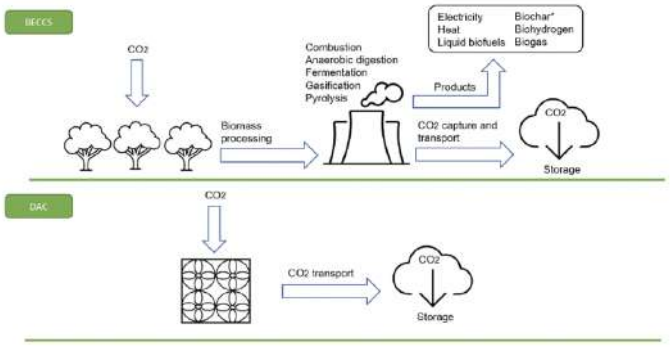


Figure 1. Overall schematic diagram of BECCS

In conventional biomass combustion, the CO_2 absorbed during plant growth (CO_2) is released back into the atmosphere upon combustion ($+CO_2$), leading to carbon neutrality. In contrast, BECCS captures the CO_2 emitted during combustion and stores it underground, achieving net carbon removal.
This can be summarized as the following:

$$\{\text{Plant } CO_2 \text{ absorbtion}(CO_2)\} - \{\text{Captured and stored combustion } CO_2\ (0)\}$$

This results in a net decrease of atmospheric CO_2, achieving a balance between emission and removal. However, this is only effective when “hidden carbon footprints” from cultivation, processing, and transport are minimized, and efficiency of capturing exceeds 90% (Fajardy and Mac Dowell, 2017). This dual demand for efficiency and complexity defines BECCS as both a breakthrough and a technological challenge.

1-2. Three Core Scientific Processes of BECCS

1. Biochemical
Carbon Fixation Plants use solar energy to convert atmospheric CO_2 and water into glucose and oxygen—an endothermic reaction $(\Delta H > 0)$ that stores chemical energy within biomass:

$$6\,CO_2(g) + 6\,H_2O(l) \rightarrow C_6H_{12}O_6(s) + 6\,O_2(g) \quad (\Delta H \approx 2830\,\text{kJ/mol per glucose unit})$$

2. Thermo-Engineered Energy Conversion
Harvested biomass (mainly cellulose, $(C_6H_{10}O_5)_n$ is combusted to release the stored chemical energy as heat—an exothermic reaction providing power for the BECCS system:

$$(C_6H_{10}O_5)_n(s) + 6n\,O_2(g) \rightarrow 6n\,CO_2(g) + 5n\,H_2O(g) + \text{Heat} \quad (\Delta H \approx -2830\,\text{kJ/mol per glucose unit})$$

3. Chemical/Geological Sequestration
The emitted CO_2 is selectively absorbed by chemical solvents and permanently captured underground through reversible acid–base reactions, involving both exothermic absorption and endothermic regeneration.

2. Structural and Engineering Analysis: Operating Principles of a BECCS Power Plant

A BECCS plant integrates a conventional biomass-fired thermal power station with post-combustion CO_2 capture facilities that separate CO_2 from flue gases. The most notable example is the Drax Power Station BECCS Project in the UK, which utilizes wood pellets to achieve carbon negativity (Drax Group, 2024*a*).

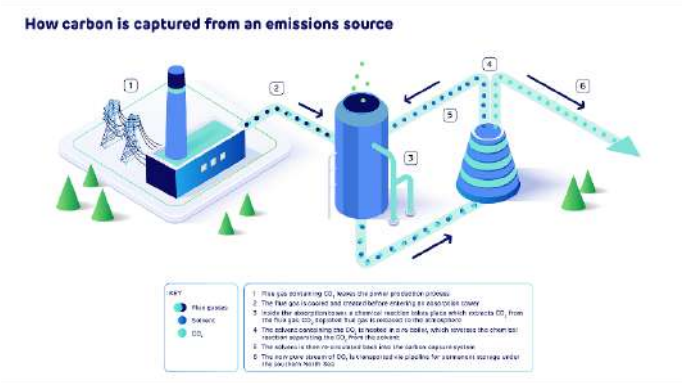


Figure 2. Schematic diagram of BECCS at Drax Power Station

2-1. Sustainable Biomass Procurement and Pretreatment

Sustainable feedstock is a prerequisite for BECCS. If food cropland or forests are exploited, the environmental value of the technology collapses. Thus, BECCS should rely on non-competitive feedstocks such as agricultural residues (e.g., rice straw), forestry waste (wood residues), and non-food energy crops (e.g., Miscanthus) that can grow on marginal lands (Intergovernmental Panel on Climate Change (IPCC), 2019).
Harvested biomass is transported to the plant and pretreated—dried, pulverized, or compressed—into various fuel forms such as pellets, chips, briquettes, torrefied biomass, oils, or gases to enhance combustion efficiency.

2-2. Process Flow: From CO_2 Capture to Storage

In a BECCS power plant, CO_2 is collected and stored through a precisely controlled, continuous process: combustion and energy production → carbon dioxide capture → absorbent regeneration and carbon dioxide separation → compression and transport → lipid storage.

Combustion & Power Generation
Pre-processed biomass in pellet form is combusted with air in a fluidized-bed boiler. This method has the advantage of burning the fuel evenly, increasing efficiency and reducing emissions of pollutants such as nitrogen oxides (NO_x). The high-temperature, high-pressure steam generated during combustion rotates a steam turbine, driving a generator and producing electricity.

CO_2 Capture
The exhaust gas generated after combustion, at approximately 40-60℃, flows into the bottom of the absorber. A cooled amine-based liquid absorbent is injected at the top. Monoethanolamine (MEA, $HO(CH_2)_2NH_2$), a typical absorbent, is weakly basic and meets with the acidic gas CO_2, forming carbamate ions through an exothermic reaction and chemically bonding. This process removes over 90% of the CO_2 in the exhaust gas. The remaining gas (mainly nitrogen) after CO_2 removal is discharged to the top of the tower.

$$2\,RNH_2\,(MEA) + CO_2 \;\rightleftharpoons\; RNH_3^+ + RNHCOO^- \quad (R = HO(CH_2)_2)$$

Solvent Regeneration & CO_2 Stripping
The rich solvent, which has a high CO_2 vapor concentration, is preheated through a heat exchanger and then transferred to the stripper. At the bottom of the stripper, steam at approximately 100-120℃ is supplied through a reboiler. This thermal energy promotes the reverse reaction of the CO_2 capture reaction according to Le Chatelier's principle. In other words, the chemical bond between the CO_2 and the absorbent is broken through

an endothermic reaction, and high-concentration CO_2 gas with a purity of over 99% is separated and released at the top of the tower. The lean solvent, which has a lower concentration after CO_2 is separated, is cooled again and returned to the absorption tower for recycling.

$$RNH_3^+ \leftrightarrow RNHCOO^- + \text{Heat} \rightarrow 2\,RNH_2\,(MEA) + CO_2$$

This regeneration process is an endothermic process that requires a continuous supply of heat energy from an external source. As such, it is the most energy-consuming step in the entire process and is a key factor in determining the efficiency and economic viability of BECCS (Rochelle, 2009).

Compression & Transport

The separated high-purity CO_2 undergoes cooling and dehydration processes, then passes through a multi-stage compressor where it is compressed to a supercritical fluid state at approximately 7.4 MPa and above 31℃, exceeding its critical point. Supercritical CO_2 possesses the properties of both a gas and a liquid, with low viscosity and high density, making it highly efficient for pipeline transportation. The compressed CO_2 is then transported to its final storage location via pipeline or special liquefied CO_2 carriers.

Geological storage

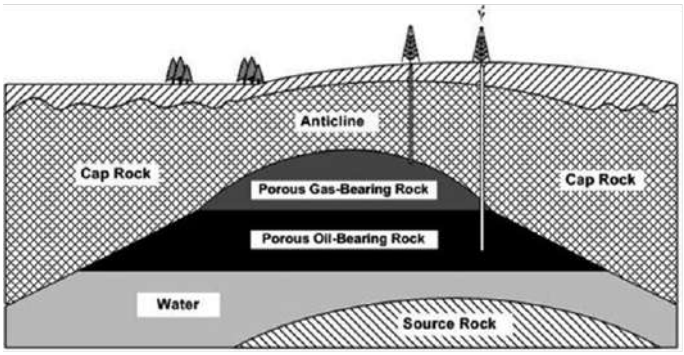


Figure 3. The structure of the anticline, an ideal CO_2 storage environment

Ultimately, CO_2 is injected and stored on land or in saline aquifers (represented by Water in Fig. 3) located at depths of 800 meters or more on the seafloor, or in depleted oil and gas fields. In fact, fossil fuels such as oil and gas exist stably in drainage trap structures like those shown in Fig. 3, and were preserved without leaking for millions to billions of years until humans artificially extracted them. Similarly, in BECCS, CO_2 is injected into the space formerly occupied by gas or oil in drainage traps (porous gas/oil bearing rock in Fig. 3) and stored in a verified natural structure.

This rock formation is considered to have the greatest potential for CO_2 storage. Injected CO_2 is permanently sequestered by several mechanisms: (1) being trapped within the rock formation (structural capture), (2) being trapped in microscopic pores (residual carbon capture), (3) dissolving into groundwater (dissolution capture), or (4) reacting with surrounding minerals over thousands of years to form carbonate minerals (mineral capture) (Global CCS Institute, 2023). A dense, impermeable cap rock acts as a cover to prevent CO_2 from leaking into the atmosphere. Norway's Sleipner storage project has successfully stored tens of millions of tons of CO_2 since 1996, demonstrating the safety of geological storage.

As such, BECCS is a complex and sophisticated system, spanning plant photosynthesis, supercritical fluid transport, and geological storage. Looking only at theoretical descriptions of ideally designed plants, BECCS appears to be a perfect solution to the climate crisis.

However, the key question remains: can this technology actually operate sustainably on a global scale? This exploration therefore requires deep consideration not only of technical feasibility, but also of the possibilities and dilemmas that this technology will pose for our ecosystems, food systems, and society as a whole.

3. Eco-Material Science Perspective: Potentials and Dilemmas of BECCS

BECCS is an innovative technology capable of producing energy simultaneously and removing carbon. By combining photosynthetic carbon fixation with CO_2 capture and storage, it establishes a carbon-net-zero system that strengthens the natural carbon cycle through engineering design (Intergovernmental Panel on Climate Change (IPCC), 2021) (Intergovernmental Panel on Climate Change (IPCC), 2018).

3-1. Potentials of BECCS

BECCS has potential in three major areas.

First, its greatest advantage is its direct effect on climate change mitigation. The Intergovernmental Panel on Climate Change (Intergovernmental Panel on Climate Change (IPCC), 2018) presents BECCS as a key technology in most scenarios that limit global temperature rise to 1.5℃ (Fajardy and Mac Dowell, 2017), predicting a CO_2 removal potential of up to 10 gigatonnes per year(Gt/year) worldwide. In other words, this technology goes beyond simply reducing emissions to actively capture carbon already in the atmosphere.

It also has advantages in terms of power stability. Unlike solar and wind power, BECCS power generation is not affected by weather or time of day, so it can serve as a base load in a renewable energy-dominated power grid (Drax Group, 2024*a*). Therefore, this technology has the potential to go beyond being a carbon-neutral power source and be used as a complementary measure to compensate for the instability caused by the intermittency of renewable energy.

Furthermore, BECCS is a technology platform that can be expanded across industries. It can capture and store CO_2 generated in a variety of fields, including bioethanol plants, the paper industry, and food processing, and the CO_2 obtained in this process can be used for carbon mineralization and environmentally friendly material production, expanding its use as a carbon resource (Global CCS Institute, 2023). In this way, BECCS can serve as the foundation for a carbon-circulating economy that connects the energy, industry, and materials sectors.

3-2. Limitations and Challenges

3-2-1. Environmental Constraints

Large-scale cultivation of energy crops may compete with food production, leading to rising food prices, deforestation, and biodiversity loss (Intergovernmental Panel on Climate Change (IPCC), 2019). Thus, sustainable feedstock management is essential.

3-2-2. Technical Constraints

Although component technologies are commercialized, overall system efficiency remains limited. CO_2 capture and regeneration consume significant energy, reducing net power efficiency by 10–15%. Ongoing research focuses on improving solvent stability, developing low-temperature regeneration methods, and optimizing integrated capture-conversion processes (Rochelle, 2009). High infrastructure costs for transport and storage

further constrain large-scale deployment (Scottish Government, 2021).

3-2-3. Geological and Safety Constraints

Geological storage depends on subsurface integrity; long-term leakage cannot be entirely ruled out. Excessive injection pressure may trigger microseismic activity, necessitating detailed site surveys and continuous monitoring.

3-2-4. Socio-Economic Constraints

BECCS faces high capital costs and low short-term profitability—over $100 per ton CO_2 stored (Global CCS Institute, 2023). Moreover, public opposition to large-scale plantations or underground storage may limit social acceptance.

Thus, policy support, international cooperation, and community engagement are essential for expansion. In conclusion, BECCS holds transformative potential for sustainable energy systems, yet the success depends on a careful balance among technological efficiency, ecological sustainability, and social legitimacy.

4. Conclusion

BECCS integrates the principle of the natural carbon cycle into human energy systems, not merely reducing emissions but may actively remove atmospheric carbon. By coupling photosynthesis, combustion, CO_2 capture, and conversion technologies, BECCS offers new opportunities for industrial sustainability and climate mitigation. Accordingly, BECCS is recognized as a key pathway toward a carbon-negative society beyond neutrality.

However, achieving this ideal requires addressing major challenges—land and water use, high capture costs, energy losses, and storage security. Without balance between technological advancement and ecological integrity, its mitigation benefits will remain limited.

Future BECCS development must therefore extend beyond engineering innovation to encompass policy frameworks, ethical considerations, and societal cooperation. With joint efforts from governments, industries, and research institutions to secure sustainable biomass and safe storage infrastructures, BECCS could become a practical and scalable solution to the global climate crisis.

Example Research Questions

In addition to this question, you can also research other topics.

1. Calculate the annual net CO_2 reduction of a 100 MW BECCS plant using Miscanthus as feedstock, assuming 90% capture efficiency and including emissions from cultivation, transport, and storage.

- annual fuel consumption: $F = \frac{\text{power generation}(MWh)}{n \times H}$
- combustion $\mathrm{CO_2}$ Amount generated: $E_1 = F \times$ carbon content coefficient
- capture $\mathrm{CO_2}$: $C = 0.9 \times E_1$ (90% assumtion)
- Emissions from all processes: $E_{\text{entire}} = E_{\text{culture}} + E_{\text{transit}} + E_{\text{Preprocessing}} + E_{\text{Compression/Transport/Injection}}$
- net reduction: $C - E_{\text{entire}}$
- Data checklist: Emission factors for sorghum cultivation, fertilizer use, heat and electricity consumption in the recovery and regeneration stages, extrusion, transportation, and storage energy

Next, we will analyze the sensitivity of each variable by changing the input values of each variable (for example, collection efficiency 85% to 95%, transportation and pretreatment emissions ±50%) and checking how the annual net CO_2 reduction varies.

2. Propose international standards or sustainability certification systems to ensure BECCS deployment does not compromise food security or biodiversity. Recommended framework:

Land: Prohibit deforestation; prioritize residues and marginal lands.

Biodiversity: Exclude protected and buffer zones; maintain soil organic carbon.

Water: Assess regional water stress; minimize irrigation; include reuse plans.

Carbon Accounting: Require full life-cycle emission reporting with third-party verification (MRV).

Policy Linkage: Integrate with national emission trading and sustainability schemes.

3. Design a carbon-negative industry model applying BECCS principles to non-power sectors such as bioethanol refineries or paper mills. Consider CO_2 concentration, heat integration, compression/transport pathways, and MRV (Monitoring, Reporting, Verification) frameworks.

References

Drax Group, 2024*a*. BECCS at Drax. https://www.drax.com/beccs/.

Drax Group, 2024*b*. Bioenergy with carbon capture, use, and storage (BECCS). https://www.drax.com/about-us/bioenergy-carbon-capture-use-and-storage-beccs/.

Fajardy, M., and N. Mac Dowell. 2017. Can BECCS deliver sustainable and resource-efficient negative emissions? Energy & Environmental Science **10**:1389–1426.

Global CCS Institute, 2023. Global status of CCS 2023. https://www.globalccsinstitute.com/resources/global-status-of-ccs-2023/.

Intergovernmental Panel on Climate Change (IPCC). 2018. Global warming of 1.5℃: an IPCC special report on the impacts of global warming of 1.5℃ above pre-industrial levels and related global greenhouse gas emission pathways. Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC). 2019. Climate change and land: an IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems. Intergovernmental Panel on Climate Change.

Intergovernmental Panel on Climate Change (IPCC). 2021. Climate change 2021: the physical science basis. Cambridge University Press.

Rochelle, G. T. 2009. Amine scrubbing for CO_2 capture. Science **325**:1652–1654.

Scottish Government, 2021. Negative emissions technologies (NETS): feasibility study. https://www.gov.scot/publications/negative-emissions-technologies-feasibility-study/.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Youngsoo Kang

M. yskang@kentech.ac.kr

Prof. YoungSoo Kang at School of Energy Engineering of KENTECH. His research area is carbon neutrality climate technology; Solar CO2 conversion into liquid fuels, Low temperature, microwave absorption and supercritical fluid degradation of waste plastics into fuels like gasoline, diesel, hydrogen gas and high value carbon based materials. He is director of three big research projects “Leaders Project on Artificial Photosynthesis of CO2 Conversion into Liquid Fuels” (one million US dollars per year for 5 years), “Industrial Alchimist Project on Continuous Flow System for Mass Production of Liquid Fuels by CO2 Conversion” (4 million US dollars per year for 5years) and “Low Temperature Degradation of Waste Plastics into Fuels” (0.3 million US dollars, 3 years). He has worked as President of Korean Chemical Society, Physical Chemistry Division during 2023 and has worked as editorial board members of 5 international journals ; Materials Today Sustainability, Nanomaterials, Rare Metals, Molecules, Energies.

KENTECH
Korea Institute of Energy Technology

III

배터리 안정성의 핵심,
배터리 관리 시스템 (BMS)

p. 050 - 072

Vol. 10 (2025)

배터리 안정성의 핵심, 배터리 관리 시스템(BMS)

Abstract 본 탐구는 전기자동차(EV) 및 에너지저장시스템(ESS)의 핵심 구성 요소인 리튬이온 배터리의 안정성과 신뢰성을 확보하기 위해 필수적인 배터리 관리 시스템(BMS, Battery Management System)의 구조와 기능을 심층적으로 분석한다. 최근 EV 보급 증가와 함께 화재 및 열폭주 사례가 사회적 문제로 대두되며, 배터리 상태를 실시간으로 감지·제어하는 BMS의 중요성이 더욱 강조되고 있다. 본 연구에서는 BMS가 수행하는 전압·전류·온도 측정, 셀 밸런싱, 충·방전 제어, 보호 기능의 원리를 설명하고, SOC(State of Charge)와 SOH(State of Health) 추정 알고리즘의 차이와 한계를 비교·분석한다. 더불어 최근 배터리 수명 연장을 위해 제안된 DFC(Delayed Full Charging) 전략을 소개하며, 사용자 패턴 기반 충전 시점 조절이 고전압 노출 시간을 줄여 배터리 열화를 지연시킬 수 있음을 논의한다. 배터리 안전성 향상을 위한 BMS 기술의 필요성과 발전 방향을 제시하며, 향후 인공지능 기반 초개인화 충전 제어, V1G·V2G 연계, 다양한 환경 조건에서의 모델 정교화 가능성을 제안한다.

Introduction

최근 전기자동차 보급이 증가함에 따라 전기차 배터리 화재 사고 또한 사회적 문제로 대두되고 있다. 전기차 화재는 단순한 기기 고장을 넘어 대규모 재산 피해와 인명 피해로 이어질 수 있으며, 이로 인해 소비자들은 전기차에 대해 불신을 가지게 된다. 이러한 화재 사고의 주요 원인으로는 배터리 셀의 물리적 손상, 과충전 등 배터리 내부의 이상 현상이 지목되고 있다. 특히 리튬이온 배터리는 에너지 밀도가 높아 작은 이상도 큰 위험으로 확대될 수 있기 때문에, 배터리 상태를 정밀하게 감지하고 관리하는 시스템의 필요성이 더욱 강조되고 있다.

한편, 최근 재생에너지 확대 정책에 따라 간헐적으로 생산되는 전력을 효율적으로 저장하고 관리할 수 있는 ESS(Energy Storage System)의 중요성이 확대되고 있다. ESS 또한 대용량 리튬이온 배터리를 사용하기 때문에 전기차와 유사한 안전 문제가 발생할 수 있으며, 실제로 국내외 ESS 화재 사례가 보고되면서 배터리 관리의 중요성은 더욱 부각되고 있다. 전기차와 ESS는 운용 환경과 목적은 다르지만, 모두 고안전·고신뢰성의 배터리 관리 기술이 요구된다는 공통점을 가진다.

특히 최근에는 화재뿐 아니라 배터리 수명 저하 문제도 산업 전반에서 중요한 이슈로 다뤄지고 있다. 리튬이온 배터리는 고전압 상태에 오래 머무를수록 전해질 분해, SEI층 손상 등 열화가 가속되기 때문에, 사용자의 충전 패턴과 운용 조건을 정밀하게 제어하는 기술의 필요성이 커지고 있다. 이러한 수명·안전 문제를 완화하기 위한 전략 중 하나가 지연 완전 충전(Delayed Full Charging, DFC)이며, 이는 배터리가 고전압 상태에 머무는 시간을 최소화해 열화 속도를 줄이는 기술로 주목받고 있다.

따라서 본 탐구 자료에서는 전기차와 ESS 배터리의 안정적 운영을 위해 핵심적인 역할을 수행하는 BMS(Battery Management System)의 기능과 SOC/SOH 추정 방식, 그리고 배터리 수명 연장 전략인 DFC 기술까지 소개하고자 한다.

1. 배경지식: BMS

BMS는 'Battery Management System'의 약자로, EV(Electric Vehicle)에 탑재된 배터리 상태를 실시간으로 모니터링하고 제어하는 전자 제어 시스템을 의미한다. BMS의 핵심 역할은 배터리 셀의 전압, 전류, 온도 등 주요 데이터를 지속적으로 측정·분석하여 배터리가 안전하고 효율적으로 작동하도록 지원하는 것이다. 또한, 충·방전 과정의 제어, 셀 간 밸런싱, 이상 상태 감지 및 보호 기능 등을 통해 배터리 성능을

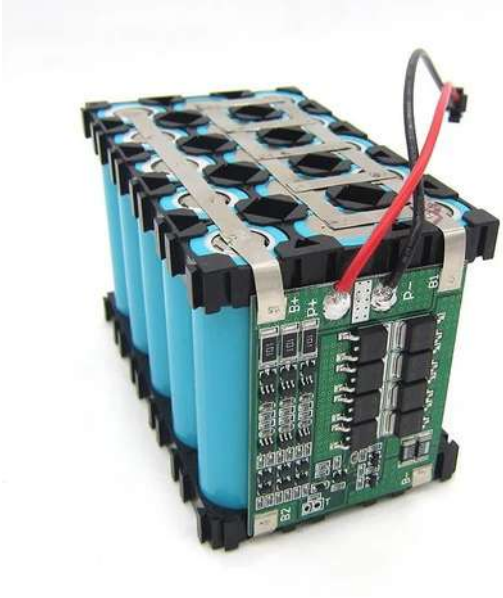


Figure 1. BMS with lithium battery

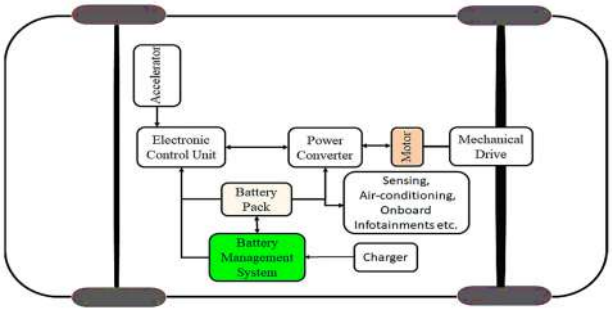


Figure 2. BMS in EV

극대화하고 수명을 연장하는 데 중요한 역할을 한다 (Habib et al., 2023; Ramkumar et al., 2022).

1-1. BMS의 주요 기능

전압, 전류, 온도 측정

각 BMS 배터리 셀과 전체 팩의 전압, 전류, 온도를 실시간으로 정밀하게 측정한다. 이를 통해 배터리 상태를 지속적으로 모니터링하며, 과충전, 과방전, 과열 등의 이상 상황을 조기에 감지하여 적절한 제어 및 보호 조치를 수행한다.

1-1-1. 셀 밸런싱(Cell Balancing)

셀 밸런싱은 배터리 팩 내 개별 셀 간 전압 차이를 조정해 모든 셀이 균일하게 충·방전되도록 하는 기능이다. 셀 간 불균형은 특정 셀의 조기 열화와 전체 성능 저하를 초래할 수 있으므로, BMS는 셀 밸런싱을 수행해 배터리 수명과 안전성을 향상시킨다.

1-1-2. 충방전 제어 및 보호

충전기 및 모터 제어기와 연동하여 충·방전 과정을 효율적으로 관리하고, 배터리를 보호한다. 과충전, 과방전, 과전류, 단락 등 위험 상황 발생 시 시스템을 차단하거나 경고 신호를 출력해 안전한 운용을 보장한다.

1-1-3. SOC(State of Charge) 및 *SOH(State of Health)* 계산

사용가능한 잔여 에너지를 나타내는 SOC와 배터리 성능의 열화 정도를 나타내는 SOH를 정밀하게 추정해 주행거리 예측과 유지보수 계획 수립에 필요한 정보를 제공한다.

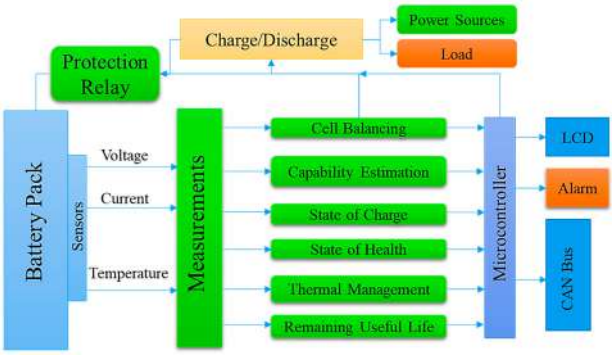


Figure 3. The role of BMS in EV

2. BMS의 기본 추정 요소

2-1. SOC

SoC는 'State of Charge'의 약자로, 말 그대로 '충전 상태'를 의미한다. 배터리의 잔존 용량을 나타내기 위해 현재 사용할 수 있는 배터리 용량을 전체 용량으로 나누어 백분율(%)로 표현한다(Dini et al., 2024).

2-1-1. 전류적산법*(Coulomb Counting Method)*

전류 적산법은 배터리로 들어오고 나가는 전류를 시간에 따라 적분하여 SOC를 추정하는 가장 널리 사용되는 방식이다. 이 방법은 단순하고 구현이 용이하다는 장점이 있지만, 측정 오차·기생 전류(parasitic current)·배터리 노화 및 온도 변화로 인한 유효 용량 변화 등의 요인으로 인해 누적 오차가 발생하기 쉽다. 이러한 문제를 줄이기 위해서는 주기적인 보정과 초기 값 업데이트가 필요하며, 특히 장기간 운용할 경우 정확도 유지를 위해 잦은 재보정이 요구된다. 전류적산법의 수학적 표현은 다음과 같다.

$$SOC(t) = SOC(t_0) + \frac{\int (I_b - I_{\text{loss}})\, dt}{C_{\text{rated}}}$$

t_0시점에서의 초기 충전상태(SOC)를 $SOC(t_0)$로 나타낸다. 여기서 I_b는 배터리 전류를 의미하며, I_{loss}는 손실 반응(loss reactions)에 의해 소비되는 전류를 의미한다. 또한 C_{rated}는 배터리의 정격 용량을 나타낸다.

2-1-2. 전압 기반 측정*(Open-Circuit Voltage(OCV) Method)*

전압 기반 SOC 추정은 배터리의 개방회로전압(OCV)을 측정한 뒤, 이를 사전 구축된 수학적 모델과 비교하여 SOC를 산출하는 방식이다. 이 방법은 비침습적이고 구조가 단순하다는 장점이 있지만, 온도 변화와 배터리 노화 특성 변화에 크게 영향을 받으며, 그 결과 SOC 추정의 정확도가 떨어질 수 있다. 또한 SOC와 OCV 사이의 비선형 관계 때문에 정확도를 확보하려면 세심한 보정이 필요하며, 특히 실제 주행과 같은 온도 변화가 잦은 환경에서는 더욱 중요하다. 배터리의 개방회로전압(OCV)과 충전상태(SOC) 사이의 관계를 모델링하기 위해 흔히 사용되는 수학적 함수는 시그모이드(sigmoid) 함수와 다항식(polynomial) 함수이며 다음과 같다.

$$SOC_{\text{sigmoid}} = \frac{1}{1 + e^{-(a \cdot OCV + b)}}$$

$$SOC_{\text{poly}} = \sum_{k=0}^{n} \alpha_k \, OCV^k$$

2-1-3. 칼만 필터*(Kalman Filter)*

이 방식은 칼만 필터와 같은 재귀 필터 알고리즘을 배터리의 수학적 모델과 결합하여 SOC를 추정하며 현재 많이 연구되고 있는 기법이다. 칼만 필터 기반 SOC 추정은 전류 적산법이나 전압법보다 더 높은 정확도를 제공하는 것으로 알려져 있다. 그러나 이 방법의 성능은 배터리 모델의 정확도에 크게 의존하기 때문에, 다양한 운용 조건에서 배터리의 특성과 거동을 정밀하게 이해하고 모델링해야 한다. 또한 칼만 필터는 계산량이 많아 실시간 처리가 필요한 자동차·산업용 임베디드 시스템과 같은 자원 제약 환경에서는 적용이 어려울 수 있다. 칼만 필터의 기본적인 상태 방정식은 다음과 같다.

$$x_k = Fx_{k-1} + Bu_k + w_k$$

x_k는 시점 k에서의 내부 에너지 상태이며 F는 시스템 동역학을 나타내는 상태 전이 행렬(state transition matrix)이다. B는 입력 u_k가 상태에 미치는 영향을 나타내는 제어 행렬(control matrix)이고 w_k는 평균이 0인 프로세스 노이즈 이다.

관측 모델(observation model)은 실제 측정된 SOC, SOH 등의 계측값과 내부 상태 간 관계를 정의하며, 다음과 같이 표현된다.

$$z_k = Hx_k + v_k$$

z_k는 시점 k에서의 실제 측정값이고 H는 상태를 측정공간으로 변환하는 관측 행렬(observation matrix)이며 v_k는 평균이 0인 측정노이즈이다.

2-1-4. 신경망 기반 알고리즘

신경망 기반 SOC 추정은 전류·전압·온도 등 배터리 데이터를 입력으로 사용하여 SOC와의 관계를 학습한 모델을 구축한 뒤, 실시간 또는 과거 데이터를 기반으로 SOC를 예측하는 방식이다. 신경망 알고리즘은 입력 변수 간 관계가 비선형적이거나 분석적 모델링이 어려운 복잡한 상황에서 전통적인 기법보다

높은 정확도를 제공할 수 있다. 하지만 이 알고리즘의 성능은 충분히 크고 다양한 데이터셋의 확보에 크게 의존하며, 학습 및 추론 과정에서 많은 계산 자원이 필요하기 때문에 자원이 제한된 시스템에서는 실제 적용에 어려움이 생길 수 있다.

2-1-5. 하이브리드 알고리즘

하이브리드 알고리즘은 두 가지 이상의 SOC 추정 방법을 결합하여 전체적인 정확도를 향상시키는 접근 방식이다. 예를 들어, 전류 적산법과 칼만 필터를 함께 사용하면 각 기법의 단점을 상호 보완하여 정확성과 강건성을 높일 수 있다. 이러한 복합적 접근은 개별 알고리즘의 한계를 완화하고 다양한 운용 조건에서도 더 신뢰할 수 있는 SOC 추정을 가능하게 한다.

2-2. SOH

배터리 SOH는 초기 출고 시점 대비 현재의 성능을 백분율로 나타내는 지표로 배터리의 열화 상태를 반영한다(Su et al., 2024).

2-2-1. 용량 기반 평가

용량 감소(capacity degradation)는 리튬이온 배터리의 SOH(상태 건강도)를 평가하는 데 가장 직접적이고 핵심적인 지표이다. 이는 장기간 사용에 따라 배터리가 에너지를 저장하고 방출하는 능력이 감소하는 것을 명확하게 반영하기 때문이다. SOH는 이러한 용량 감소를 기반으로 정의되며, 다음과 같이 표현된다.

$$SOH = \frac{Q_{\text{max}}}{Q_n} \times 100\%$$

여기서 Q_max는 현재 조건에서의 최대 사용 가능 용량, Q_n은 정격 용량을 의미한다.

2-2-2. 내부 저항 기반 평가*(Incremental Capacity Analysis (ICA))*

리튬이온 배터리의 노화는 내부저항 증가로 나타나는 것이 일반적인 특징이며, 이는 전기화학 시스템 내부의 다양한 임피던스 요소 변화에서 근본적으로 기인한다. 내부저항 증가 기반의 SOH 정의는 다음과 같은 식으로 표현된다.

$$SOH = \frac{R_E - R_{\text{real}}}{R_E - R_N}$$

R_E는 수명 종료(End of Life, EOL) 시의 내부저항이며, R_{real}은 현재 측정된 내부저항, R_N은 정격 내부저항을 의미한다. 일반적으로 R_E는 제조사에서 정의하며, 전기차용 배터리의 경우 보통 정격 내부저항 R_N의 두 배 정도로 설정된다. 한편 R_{real} 값은 펄스 시험(pulse testing)을 통해 추정된다.

2-2-3. 사이클 수명 기반 평가

충·방전 곡선(charge–discharge curves)은 배터리 SOH를 평가하는 데 중요한 통계적 특징들을 제공한다. 이러한 곡선에서 통계적 특징을 추출하는 과정은 주로 시간 영역(time-domain) 분석과 엔트로피(entropy)

기반 방법에 의존한다. 시간 영역 지표에는 분산(variance)과 표준편차(standard deviation)가 포함되며, 엔트로피 기반 지표로는 퍼지 엔트로피(fuzzy entropy)와 샘플 엔트로피(sample entropy)가 있으며, 이는 배터리의 비선형적 특성을 포착한다. 분산은 데이터의 산포 정도를 나타내는 전통적인 지표로, 전압 또는 전류의 안정성을 직접적으로 반영한다.배터리 내부에서 전극 물질 분포 불균일, 내부 접촉 불량 등의 문제가 발생하면 분산이 증가하며, 이러한 통계적 특징의 정량적 분석을 통해 배터리 건강 상태를 간접적으로 평가할 수 있다. 퍼지 엔트로피 (Fuzzy Entropy)는 비선형 동적 특성을 나타내는 지표로, 내부 화학 반응의 복잡성을 효과적으로 반영한다. 배터리가 노화됨에 따라 내부 반응 메커니즘이 더 복잡하고 불규칙해지며, 그 결과 퍼지 엔트로피의 값이 크게 변동한다. 이러한 변화는 배터리의 열화 상태를 정량적으로 평가하는 데 활용될 수 있다.

2-2-4. 열화 모델 기반 평가

온도는 리튬이온 배터리의 성능과 수명에 큰 영향을 미치며, 배터리 내부의 전기화학 반응, 열역학적 과정, 그리고 소재 특성에 변화를 일으킨다. 온도는 배터리 열화에 양면적 영향을 주는데, 높은 온도는 전극 소재의 열화, 전해액 분해, SEI 막 성장 등을 가속시키는 반면, 낮은 온도에서는 리튬 이온 이동도가 감소하고 분극이 증가한다. 따라서 배터리를 25℃에서 45℃ 사이의 최적 온도 범위에서 유지하는 것이 매우 중요하며, 이 범위를 벗어나면 용량 감소와 수명 단축이 가속될 수 있다. SOH를 정확하게 평가하기 위해서는 온도 관련 특성의 효과적인 모니터링과 분석이 필수적이다. DTV(Differential Thermal Voltammetry)는 전압–온도 결합 분석을 통해 배터리 열화 메커니즘을 식별하는 중요한 방법이다. 충·방전 과정에서의 전압 및 온도 데이터를 미분함으로써 생성되는 DTV 곡선은 내부 화학 반응의 발열량, 상변화(phase change) 등의 정보를 반영하며, 열화 메커니즘 분석에 도움을 준다.

3. 배터리의 수명 연장을 위한 기술: DFC

오래 유지하는 기술이다. 사용자의 충전 습관, 기상 시간 등의 사용 패턴을 파악하여 배터리가 완충 상태에서 충전되는 시간을 최소로 한다. 보통 80%까지 빠르게 충전한 다음 이를 유지 한다. 이후 파악한 사용자의 패턴에 맞춰 100% 충전된 상태 즉, 배터리에 고전압이 걸린 상태를 최소한으로 유지한다. 아래는 충전하는 기계에서의 DFC 알고리즘이다.

Algorithm 1 *DFC* Algorithm

Require: Current SOC *soc*, Extra time buffer t_{margin}
1: **Define:** Pre-departure time as 0.5 hour plus t_{margin} prior to the departure moment
2: **Initialization:** Start real-time pre-departure detection when EV is plugged in
3: **while** pre-departure time is not predicted **do**
4: **if** *soc* < 80% **then**
5: Proceed with slow charging
6: **else**
7: Pause charging
8: **end if**
9: **end while**
10: Initiate fast charging

Figure 4. The DFC algorithm of charging machine

3-1. DFC의 필요성

앞서 설명한 것처럼 배터리가 충전될수록 고전압 환경이 만들어진다. 고전압 상태는 내부의 전해질 분해를 가속 시킨다. 리튬 이온 배터리의 경우 셀당 최대 전압은 3.8V 4.7V이다. 이 범위를 넘게 되면 내부에서 산화가 발생하고 수명이 단축된다. 국립 재생 에너지 연구소(NREL)의 연구에 따르면 DFC를 통해 20% 80%의 충전 상태를 유지한 배터리의 수명이 증가한다. 배터리의 완전 충전 및 방전 싸이클의 횟수가 커질 수로 성능이 떨어진다(Lee et al., 2024).

과도한 충전은 배태리에 열을 발생시켜서 산화 반응 및 전해질의 분해를 가속한다. 분해된 전해질은 가스를 만들고 이는 폭발의 위험성도 갖는다. 배터리 내부 특히 음극 표면에 있는 SEI(Solid Electrolyte Interphase)층이라는 보호막을 분해, 변형시킴으로써 활성 리튬 이온의 수를 영구적으로 감소시킨다. 또한 내부의 구조를 변형시키거나 부산물을 만듦으로써 저항을 증가시킨다. 증가한 저항은 더 높은 열을 만드는 악순환이 반복된다.

3-1-1. 배터리 수명 측면

Hoke가 개발한 방법에 따르면 새벽에 충전하고 6.6kW로 80%까지 충전한 후 출발 30분 전 18kW의 빠른 속도로 충전하는 것이 다른 방법과 비교했을 때 50%가량 배터리의 수명을 늘리는 효과가 있다. 나아가 다음날 필요한 만큼만 충전하게 된다면 150% 정도 배터리의 수명이 늘어난다. 평균 SOC가 10% 감소한다면 배터리 수명은 8.95% 증가한다.

3-1-2. 경제적 측면

384개의 SOC 및 다양성 열화 요인 조합을 가진 시나리오에서 시뮬레이션 테스트한 결과 총 운영 비용이 15% 절감한 효과있다. 또한, 인프라 계획, 자원 배분, 소비자 만족도, 충전소의 수익성을 포함하는 스마트 그리드 현대화 설계의 필수적이다.

3-2. DFC가 배터리 수명에 미치는 영향

충전량, 충전 속도, 열을 관리하여 배터리의 수명을 늘려준다. 상변화 물질(PMS, Phase Change Material)을 활용하여 충전 중 발생하는 열을 관리한다. 두께를 늘릴수록 효과가 늘어나지만, 크기 및 두께의 한계가 있다. 팽창 흑연(EG, Expandable Graphite)를 6% 첨가하여 열전도율을 높여서 해결했다. 배터리가 완전 충전 및 방전이 싸이클의 횟수를 줄이기 위해 사용자의 패턴에 맞춰 충전량을 조절한다. 80%까지는 빠르게 충전한 후, 이 범위에서 머무르게 한다. 이후, 적절한 시간 대에 100%까지 충전한다. 이를 통해 배터리가 고전압 상태에 머무르는 시간을 최소화하여 전해질 분해, 구조 변형 등의 문제를 감소시킨다.

3-3. DFC의 발전 방향

현재는 주로 핸드폰에서 직접 최대 충전량을 설정하거나 충전 시간을 설정하는 등의 방식으로 상용되고 있다. 하지만, 이는 최적화와는 거리가 먼 기술이다. DFC는 인공지능과 함께 발전해 나가야 하는 기술이다. 더불어 초개인화를 위하여 GPS를 이용한 위치 정보, 웨어러블 기기를 활용한 생체 정보를 활용할 수 있다. 그리드와 연동시켜 전기 자동차를 효율적으로 충전할 수 있습니다. 예를 들면 밤 10시에 충전 플러그를 연결해도 요금이 싼 시간대가 심야 시간대인 것을 학습하고, 새벽에 충전을 시작하는 것이 있다. 나아가 새벽에 완충시킨 전기를 오전 피크 시간대에 다시 되팖으로써 경제적 이득 및 그리드 망을 안정화할 수 있다. 추가적으로 현재의 리튬 이온화 배터리가 아닌 다른 하드웨어에 맞춰서 사용될 수 있다.

Conclusions

이번 탐구 자료는 전기자동차의 보급률 증가와 에너지 저장 시스템의 안정성이 중요해짐에 따라, 두 시스템의 핵심 구성 요소인 리튬이온 배터리의 안전성과 신뢰성을 확보하기 위한 BMS에 대해 심층적으로 분석한다. 전기차 화재 및 ESS 사고 사례에서 보듯, BMS의 발전은 배터리 시스템의 안정성과 직결되는 매우 중요한 과제이다. BMS는 배터리의 전압, 전류, 온도를 실시간으로 모니터링한다. 이를 바탕으로 내부의 균형 및 충방전 제어를 통해 시스템의 안정성을 높이는 역할을 한다. 특히 이번 탐구 자료에서 상세히 다룬 SOC(충전 상태)와 SOH(건강 상태) 추정 기술은 BMS의 핵심이라고 할 수 있다. 전류적산법, OCV, 칼만 필터, 신경망같은 알고리즘 들은 배터리의 잔여 용량과 수명을 정확하게 예측하여 사용자의 편의성을 높이고 배터리 성능을 유지하기 위한 필수적인 기술이다. 더 나아가, DFC 기술 사례는 BMS가 단순히 배터리를 보호하는 것을 넘어, 사용자의 데이터와 패턴을 학습하여 배터리 수명을 연장하고 경제적 이득을 취할 수 있는 기술로 진화하고 있음을 보여준다. 결론적으로, BMS는 전기차, ESS의 안정성을 높이고. 배터리의 신뢰성을 보장하기 위한 핵심 기술이다. 앞으로 BMS 기술은 인공지능, 초개인화 사회, V1G, V2G 등과 결합하여 더욱 정교해질 것이다. 궁극적으로 이는 미래 친환경 에너지 및 모빌리티 시대를 앞당기는 견고한 기술적 기반이 될 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

***1. NASA*의 오픈 배터리 데이터셋을 활용해 직접 *SOC*와 *SOH*를 인공지능 모델*(e.g. LSTM)*을 이용하여 구해보자.**

***2. ESS*에서의 *BMS*는 *EV*의 *BMS* 구조와 어떤 점이 다를까?**

ESS와 EV의 배터리 동작 시간, 구성, 충전-방전 시간과 방법 등을 고려하여 생각해보자.

***3.* 계절의 변화로 인한 환경 온도의 변화가 *BMS*의 성능에 영향을 줄까? 영향을 준다면 어떻게 이를 방지할 수 있을까?**

***4.* 스마트폰 데이터, 생활 공간 조명의 변화와 같은 데이터를 활용해 개인의 생활패턴에 초점을 맞춘 *DFC* 시스템을 구성해보자.**

References

Dini, P., A. Colicelli, and S. Saponara. 2024. Review on Modeling and SOC/SOH Estimation of Batteries for Automotive Applications. Batteries **10**:34.

Habib, A. K. M., M. K. Hasan, G. Issa, D. Singh, S. Islam, and T. Ghazal. 2023. Lithium-Ion Battery Management System for Electric Vehicles: Constraints, Challenges, and Recommendations. Batteries **9**:152.

Lee, Y., W. Song, J. Song, and Y. Noh. 2024. Extending EV Battery Lifetime: Digital Phenotyping Approach for Departure Time Prediction. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **8**:1–30.

Ramkumar, M., C. Reddy, A. Ramakrishnan, K. Raja, S. Pushpa, S. Jose, and J. Mani. 2022. Review on Li-Ion Battery with Battery Management System in Electrical Vehicle. Advances in Materials Science and Engineering **2022**:1–8.

Su, L., Y. Xu, and Z. Dong. 2024. State-of-health estimation of lithium-ion batteries: A comprehensive literature review from cell to pack levels. Energy Conversion and Economics **5**:n/a–n/a.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Seok Gwang Doo

M. sgdoo@kentech.ac.kr

Gwang Doo received the B.S. degree in mineral & petroleum engineering from Seoul National University, Seoul in 1986, and the M.S. and Ph.D. degrees in Electrochemistry from the same university in 1988 and 1993, respectively. In 1993, he joined Samsung Advanced Institute of Technology, the central R&D center of Samsung Electronics, where he was Group Leader of Battery Group since 2009, and he became Lab Director of Energy Lab at SAIT. And in the end of 2018, he was transferred to Samsung SDI as head of advanced materials Group at R&D center. Currently he is a professor of Department of Energy Engineering at KENTCH and director of KENTECH Institute of Energy Materials and Devices. His current research interests include advanced Li-ion battery materials and next-generation batteries, such as all solid-state batteries, aqueous Zn ion batteries for EV's & ESS.

KENTECH
Korea Institute of Energy Technology

Vol. 10 (2025)

The Key to Battery Safety: Battery Management System (BMS)

Abstract This study provides an in-depth analysis of the structure and functions of the Battery Management System (BMS), an essential component for ensuring the stability and reliability of lithium-ion batteries in Electric Vehicles (EVs) and Energy Storage Systems (ESS). With the recent proliferation of EVs, incidents of fire and thermal runaway have emerged as significant social issues, underscoring the critical importance of BMS in monitoring and controlling battery status in real-time. This research elucidates the principles behind key BMS functions, including voltage, current, and temperature measurement, cell balancing, charge/discharge control, and protection mechanisms. Furthermore, it compares and analyzes the differences and limitations of State of Charge (SOC) and State of Health (SOH) estimation algorithms. Additionally, the study introduces the Delayed Full Charging (DFC) strategy, recently proposed to extend battery lifespan, and discusses how adjusting charging timing based on user patterns can mitigate battery degradation by reducing the duration of exposure to high voltage. Finally, this inquiry outlines the necessity and future development directions of BMS technology for enhanced battery safety, proposing possibilities for AI-based hyper-personalized charging control, integration with V1G/V2G systems, and model refinement under various environmental conditions.

Introduction

With the recent surge in the adoption of Electric Vehicles (EVs), battery fire incidents have emerged as a pressing social issue. These fires extend beyond simple mechanical malfunctions, often leading to extensive property damage and casualties, which in turn fosters consumer distrust in EVs. Internal battery anomalies, such as physical damage to battery cells and overcharging, are identified as the primary causes of these accidents. In particular, given the high energy density of lithium-ion batteries where even minor anomalies can escalate into major hazards the necessity for systems capable of precisely monitoring and managing battery status is becoming increasingly critical.

Concurrently, driven by recent policies to expand renewable energy, the importance of Energy Storage Systems (ESS) capable of efficiently storing and managing intermittently generated power is growing. Since ESS also utilizes high-capacity lithium-ion batteries, it poses safety risks similar to those of Electric Vehicles (EVs). Indeed, with reports of ESS fire incidents both domestically and internationally, the critical importance of battery management is being further underscored. Although EVs and ESS differ in their operating environments and purposes, they share a common requirement for high-safety and high-reliability battery management technologies.

Recently, in addition to fire risks, battery degradation has emerged as a critical issue throughout the industry. Since the degradation of lithium-ion batteries—caused by factors such as electrolyte decomposition and damage to the Solid Electrolyte Interphase (SEI) layer—accelerates the longer they remain in a high-voltage state, there is a growing necessity for technologies that precisely control user charging patterns and operating conditions. One strategy proposed to mitigate these lifespan and safety challenges is Delayed Full Charging (DFC). This technique is garnering attention for its ability to minimize the duration a battery remains at a high-voltage state, thereby decelerating the degradation process.

Therefore, this study aims to introduce the Battery Management System (BMS), which plays a pivotal role in the stable operation of EV and ESS batteries. Specifically, it covers the core functions of the BMS, methodolo-

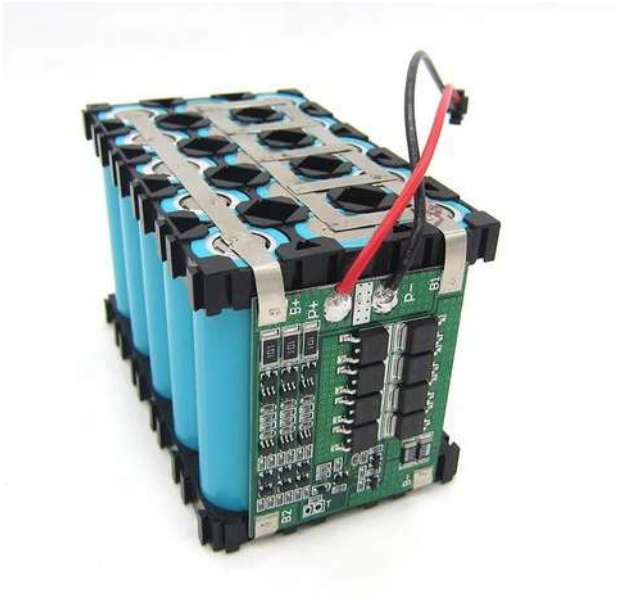


Figure 1. BMS with lithium battery

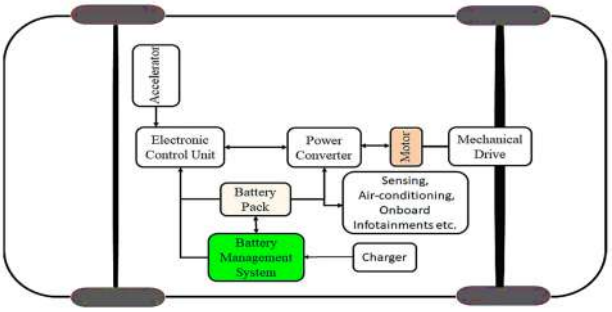


Figure 2. BMS in EV

gies for estimating SOC and SOH, and the Delayed Full Charging (DFC) technology as a strategic approach to extending battery lifespan.

1. Background knowledge : BMS

BMS, an acronym for 'Battery Management System,' refers to an electronic control system designed to monitor and regulate the status of Electric Vehicle (EV) batteries in real-time. The core function of the BMS is to ensure the safe and efficient operation of the battery by continuously measuring and analyzing critical data, including cell voltage, current, and temperature. Furthermore, it plays a vital role in maximizing battery performance and extending lifespan through mechanisms such as charge/discharge control, cell balancing, anomaly detection, and protection protocols. (Habib et al., 2023; Ramkumar et al., 2022).

1-1. Main Function of BMS

Voltage, Current, and Temperature Measurement

The BMS precisely measures the voltage, current, and temperature of individual battery cells and the entire pack in real-time. Through this process, it continuously monitors the battery's status and enables early detection of anomalies such as overcharging, overdischarging, and overheating, thereby executing appropriate control and protective measures.

1-1-1. Cell Balancing

Cell balancing is a function designed to adjust voltage differences among individual cells within a battery pack, ensuring that all cells are charged and discharged uniformly. As cell imbalance can lead to the premature degradation of specific cells and a decline in overall performance, the BMS performs cell balancing to enhance both battery lifespan and safety.

1-1-2. Charge,Discharge Control and Protection

It efficiently manages the charge and discharge processes and protects the battery by interfacing with chargers and motor controllers. Furthermore, it ensures safe operation by shutting down the system or issuing warning signals in the event of hazardous conditions such as overcharge, over-discharge, overcurrent, and short circuits.

1-1-3. Calcualtion of SOC(State of Charge) and SOH(State of Health)

It precisely estimates the State of Charge (SOC), which indicates the remaining usable energy, and the State of Health (SOH), which represents the degree of battery performance degradation, to provide the necessary information for predicting driving range and establishing maintenance plans.

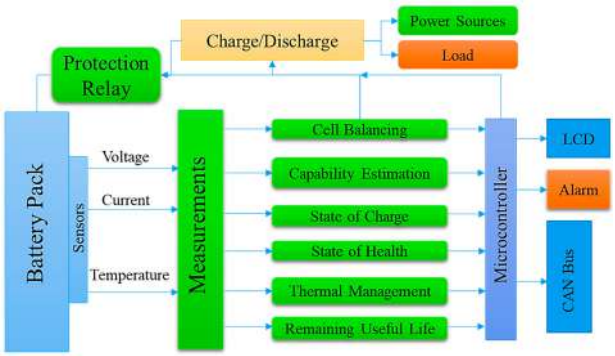


Figure 3. The role of BMS in EV

2. Basic Estimation Factor of BMS

2-1. SOC

SoC stands for 'State of Charge,' which literally refers to the charge status. It indicates the remaining battery capacity, expressed as a percentage(%) calculated by dividing the current usable capacity by the total capacity.(Dini et al., 2024).

2-1-1. Coulomb Counting Method

The Ampere-hour (Ah) counting method is the most widely used technique for estimating SOC by integrating the current flowing into and out of the battery over time. While this method offers the advantages of simplicity and ease of implementation, it is susceptible to cumulative errors caused by factors such as measurement errors, parasitic currents, and changes in effective capacity due to battery aging and temperature variations. To mitigate these issues, periodic calibration and initial value updates are necessary; in particular, frequent recalibration is required to maintain accuracy during long-term operation. The mathematical expression for the Ah counting method is as follows:

$$SOC(t) = SOC(t_0) + \frac{\int (I_b - I_{\text{loss}})\, dt}{C_{\text{rated}}}$$

The initial state of charge at time t_0 is denoted as $SOC(t_0)$. The variable I_b represents the battery current, whereas I_{loss} denotes the current consumed by parasitic loss reactions. The rated battery capacity is expressed as C_{rated}, which serves as the nominal capacity parameter for state-of-charge estimation.

2-1-2. Open-Circuit Voltage(OCV) Method

Voltage-based SOC estimation is a method that calculates SOC by measuring the battery's Open Circuit Voltage (OCV) and comparing it with a pre-established mathematical model. While this method offers the advantages of being non-invasive and simple in structure, it is significantly affected by temperature variations and changes in battery aging characteristics, which can compromise the accuracy of SOC estimation. Furthermore, due to the non-linear relationship between SOC and OCV, careful calibration is required to ensure accuracy; this is particularly critical in environments with frequent temperature fluctuations, such as real-world driving conditions. The mathematical functions commonly used to model the relationship between the battery's Open Circuit Voltage (OCV) and State of Charge (SOC) are the sigmoid function and the polynomial function, expressed as follows:

$$SOC_{\text{sigmoid}} = \frac{1}{1 + e^{-(a \cdot OCV + b)}}$$

$$SOC_{\text{poly}} = \sum_{k=0}^{n} \alpha_k \, OCV^k$$

2-1-3. Kalman Filter

This method estimates SOC by combining recursive filter algorithms, such as the Kalman filter, with the battery's mathematical model and is currently a subject of extensive research. Kalman filter-based SOC estimation is known to offer higher accuracy compared to the Ampere-hour counting method or voltage-based methods. However, since the performance of this method relies heavily on the accuracy of the battery model, it requires a precise understanding and modeling of the battery's characteristics and behavior under various operating conditions. Furthermore, due to its high computational load, the Kalman filter can be challenging to implement in resource-constrained environments such as automotive and industrial embedded systems that require real-time processing. The basic state equations of the Kalman filter are as follows:

$$x_k = Fx_{k-1} + Bu_k + w_k$$

The variable x_k denotes the internal energy state at time step k, and F represents the state transition matrix that characterizes the system dynamics. The matrix B is the control matrix describing how the input u_k influences the state, while w_k denotes the process noise, which is assumed to have zero mean.

The observation model defines the relationship between actual measured values, such as SOC and SOH, and the internal states, and is expressed as follows:

$$z_k = Hx_k + v_k$$

The variable z_k denotes the actual measurement obtained at time step k. The matrix H is the observation matrix that maps the system state to the measurement space, and v_k represents the measurement noise, which is assumed to have zero mean.

2-1-4. Neural Network based Algorithms

Neural network-based SOC estimation constructs a model that learns the relationship between SOC and battery inputs such as current, voltage, and temperature. it then predicts SOC based on real-time or historical data. Neural network algorithms can deliver higher accuracy than traditional techniques, particularly in complex scenarios where the relationships between input variables are non-linear or where analytical modeling is difficult. However, the performance of these algorithms relies heavily on the availability of sufficiently large and diverse datasets. Additionally, the high computational resources required for training and inference can pose challenges for practical implementation in resource-constrained systems.

2-1-5. Hybrid Algorithms

Hybrid algorithms represent an approach that combines two or more SOC estimation methods to enhance overall accuracy. For instance, integrating the Ampere-hour counting method with a Kalman filter allows the strengths of each technique to complement the other, thereby improving accuracy and robustness. This composite approach mitigates the limitations of individual algorithms and enables more reliable SOC estimation across a variety of operating conditions.

2-2. SOH

Battery SOH is an indicator that expresses current performance as a percentage relative to the initial state at the time of shipment, thereby reflecting the battery's degradation status.(Su et al., 2024).

2-2-1. Capacity based Evaluation

Capacity degradation is the most direct and critical indicator for evaluating the State of Health (SOH) of lithium-ion batteries. This is because it clearly reflects the diminishing ability of the battery to store and release energy over long-term use. SOH is defined based on this capacity degradation and is expressed as follows:

$$SOH = \frac{Q_{\max}}{Q_n} \times 100\%$$

In this context, $Q_{\max}$ refers to the maximum available capacity under the given operational conditions, whereas Q_n denotes the nominal rated capacity of the battery.

2-2-2. Incremental Capacity Analysis (ICA)

The aging of lithium-ion batteries typically manifests as an increase in internal resistance, which is fundamentally attributed to changes in various impedance elements within the electrochemical system. The definition of SOH based on this increase in internal resistance is expressed as follows:

$$SOH = \frac{R_E - R_{\text{real}}}{R_E - R_N}$$

The parameter R_E denotes the internal resistance at the end of life (EOL). The quantity R_{real} represents the currently measured internal resistance, while R_N indicates the nominal internal resistance. Typically, R_E is defined by the manufacturer and, in the case of electric vehicle batteries, is often set to approximately twice the nominal resistance R_N. The value of R_{real} is generally estimated through pulse testing.

2-2-3. Cycle Life Based Evaluation

Charge-discharge curves offer critical statistical features for evaluating battery SOH. The process of extracting statistical features from these curves primarily relies on time-domain analysis and entropy-based methods. Time-domain indicators include variance and standard deviation, while entropy-based indicators—such as fuzzy entropy and sample entropy—capture the non-linear characteristics of the battery. Variance, a traditional metric indicating data dispersion, directly reflects the stability of voltage or current. Variance increases when internal issues arise, such as non-uniform distribution of electrode materials or poor internal contact. Consequently, the quantitative analysis of these statistical features allows for the indirect evaluation of the battery's health status. Fuzzy Entropy serves as an indicator of non-linear dynamic characteristics, effectively reflecting the complexity of internal chemical reactions. As the battery ages, internal reaction mechanisms become increasingly complex and irregular, resulting in significant fluctuations in fuzzy entropy values. These variations can be utilized to quantitatively assess the battery's degradation state.

2-2-4. Degradation Model Based Evaluation

Temperature significantly influences the performance and lifespan of lithium-ion batteries, inducing changes in internal electrochemical reactions, thermodynamic processes, and material properties. Temperature exerts a dual impact on battery degradation: high temperatures accelerate electrode material degradation, electrolyte decomposition, and SEI layer growth, while low temperatures reduce lithium-ion mobility and increase polarization. Consequently, maintaining the battery within the optimal temperature range of 25℃ to 45℃ is critical, as operating outside this range can accelerate capacity loss and shorten lifespan. Effective monitoring and analysis of temperature-related characteristics are essential for the accurate evaluation of SOH. Differential

Thermal Voltammetry (DTV) is a key method for identifying battery degradation mechanisms through combined voltage-temperature analysis. DTV curves, generated by differentiating voltage and temperature data during charge and discharge processes, reflect information such as heat generation from internal chemical reactions and phase changes, thereby facilitating the analysis of degradation mechanisms.

3. Technology for Battery Lifespan Extension: DFC

This technique is designed to prolong battery lifetime by minimizing the duration during which the battery remains at a fully charged state. By analyzing user specific behavioral patterns, such as charging habits and wake up times, the system limits the time the battery stays at high voltage. Typically, the battery is rapidly charged up to approximately 80% and then maintained at this level. Subsequently, the system completes the remaining portion of the charge to reach 100% in a way that ensures the fully charged state is sustained for the shortest possible duration, based on the predicted user usage pattern. The following describes the DFC (Dynamic Full Charge) algorithm implemented in charging systems.

Algorithm 1 *DFC* Algorithm

Require: Current SOC *soc*, Extra time buffer t_{margin}
Define: Pre-departure time as 0.5 hour plus t_{margin} prior to the departure moment
Initialization: Start real-time pre-departure detection when EV is plugged in
while pre-departure time is not predicted **do**
    if *soc* < 80% **then**
        Proceed with slow charging
    else
        Pause charging
    end if
end while
Initiate fast charging

Figure 4. The DFC algorithm of charging machine

3-1. Necessity of DFC

As previously explained, a high-voltage environment is created as the battery charges. This high-voltage state accelerates the decomposition of the internal electrolyte. For lithium-ion batteries, the maximum voltage per cell typically ranges from 3.8V to 4.7V. Exceeding this range triggers internal oxidation and shortens the battery's lifespan. According to research by the National Renewable Energy Laboratory (NREL), the lifespan of a battery is extended when its State of Charge (SOC) is maintained between 20% 80% through DFC. Furthermore, performance degrades as the frequency of full charge and discharge cycles increases.(Lee et al., 2024).

Overcharging generates heat within the battery, thereby accelerating oxidation reactions and the decomposition of the electrolyte. This decomposed electrolyte produces gas, which poses a risk of explosion. Furthermore, by degrading or deforming the Solid Electrolyte Interphase (SEI) layer, a protective film on the anode surface, overcharging permanently reduces the number of active lithium ions. It also increases internal resistance by altering the internal structure or forming byproducts. This increased resistance subsequently generates higher heat, perpetuating a vicious cycle.

3-1-1. Battery Lifespan Perspective

According to the method developed by Hoke, charging at dawn to 80% at 6.6 kW, followed by rapid charging at 18 kW just 30 minutes prior to departure, is effective in extending battery lifespan by approximately 50% compared to other methods. Furthermore, if the battery is charged only to the level required for the next day's use, the lifespan can increase by roughly 150%. Specifically, a 10% reduction in the average State of Charge (SOC) leads to an 8.95% increase in battery lifespan.

3-1-2. Economic Perspective

Simulation tests conducted on scenarios featuring 384 combinations of SOC and various degradation factors demonstrated a 15% reduction in total operating costs. Furthermore, it is essential for the design of smart grid modernization, encompassing infrastructure planning, resource allocation, consumer satisfaction, and the profitability of charging stations.

3-2. Impact of DFC on Battery Lifespan

It extends battery lifespan by regulating charge levels, charging speeds, and heat. Phase Change Materials (PCM) are utilized to manage the heat generated during the charging process. While increasing the material's thickness enhances its effectiveness, there are practical limitations regarding size and thickness. To address this, 6% Expandable Graphite (EG) was added to increase thermal conductivity. Furthermore, the charging level is adjusted based on user patterns to reduce the frequency of full charge and discharge cycles. The battery is rapidly charged to 80% and maintained in this range. Subsequently, it is fully charged to 100% at the most appropriate time. This approach minimizes the duration the battery remains in a high-voltage state, thereby reducing issues such as electrolyte decomposition and structural deformation.

3-3. Future Outlook for DFC

Currently, commercial applications are primarily limited to users manually setting maximum charge limits or scheduling charging times via mobile phones. However, this approach falls short of true optimization. DFC is a technology that must evolve in tandem with artificial intelligence (AI). Furthermore, to achieve hyper-personalization, the system can utilize location data via GPS and biometric data from wearable devices. By integrating with the power grid, DFC enables the efficient charging of electric vehicles. For instance, even if the vehicle is plugged in at 10:00 PM, the system can learn that electricity rates are lower during late-night hours and automatically delay the start of charging until dawn. Moreover, by selling the electricity fully charged at dawn back to the grid during morning peak hours, it can generate economic benefits and contribute to grid stabilization. Finally, this technology holds the potential to be adapted for future hardware beyond current lithium-ion batteries.

Conclusions

This inquiry provides an in-depth analysis of the Battery Management System (BMS), a critical component for ensuring the safety and reliability of lithium-ion batteries in Electric Vehicles (EVs) and Energy Storage Systems (ESS). This analysis is particularly timely given the rising adoption of EVs and the growing importance of ESS stability. As evidenced by instances of EV fires and ESS accidents, the advancement of BMS technology is a crucial task directly linked to the stability of battery systems. The BMS monitors battery voltage, current, and temperature in real-time. Based on this data, it enhances system stability through internal balancing and the control of charge and discharge processes.

In particular, the estimation technologies for State of Charge (SOC) and State of Health (SOH), discussed in detail in this study, constitute the core of the BMS. Algorithms such as the Ampere-hour counting method, OCV, Kalman filters, and neural networks are essential technologies. They accurately predict remaining battery capacity and lifespan, thereby enhancing user convenience and maintaining battery performance. Furthermore, the case of DFC technology demonstrates that the BMS is evolving beyond simple battery protection. It is advancing into a technology capable of extending battery lifespan and securing economic benefits by learning from user data and usage patterns. In conclusion, the BMS is a pivotal technology for enhancing the stability of EVs and ESS and for guaranteeing battery reliability. Looking ahead, BMS technology is set to become even more sophisticated by integrating with artificial intelligence, the hyper-personalized society, and V1G/V2G systems. Ultimately, this will serve as a solid technological foundation that accelerates the advent of the future eco-friendly energy and mobility era.

Example Questions

Research may be conducted freely on any topic, including but not limited to the items listed below.

1. Using NASA's open battery datasets, develop an artificial intelligence model (e.g., LSTM) to estimate the State of Charge (SOC) and the State of Health (SOH).

2. How does the Battery Management System (BMS) in an Energy Storage System (ESS) differ from that in an Electric Vehicle (EV)?

Consider aspects such as operating duration, system configuration, and charging–discharging methods and times for ESS and EV applications.

3. Do seasonal variations in ambient temperature influence the performance of a BMS? If so, discuss possible mitigation strategies to reduce such effects.

4. Design a Dynamic Full Charge (DFC) system that leverages personal lifestyle data such as smartphone usage patterns or changes in indoor lighting conditions to tailor charging behavior to individual daily routines.

References

Dini, P., A. Colicelli, and S. Saponara. 2024. Review on Modeling and SOC/SOH Estimation of Batteries for Automotive Applications. Batteries **10**:34.

Habib, A. K. M., M. K. Hasan, G. Issa, D. Singh, S. Islam, and T. Ghazal. 2023. Lithium-Ion Battery Management System for Electric Vehicles: Constraints, Challenges, and Recommendations. Batteries **9**:152.

Lee, Y., W. Song, J. Song, and Y. Noh. 2024. Extending EV Battery Lifetime: Digital Phenotyping Approach for Departure Time Prediction. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies **8**:1–30.

Ramkumar, M., C. Reddy, A. Ramakrishnan, K. Raja, S. Pushpa, S. Jose, and J. Mani. 2022. Review on Li-Ion Battery with Battery Management System in Electrical Vehicle. Advances in Materials Science and Engineering **2022**:1–8.

Su, L., Y. Xu, and Z. Dong. 2024. State-of-health estimation of lithium-ion batteries: A comprehensive literature review from cell to pack levels. Energy Conversion and Economics **5**:n/a–n/a.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Seok Gwang Doo

M. sgdoo@kentech.ac.kr

Gwang Doo received the B.S. degree in mineral & petroleum engineering from Seoul National University, Seoul in 1986, and the M.S. and Ph.D. degrees in Electrochemistry from the same university in 1988 and 1993, respectively. In 1993, he joined Samsung Advanced Institute of Technology, the central R&D center of Samsung Electronics, where he was Group Leader of Battery Group since 2009, and he became Lab Director of Energy Lab at SAIT. And in the end of 2018, he was transferred to Samsung SDI as head of advanced materials Group at R&D center. Currently he is a professor of Department of Energy Engineering at KENTCH and director of KENTECH Institute of Energy Materials and Devices. His current research interests include advanced Li-ion battery materials and next-generation batteries, such as all solid-state batteries, aqueous Zn ion batteries for EV's & ESS.

KENTECH
Korea Institute of Energy Technology

IV

비귀금속 촉매 기반 수전해 기술 -

분리막과 나노 관점에서 이해하는 그린수소 생산

p. 074 - 095

Vol. 01 (2026)

비귀금속 촉매 기반 수전해 기술- 분리막과 나노 관점에서 이해하는 그린 수소 생산

Abstract

수전해 기술은 전기를 이용해 물을 분해함으로써 수소를 생산하는 방법으로, 이산화탄소를 직접 배출하지 않는다는 점에서 차세대 그린 수소 생산 기술로 주목받고 있다. 특히 재생에너지와 결합할 경우, 수전해는 변동적인 전력을 저장 가능한 화학 에너지 형태로 전환하는 핵심 수단이 될 수 있다. 그러나 수전해 기술은 종종 촉매 성능 중심으로 단순화되어 설명되며, 실제 장치에서 작동하는 구조적 특성은 상대적으로 덜 조명되는 경향이 있다. 본 탐구자료에서는 수전해 반응이 촉매 단독이 아닌, MEA(Membrane Electrode Assembly) 안에서 이루어진다라는 점에 주목한다. 특히 귀금속 촉매에서 비귀금속 촉매로의 전환 과정에서, 분리막이 반응 환경과 이온 전달을 결정하는 핵심 요소로 작용함을 설명한다. 비귀금속 촉매는 비용과 자원 측면에서 높은 잠재력을 지니지만, 산성 환경에서의 불안정성으로 인해 알칼라인 조건과 음이온 전달 분리막과의 결합이 필수적이다. 이를 바탕으로 본 글은 수전해 장치의 기본 구조, 귀금속 및 비귀금속 촉매의 특성 차이, 그리고 삼상 경계 개념을 통해 비귀금속 촉매 수전해가 '촉매 선택의 문제'를 넘어 '반응 환경과 구조 설계의 문제'임을 제시한다. 이러한 구조적 이해는 이후 비귀금속 촉매 기반 수전해 기술의 잠재력과 한계, 그리고 이를 극복하기 위한 재료과학적 연구 흐름을 이해하기 위한 기초를 제공한다.

Introduction

수전해 기술은 전기를 이용해 물을 분해함으로써 수소를 생산하는 방법으로, 수소를 장기 저장 가능한 매개체로 활용할 수 있다는 장점도 주목받고 있다. 특히 재생에너지와 결합할 경우, 이산화탄소를 직접 배출하지 않는다는 점에서 차세대 수소 생산 기술이라는 점도 주목받는 바이다.

이러한 설명에서 수전해는 종종 단순한 화학 반응처럼 소개된다. 물에 전압을 걸면 수소와 산소가 발생하고, 촉매를 사용하면 반응이 더 잘 일어난다라는 설명이 일반적이다. 하지만 실제 수전해 장치는 이러한 간단한 설명보다 훨씬 복잡한 구조를 가진다.

더불어 수전해의 성능과 안정성은 단순히 전압이나 촉매 성능만으로 결정되지 않는다. 촉매가 어떤 구조 안에 놓여 있고, 어떤 환경에서 작동하는지가 수전해 기술의 핵심이다. 수전해는 '좋은 촉매를 선택하는 문제'를 넘어, 반응이 일어나는 환경과 구조를 설계하는 시스템 공학적 문제로 이해되어야 한다. 이러한 관점은 귀금속 촉매에서 비귀금속 촉매로의 전환 과정에서 더욱 중요해진다. 비귀금속 촉매는 비용과 자원 측면에서 높은 잠재력을 지니지만, 기존의 산성 환경에서는 안정적으로 작동하기 어렵다. 따라서 비귀금속 촉매의 활용 가능성은 촉매 물성 자체보다도, 이를 가능하게 하는 분리막과 이온 전달 구조, 즉 수전해 장치 전체의 설계 방식에 의해 좌우된다.

이에 본 탐구 자료는 수전해 반응이 촉매 단독이 아닌 전극–촉매–분리막이 결합된 구조(Membrane Electrode Assembly, MEA) 안에서 이루어진다는 점에 주목하고, 분리막이 이온 전달과 pH 환경을 통해 반응 조건을 어떻게 규정하는지를 중심으로 비귀금속 촉매 기반 수전해 기술을 구조적·나노적 관점에서 분석하고자 한다.

1. 배경지식 : 그린 수소와 수전해

1-1. 그린 수소란 무엇인가?

기후 위기 해결을 위한 탄소 중립의 핵심 에너지원으로 '수소'가 주목받고 있다. 하지만 생산 방식에 따라 친환경성에는 큰 차이가 있다. 천연가스를 개질하여 만드는 '그레이 수소'는 필연적으로 이산화탄소(CO_2)를 배출한다. 반면, 태양광이나 풍력 같은 재생에너지 전력으로 물(H_2O)을 전기분해하여 얻는 '그린 수소(Green Hydrogen)'는 생산 과정에서 오직 산소(O_2)만 배출하는 가장 청정한 에너지원이다 (International Energy Agency, 2019).

1-2. 물을 쪼개는 기술, 수전해(Water Electrolysis)

수전해는 전기화학적 반응을 통해 물을 수소와 산소로 분리하는 기술이다. 열역학적으로 물은 매우 안정한 물질이기에, 이를 분해하려면 외부에서 에너지를 가해야 한다.

$$2\,H_2O(l) \rightarrow 2\,H_2(g) + O_2(g)$$

이때 필요한 에너지를 전기의 형태로 공급하는 것이 수전해다. 핵심은 이론적으로 필요한 전압(1.23V)보다 실제로는 더 높은 전압(과전압)이 필요하다는 점인데, 이 과전압을 얼마나 낮추느냐가 기술력의 척도가 된다 (Zeng and Zhang, 2010).

2. 수전해 장치의 심장 : MEA

2-1. MEA(Membrane Electrode Assembly)의 결합 구조

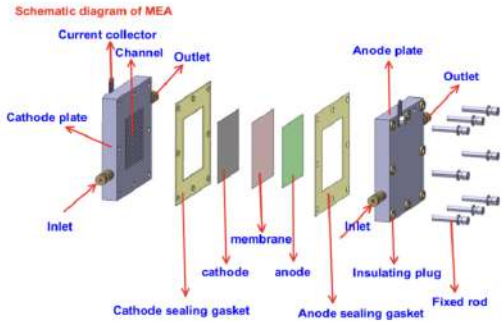


Figure 1. MEA 장치의 구조

수전해 셀의 핵심 부품인 MEA는 전극, 촉매, 분리막이 샌드위치처럼 압착된 구조다.
• 전극(Electrode): 외부 도선에서 오는 전자를 공급하거나 내보내는 집전체(Current Collector) 역할을 한다. 기체가 잘 빠져나가도록 다공성 구조를 가진다.
• 전극 촉매(Catalyst): 수전해 반응이 일어나는 부분으로, 촉매를 이용하여 반응의 활성화 에너지를 낮춰 반응 속도를 높인다. 촉매가 없으면 물 분해 반응은 너무 느려서 실용적이지 않다.
• 분리막(Membrane): 양극과 음극을 물리적으로 격리하여 생성된 수소와 산소 기체가 섞여 폭발하는 것을 막는다. 동시에 특정 이온만을 선택적으로 통과시켜 전기 회로를 완성한다 (O'Hayre et al., 2016).

2-2. 분리막 : 이온 전도와 pH 환경의 결정자

분리막은 단순한 차단막이 아니다. 전기화학 반응이 지속되려면 전자는 도선을 통해, 이온은 분리막을 통해 이동하며 닫힌 회로(Closed Loop)를 형성해야 한다. 또한 분리막의 이온 교환 특성(H^+ 전달 또는 OH^- 전달)에 따라 전해질의 pH가 결정된다.

• 양이온 교환막(PEM): H^++를 전달하며, 강한 산성 환경을 만든다.
• 음이온 교환막(AEM): OH^-를 전달하며, 알칼라인(염기성) 환경을 만든다. 이 pH 환경은 어떤 금속 촉매를 사용할 수 있는지를 결정하는 가장 중요한 변수가 된다 (Merle et al., 2011).

3. AEM과 비귀금속 촉매

귀금속 촉매에서 비귀금속 촉매(Non-Precious Metal Catalyst, NPMC)로의 전환은 단순한 비용 절감 전략이 아니라, 그린 수소 생산을 실험실 규모에서 기가와트(GW)급 산업 규모로 확장하기 위한 구조적 전환을 의미한다.
현재 상용 수전해 기술은 촉매 소재의 희소성과 공급 제약으로 인해 대규모 확장이 근본적으로 제한되고 있다. 니켈(Ni), 철(Fe), 코발트(Co) 등 전이금속 기반의 비귀금속 촉매가 연구되고 있지만 이 금속들은 PEM의 조건인 산성에서 견디지 못한다.
그래서 등장한 것이 알칼라인 환경을 이용하는 AEM(음이온 교환막)수전해이고, 이번 자료에서 가장 중점으로 확인할 기술이다. 알칼라인 조건에서는 비귀금속 촉매도 부식되지 않고 안정적으로 작동할 수 있다 (Zou and Zhang, 2015). 즉, 값싼 촉매를 쓰기 위해 시스템 전체의 화학적 환경(pH)과 이온 이동 방식($H^+ \rightarrow OH^-$)을 바꾼 것이 비귀금속 수전해 기술의 핵심 전략이다.
비귀금속 촉매는 이러한 병목을 해결할 수 있는 현실적인 대안으로 부상하고 있다.풍부한 매장량과 낮은 가격, 그리고 합금·나노구조화를 통한 성능 조절 가능성 덕분에 차세대 수전해 시스템의 핵심 구성요소로 주목받는다.
그러나 동시에 활성, 안정성, 전기전도성 측면에서 비귀금속 보다는 귀금속이 우위에 있다는 한계를 내포하고 있기에 이를 극복하지 못할 경우 산업적 적용에는 한계가 존재한다. 이 장에서는 비귀금속 촉매의 잠재력과 더불어, 그 이면에 존재하는 기술적 딜레마를 체계적으로 분석한다.

3-1. AEM과 비귀금속 촉매 기술의 잠재력 : PEM과 귀금속 촉매 기술에 비교하여

AEM에 활용될 수 있는 비귀금속 촉매의 도입은 수소 경제의 세 가지 핵심 축인 경제성, 확장성, 그리고 자원 안보 측면에서 막대한 잠재력을 지닌다.
경제성 측면에서, 수전해 스택(stack) 비용이 약 35–45% 이상이 촉매와 전극 소재 마련에서 발생한다. 현재 PEM 수전해에 필수적인 귀금속 이리듐(Ir)은 지구 연간 생산량이 약 7–8톤에 불과하며, kg당 가격이 수천만 원을 호가한다. 반면 AEM 수전해에 사용되는 비귀금속인 니켈(Ni), 철(Fe), 코발트(Co)등 비교적 흔한 금속으로 대체할 경우 kg당 수만 원 수준으로 촉매 원가를 이론적으로 100–1000배 이상 절감할 수 있다. 국제 에너지 기구와 DOE 분석에 따르면, 이러한 촉매 비용 절감은 장기적으로 봤을 때 최종 수소 생산 단가(LCOH, Levelized Cost of Hydrogen)를 kg당 1.5–2.0 USD 미만으로 낮추는 핵심 동력이 된다.
확장성 측면에서 비귀금속 촉매는 대규모 설비 확장에 필수적인 재료 접근성을 제공한다. 귀금속 기반 PEM 수전해는 설비 용량이 수백 MW를 초과할 경우 촉매 수급 자체가 병목으로 작용한다. 반면 AEM 비귀금속 기반 촉매는 그 원료 금속이 전 세계적으로 매장량이 풍부하여, 수 GW급 설비로의 확장이 자원적 제약 없이 가능하다. 공급망 안정성 및 자원 안보측면에서도 비귀금속 촉매는 전략적 의미를 갖는다. 귀금속인 백금족 금속은 특정 국가(남아공, 러시아 등)에 매장량이 집중되어 있어 지정학적 리스크가 크다. 반면 Ni, Fe, Mn 등 비귀금속 촉매의 원료 금속들은 전 세계적으로 고르게 분포되어 있어 탄소 중립 국가들이 자국 기술로 독자적인 에너지 공급망을 구축할 수 있게 한다.
시스템 설계 유연성 또한 중요한 장점이다. 비귀금속 촉매는 주로 알칼리성 환경(AEC, AEMWE)에서 안정적으로 작동하며, 이는 산성에서의 부식을 고려하여 사용해야하는 고가의 티타늄 분리판 대신 산성에 강한 스테인리스강, 니켈 합금, 고분자 소재를 사용할 수 있게 한다. 결과적으로 촉매 비용 절감이 시스템 하드웨어 비용 절감으로 연쇄 확산되는 구조를 형성한다.

3-2. 기술적 딜레마 및 한계점

비귀금속 촉매는 경제성과 확장성 측면에서 큰 잠재력을 지니고 있음에도 불구하고, 아직까지 귀금속 촉매를 전면적으로 대체하지는 못하고 있다.
그 근본적인 이유는 촉매 설계에서 중요한 세 가지 요소, 즉 촉매 활성(Activity), 화학적·전기화학적 안정성이 귀금속보다 현저히 떨어지기 때문이다. 일반적으로 촉매의 활성을 높이기 위해서는 표면 반응성이 큰 불안정한 상태를 유도해야 하지만, 이러한 상태는 장기 운전에서 촉매의 용출이나 구조 붕괴로 이어질 가능성이 높다. 이러한 측면이 비귀금속 촉매 상용화를 가로막는 핵심 기술적 딜레마이다(Seh et al., 2017).

3-3-1. 산성 환경에서의 열역학적 불안정성

가장 큰 한계는 화학적 안정성이다. 현재 가장 효율이 좋은 PEM 수전해는 강산성(pH <1)환경에서 작동하는데, 대부분의 비귀금속(Ni, Fe, Mn 등)은 이 환경에서 이온 형태로 용출(Leaching)되어 사라진다. 이는 촉매의 수명을 급격히 단축시키며, 용출된 금속 이온이 교환막을 오염시켜 시스템 전체의 성능저하를 유발하는 딜레마를 발생시킨다. 이 현상은 단순히 촉매 질량이 줄어드는 문제에 그치지 않는다. 용출된 금속 이온은 고분자 전해질막 내부로 이동해 이온 교환 기능을 저해하거나 막을 오염시키며, 그 결과 셀 저항 증가, 전압 상승, 효율 저하로 이어진다. 즉, 비귀금속 촉매의 불안정성은 촉매 단일 소재의 문제가 아니라 시스템 전체의 수명과 신뢰성을 동시에 위협하는 요인으로 작용한다.

3-3-2 산소 발생 반응*(OER)*의 높은 과전압

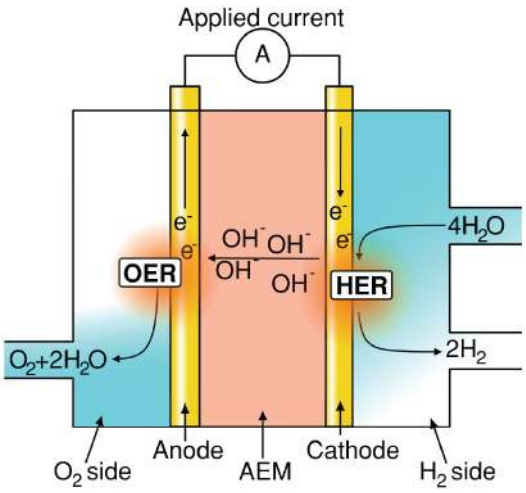


Figure 2. OER and HER in AEM

수전해는 전기를 이용해 물(H_2O)을 수소(H_2)와 산소(O_2)로 분해하는 기술이다. 이 과정은 두 개의 반쪽 반응으로 나뉘어 진행되는데, 전극 한쪽에서는 물로부터 수소를 만드는 수소 발생 반응(HER)이 일어나고, 다른 한쪽에서는 물로부터 산소를 만드는 산소 발생 반응(OER)이 일어난다.
HER는 비교적 단순한 전자 이동 과정을 가지기 때문에 반응이 쉽게 일어나지만, OER은 하나의 산소 분자를 만들기 위해 4개의 전자와 양성자가 순차적으로 이동해야 하는 복잡한 반응이다. 이 때문에 OER은 더 많은 에너지를 필요로 하며, 수전해 전체 효율을 결정짓는 핵심 단계로 작용한다.
따라서 우리가 주목해야하는 부분은 수전해 전체 효율을 결정짓는 산소 발생 반응이다. 비귀금속 촉매는 수소 발생 반응(HER)에서는 귀금속과 대등한 성능을 보이나, 산소 발생 반응에서는 복잡한 4전자 이동

경로로 인해 훨씬 높은 에너지 장벽(과전압)을 갖는다. 이로 인해 투입되는 전기에너지 중 열로 손실되는 비중이 커지며 전체 에너지 효율(LHV 기준)이 저하되는 기술적 한계가 있다.
수전해 시스템의 전체 에너지 효율을 결정짓는 반응은 산소 발생 반응(Oxygen Evolution Reaction, OER)이다. OER은 물 분자로부터 산소를 생성하는 과정에서 4개의 전자와 양성자가 순차적으로 이동하는 복잡한 반응 경로를 가지며, 이로 인해 본질적으로 높은 활성화 에너지를 요구한다.
비귀금속 촉매는 알칼리 조건에서 수소 발생 반응(HER)에서는 귀금속에 근접한 성능을 보이는 경우가 많지만, OER에서는 반응 중간체(M–OH, M–O, M–OOH)를 동시에 안정화시키기 어렵기 때문에 높은 과전압(overpotential)이 요구된다.
이 과전압은 전극에 인가된 전기에너지 중 상당 부분이 유효한 화학에너지로 전환되지 못하고 열로 손실됨을 의미하며, 결과적으로 수전해 시스템의 전체 효율(LHV 기준)을 저하시킨다(Seh et al., 2017).

3-3-4 전도성 및 내구성의 상관관계

비귀금속 촉매로 주로 사용되는 금속 산화물이나 수산화물은, 본질적으로 금속 상태에 비해 전기 전도도가 낮다. 낮은 전도성은 전극 내부에서 전자 전달 저항을 증가시켜 반응 속도를 제한하는 요인이 된다. 이를 보완하기 위해 전도성 탄소(C, carbon black, CNT 등)를 지지체로 사용하는 경우가 많지만, 이 접근법은 또 다른 문제를 발생시킨다. 산소 발생 반응이 일어나는 고전압 조건(일반적으로 1.5 V vs. RHE)에서는 탄소가 열역학적으로 불안정해져 이산화탄소(CO_2)로 산화되는 지지체 부식(Carbon corrosion)이 발생한다. 해당 문제들로 인해 촉매 활성점이 탈락하거나 전극 구조가 붕괴되어, 장기 운전 시 내구성이 급격히 저하되는 단점이 생긴다.

3-3-5 사회/경제적 한계

기술적 성능이 실험실 수준에서 우수하더라도, 이를 산업적으로 적용하기 위해서는 제조 공정, 장기 신뢰성, 그리고 규제·표준 체계가 함께 고려되어야 한다. 비귀금속 촉매는 조성 변화에 민감하고, 대면적 전극으로 확장할 경우 균일한 성능을 확보하기 어렵다는 문제가 있다. 또한 장기 운전 데이터가 귀금속 촉매에 비해 부족해, 금융·정책적 관점에서 투자 리스크가 높게 평가되는 경향이 있다.

4. 재료과학적 돌파구 : 한계 극복을 위한 최신 연구흐름

이러한 한계를 극복하기 위해 최근 연구는 재료의 조성을 바꾸는 수준을 넘어, 원자와 전자의 배치를 설계하는 방향으로 발전하고 있다. 나노기술과 계산과학의 결합은 비귀금속 촉매의 근본적 약점을 보완할 수 있는 새로운 가능성을 제시한다.

4-1. 삼상 경계(Triple Phase Boundary, TPB)와 촉매 구조

촉매 반응은 표면에서 일어난다. 이를 나노미터(nm) 수준에서 확대해 보면, 반응은 다음 세 가지 요소가 만나는 접점에서만 발생함을 알 수 있다 (Bard and Faulkner, 2001). 고체 촉매: 반응 장소 제공 전해질(이온): 이온(OH^-등) 공급 통로 기체/액체(물): 반응물 공급 및 생성물 배출 통로 이 세 가지가 만나는 지점을 삼상 경계(TPB)라고 한다. 촉매 입자가 아무리 많아도 이온이 닿지 않거나 전자가 공급되지 않는 '고립된 촉매'는 반응에 참여하지 못한다. 따라서 나노 구조 제어를 통해 TPB 면적을 극대화하는 것이 성능 향상의 지름길이다.

4-2. 전자 구조 제어를 통한 활성 극대화

도핑 및 합금화 전략은 니켈 격자구조에 철이나 망간을 미량 첨가하여 금속의 전자 밀도와 d-orbital 점유 상태를 조절하는 접근법이다(Trotochaud et al., 2014). 이는 반응 중간체(M-O, M-OH, M-OOH)와 촉매 표면 사이의 결합에너지를 최적화함으로써 너무 강하지도 약하지도 않는 이상적인 결합을 유도하여

산소 발생 반응의 반응 장벽을 낮추는 Sabatier 원리의 실현을 목표로 한다(Nørskov et al., 2005). 또한 단일 원자 촉매(Single-Atom Catalysts, SACs)는 금속을 개별 원자 단위로 지지체에 고정시켜, 모든 금속 원자를 활성점으로 활용한다. 이 방식은 금속 사용 효율을 극대화함과 동시에, 귀금속 수준의 반응성을 비귀금속으로 구현할 수 있는 가능성을 보여준다(Seh et al., 2017).

4-3. 나노 구조 공학을 통한 표면적 확보

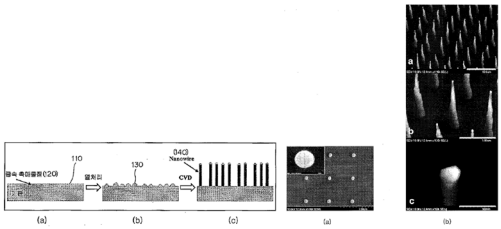


Figure 3. nano - wire의 성장 과정

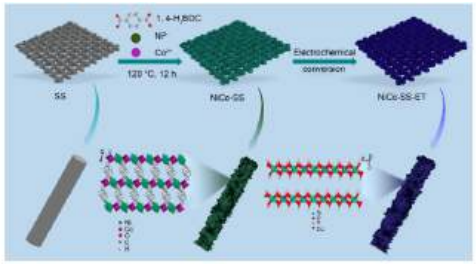


Figure 4. nano - sheet의 성장 과정

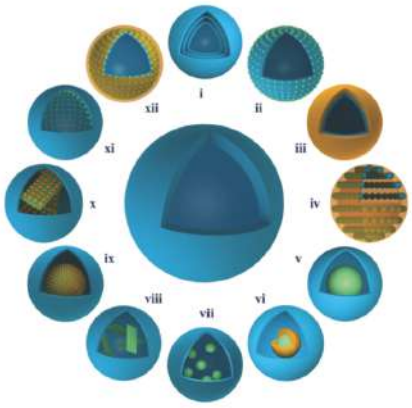


Figure 5. 기본 중공구조 및 여러가지 변형 중공구조

촉매를 나노 와이어(Nano - wire), 나노 시트(Nano - sheet) 또는 중공 구조(Hollow structure)로 제작하여 반응이 일어날 수 있는 유효 표면적을 극대화한다. 이는 같은 양의 비귀금속을 쓰더라도 반응에 참여할 수 있는 활성점의 개수를 증가시켜 전류 밀도를 수 배 이상 끌어올릴 수 있다.

4-3. 층상 이중 수산화물(LDH) 및 하이브리드 재료

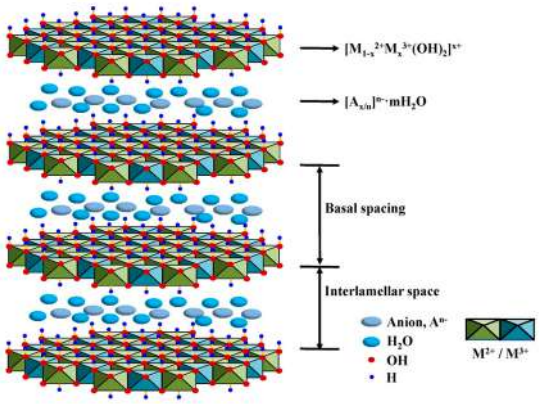


Figure 6. Metal LDH의 구조

최근 주목받는 Ni–Fe LDH는 층상 구조를 통해 이온 이동을 원활하게 하고, 활성점을 구조적으로 안정화한다. 일부 연구에서는 알칼리 수전해 조건에서 귀금속 기반 촉매를 능가하는 OER 성능을 보고하고 있다(Gong et al., 2013). 또한 그래핀이나 전도성 고분자와의 하이브리드화는 촉매 부식을 물리적으로 억제하는 보호막 역할을 수행한다.

5. Conclusion

비귀금속 촉매 기술은 지속 가능한 수소 생산 체계의 경제적 완성도를 결정짓는 마지막 퍼즐이다. 단순히 저렴한 재료를 찾는 것을 넘어, 원자 단위에서의 전자 구조 설계와 나노 공학적 접근을 통해 귀금속 수준의 활성과 내구성을 동시에 확보하려는 시도가 이어지고 있다. 하지만 실험실 환경에서의 성과가 대규모 산업 현장으로 이어지기 위해서는 가속 노화 테스트(AST)를 통한 장기 안정성 검증과 대면적 촉매 코팅 공정 기술이 병행되어야 한다. 향후 비귀금속 기반 AEM 수전해 기술이 상용화에 성공한다면, 수소는 더 이상 비싸서 한정적으로 이용되는 에너지가 아닌 손쉽게 얻을 수 있는 보편적 에너지가 될 것이다. 결국, 그린 수소의 미래는 나노 촉매의 표면 위에서 결정될 것이며, 이는 재료과학과 화학공학의 융합이 만들어낼 기후 위기 극복의 실전적 해답이 될 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

***Q1. 1kg*의 수소를 생산하는 데 이론적으로 필요한 전력은 *33.3kWh*이다. 하지만 실제 *PEM* 수전해 장치는 과전압 때문에 *50kWh*를 소모한다. 만약 새로운 비귀금속 *Ni-Fe* 촉매를 개발하여 전압 효율을 *5%* 개선한다면, 연간 *1,000*톤의 수소를 생산하는 공장에서 절감할 수 있는 전기료는 얼마인가? (전기료는 *150*원/*kWh*로 가정)**

탐구 가이드: 효율 개선이 단지 과학적 성취를 넘어 실제 산업 현장에서 어떤 경제적 가치(ROI)를 창출하는지 분석해 보자.

***Q2.* 비귀금속 촉매는 알칼리 환경에서 활성은 높지만, *1,000*시간 이상 구동 시 표면 구조가 변하며 성능이 저하되는 내구성 문제가 있다. 이를 테스트하기 위한 가속 열화 시험*(Accelerated Stress Test)* 방법을 설계해 보자.**

탐구 가이드: 전압을 빠르게 반복해서 변화시키거나 온도를 높였을 때, 촉매 입자가 용출(Leaching)되는지 확인하는 방법을 화학적 원리로 설명해 보자.

***Q3.* 아래의 세 지역에 가장 적합한 수전해 기술을 매칭하고 그 이유를 재료/환경적 특성과 연결해 설명하시오.**

- A 지역: 대규모 풍력 발전 단지가 있어 전력 변동이 매우 심함.
- B 지역: 인근에 대형 원자력 발전소가 있어 저렴한 고온 폐열이 풍부함.
- C 지역: 축산 농가가 많아 바이오가스 플랜트가 이미 구축되어 있음.

탐구 가이드: 각 수전해 기술(PEMWE, SOEC, AEC)의 응답성과 작동 온도 특성을 고려하여 최적의 시스템 시나리오를 작성해 보자.

References

Bard, A. J., and L. R. Faulkner. 2001. Electrochemical Methods: Fundamentals and Applications. 2nd edition. Wiley.

Carmo, M., D. L. Fritz, J. Mergel, and D. Stolten. 2013. A comprehensive review on PEM water electrolysis. International Journal of Hydrogen Energy **38**:4901–4934.

Dekel, D. R. 2018. Review of cell performance in anion exchange membrane fuel cells. Journal of Power Sources **375**:158–169.

Gong, M., Y. Li, H. Wang, Y. Liang, J. Z. Wu, J. Zhou, J. Wang, T. Regier, F. Wei, and H. Dai. 2013. An Advanced Ni–Fe Layered Double Hydroxide Electrocatalyst for Water Oxidation. Journal of the American Chemical Society **135**:8452–8455.

International Energy Agency, 2019. The Future of Hydrogen.

Merle, G., M. Wessling, and K. Nijmeijer. 2011. Anion exchange membranes for alkaline fuel cells. Journal of Membrane Science **377**:1–35.

Nørskov, J. K., T. Bligaard, A. Logadottir, J. R. Kitchin, J. Chen, S. Pandelov, and U. Stimming. 2005. Trends in the Exchange Current for Hydrogen Evolution. Journal of The Electrochemical Society **152**:J23–J26.

O'Hayre, R., S.-W. Cha, W. Colella, and F. B. Prinz. 2016. Fuel Cell Fundamentals. 3rd edition. Wiley.

Seh, Z. W., J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov, and T. F. Jaramillo. 2017. Combining theory and experiment in electrocatalysis: Insights into materials design. Science **355**:eaad4998. https://www.science.org/doi/abs/10.1126/science.aad4998.

Trotochaud, L., S. L. Young, J. K. Ranney, and S. W. Boettcher. 2014. Nickel–Iron Oxyhydroxide Oxygen-Evolution Electrocatalysts: The Role of Intentional and Incidental Iron Incorporation. Journal of the American Chemical Society **136**:6744–6753.

Zeng, K., and D. Zhang. 2010. Recent progress in alkaline water electrolysis for hydrogen production. Progress in Energy and Combustion Science **36**:307–326.

Zou, X., and Y. Zhang. 2015. Noble metal-free hydrogen evolution catalysts for water splitting. Chemical Society Reviews **44**:5148–5180.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Changhee Kim

M. chkim@kentech.ac.kr

Prof. Prof. Chang-Hee Kim received his Ph.D. degree in the Department of Materials Science andEngineering from the Korea Advanced Institute of Science and Technology (KAIST), Korea. Since joining the Hydrogen Research Department at the Korea Institute of Energy Research (KIER) in 2004, he has been actively involved in multidisciplinary research projects on green hydrogen production technologies and also served as the Head of the Hydrogen Research Department, leading multiple national flagship projects. He made a significant contribution to advancing the technological capabilities of Korean companies in the field of green hydrogen production and achieved technology transfer to Korean industry with a cumulative value of approximately 3.8 billion KRW. He was honored with the Presidential Citation Ribbon from the Korean government and the Minister’s Citation from the Ministry of Science and ICT in 2021 for his significant contributions in this field. He is currently a full professor in the Hydrogen Track of the School of Energy Technology at the Korea Institute of Energy Technology (KENTECH), and actively serves as the Chair of the POSCO Advisory Council, the Korean representative member of the IEA Water Electrolysis Task, a committee member of the R&D Advisory Committee of Korea Hydro & Nuclear Power (KHNP), and as Vice President of the Korean Hydrogen and New Energy Society as well as the Korean Society of Environmental Engineers.

KENTECH
Korea Institute of Energy Technology

Vol. 01 (2026)

Non-Precious Metal Catalyst–Based Water Electrolysis: Understanding Green Hydrogen Production from Membrane and Nanoscale Perspectives

Abstract

Water electrolysis is a technology that produces hydrogen by splitting water using electricity. It has attracted significant attention as a next-generation green hydrogen production method since it does not directly emit carbon dioxide. When coupled with renewable energy sources, electrolysis can serve as a key pathway to convert intermittent electrical power into storable chemical energy. However, discussions of water electrolysis are often simplified to focus primarily on catalyst performance, while the structural characteristics regarding real device operation receive comparatively less attention.

This study emphasizes that electrochemical water-splitting reactions do not occur on catalysts alone, but within an integrated Membrane Electrode Assembly (MEA) structure. In particular, during the transition from noble-metal-based catalysts to non-precious metal catalysts, the membrane plays a central role in determining both the reaction environment and ion transport pathways. Although non-precious metal catalysts offer substantial advantages in terms of cost and resource availability, their instability under acidic conditions necessitates operation in alkaline environments in combination with anion exchange membranes.

Based on this perspective, the work examines the fundamental structure of water electrolysis systems, the contrasting characteristics of noble and non-precious metal catalysts, and the concept of the triple-phase boundary. Through this analysis, non-precious metal-based water electrolysis is framed not merely as a catalyst selection problem, but as a challenge of reaction environment and structural design. This structural understanding provides a foundation for exploring the potential and limitations of non-precious metal electrolysis technologies, as well as recent materials science strategies aimed at overcoming these challenges.

Introduction

Water electrolysis is a technology that produces hydrogen by splitting water using electricity. It is recognized for its ability to utilize hydrogen as a long-term energy storage medium. In particular, when coupled with renewable energy sources, water electrolysis has attracted attention as a next-generation hydrogen production technology because it enables hydrogen generation without direct carbon dioxide emissions.

In such explanations, water electrolysis is often presented as a simple chemical reaction: applying a voltage to water produces hydrogen and oxygen, and the use of a catalyst merely accelerates the process. In reality, however, practical electrolysis systems exhibit far more complex structural characteristics than this simplified description suggests.

Moreover, the performance and stability of water electrolysis systems are not determined solely by operating voltage or intrinsic catalyst activity. Rather, the structural context in which a catalyst is embedded and the environment under which it operates are central to electrolysis performance. Water electrolysis should therefore be understood not simply as a problem of selecting efficient catalysts, but as a systems engineering challenge involving the design of reaction environments and device architectures. This perspective becomes particularly important in the transition from noble-metal catalysts to non-precious metal catalysts. While non-precious metal

catalysts offer significant advantages in terms of cost and resource availability, they cannot operate stably under conventional acidic conditions. Consequently, the viability of non-precious metal electrolysis is governed less by catalyst properties alone than by the overall system design, including membrane selection and ion transport architecture that enable stable operation.

Accordingly, this study focuses on the fact that water electrolysis reactions occur not on catalysts alone but within an integrated structure consisting of the electrode, catalyst, and membrane, known as the Membrane Electrode Assembly (MEA). Emphasis is placed on how the membrane governs reaction conditions through ion transport and pH regulation, and non-precious metal catalyst–based water electrolysis is analyzed from both structural and nanoscale perspectives.

1. Background: Green Hydrogen and Water Electrolysis

1-1. What Is Green Hydrogen?

Hydrogen has gained attention as a key energy carrier for achieving carbon neutrality in response to the climate crisis. However, its environmental impact varies significantly depending on the production method. Gray hydrogen, which is produced through the reforming of natural gas, inevitably results in the emission of carbon dioxide(CO_2). n contrast, green hydrogen, which is produced by electrolyzing water (H_2O) using electricity generated from renewable energy sources such as solar or wind power, is considered the cleanest form of hydrogen, as its production process releases only oxygen (O_2) (International Energy Agency, 2019).

1-2. Water Electrolysis: The Technology for Splitting Water

Water electrolysis is a technology that separates water into hydrogen and oxygen through electrochemical reactions. Because water is thermodynamically a highly stable substance, external energy must be supplied in order to decompose it.

$$2\,H_2O(l) \rightarrow 2\,H_2(g) + O_2(g)$$

In water electrolysis, the required energy is supplied in the form of electricity. A key challenge is that, although the thermodynamically required voltage is 1.23 V, practical systems require a higher operating voltage due to overpotentials. The ability to minimize these overpotentials therefore serves as a critical measure of technological performance (Zeng and Zhang, 2010).

2. The Heart of a Water Electrolysis System: The MEA

2-1. The Integrated Structure of the Membrane Electrode Assembly (MEA)

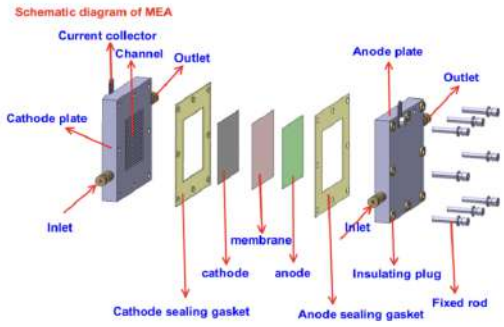


Figure 1. The Structure of MEA

The MEA, which is the core component of a water electrolysis cell, consists of a sandwich-like structure in which the electrodes, catalysts, and membrane are compressed together.
• Electrode: The electrode functions as a current collector that supplies electrons from, or transfers electrons to, the external circuit. It typically possesses a porous structure to allow the efficient transport and release of generated gases.
• Catalyst: The catalyst is the site at which the electrochemical water-splitting reactions occur. By lowering the activation energy of the reactions, the catalyst significantly enhances reaction rates. Without a catalyst, the water-splitting process proceeds too slowly to be practical for real-world applications.
• Membrane: The membrane physically separates the anode and cathode, preventing the mixing of generated hydrogen and oxygen gases and thereby reducing the risk of explosive reactions. At the same time, it selectively allows the transport of specific ions, enabling the completion of the electrochemical circuit (O'Hayre et al., 2016).

2-2. The Membrane: A Determinant of Ion Transport and pH Environment

The membrane is not merely a physical barrier. For electrochemical reactions to proceed continuously, electrons must travel through the external circuit while ions migrate through the membrane, together forming a closed electrochemical loop. In addition, the pH of the electrolyte is determined by the ion-exchange characteristics of the membrane, specifically whether it conducts protons (H^+) or hydroxide ions (OH^-).
• Proton Exchange Membrane (PEM): Conducts protons (H^+), thereby creating a strongly acidic operating environment.
• Anion Exchange Membrane (AEM): Conducts hydroxide ions (OH^-), thereby creating an alkaline (basic) operating environment. This pH environment becomes the most critical factor in determining which metal catalysts can be stably and effectively employed in the electrolysis system (Merle et al., 2011).

3. AEM and Non-Precious Metal Catalysts

The transition from noble-metal catalysts to non-precious metal catalysts (NPMCs) represents not merely a cost-reduction strategy, but a structural shift required to scale green hydrogen production from laboratory-scale systems to gigawatt (GW)-level industrial deployment.
Current commercial water electrolysis technologies face fundamental limitations in large-scale deployment due to the scarcity and supply constraints of catalyst materials. Although non-precious metal catalysts based on transition metals such as nickel (Ni), iron (Fe), and cobalt (Co) are actively being investigated, these metals cannot withstand the acidic conditions imposed by proton exchange membrane (PEM) systems.
To address this limitation, anion exchange membrane (AEM) water electrolysis, which operates under alkaline conditions, has emerged as a promising alternative and is the primary focus of this study. Under alkaline environments, non-precious metal catalysts can operate stably without severe corrosion, enabling their practical application in electrolysis systems (Zou and Zhang, 2015). In essence, the core strategy of non-precious metal–based water electrolysis lies in altering the chemical environment (pH) and the dominant ion transport mechanism—from proton conduction to hydroxide ion conduction—across the entire system in order to enable the use of low-cost catalysts.
Non-precious metal catalysts have emerged as a practical solution to these bottlenecks. Owing to their abundant availability, low cost, and tunable performance through alloying and nanostructural engineering, they are increasingly regarded as key components of next-generation water electrolysis systems.
However, non-precious metal catalysts inherently lag behind noble-metal catalysts in terms of catalytic activity, stability, and electrical conductivity. If these limitations are not effectively addressed, their industrial application will remain constrained. This section therefore systematically examines not only the potential of non-precious

metal catalysts but also the underlying technological dilemmas that accompany their use.

3-1. Potential of AEM and Non-Precious Metal Catalyst Technologies: A Comparison with PEM and Noble-Metal Catalyst Systems

The incorporation of non-precious metal catalysts into AEM-based electrolysis systems offers substantial potential across the three core pillars of the hydrogen economy: economic viability, scalability, and resource security.

From an economic perspective, approximately 35–45% of the total cost of a water electrolysis stack arises from catalyst and electrode materials. In PEM electrolysis systems, the use of noble-metal catalysts such as iridium (Ir) is unavoidable; however, the global annual production of iridium is only about 7–8 tons, and its market price reaches tens of thousands of U.S. dollars per kilogram. In contrast, replacing noble metals with relatively abundant non-precious metals such as nickel (Ni), iron (Fe), and cobalt (Co) in AEM electrolysis systems could theoretically reduce catalyst material costs by a factor of 100–1000, lowering prices to the order of tens of dollars per kilogram. According to analyses by the International Energy Agency (IEA) and the U.S. Department of Energy (DOE), such reductions in catalyst cost constitute a key driver for lowering the levelized cost of hydrogen (LCOH) to below 1.5–2.0 USD per kilogram in the long term.

From a scalability perspective, non-precious metal catalysts provide the material accessibility required for large-scale system deployment. In noble-metal-based PEM electrolysis, catalyst supply itself becomes a critical bottleneck once system capacities exceed several hundred megawatts. In contrast, AEM electrolysis systems employing non-precious metal catalysts rely on metals with globally abundant reserves, enabling expansion to gigawatt-scale facilities without fundamental resource constraints.

From the perspective of supply chain stability and resource security, non-precious metal catalysts also carry significant strategic value. Platinum-group metals, which are essential for noble-metal-based systems, are geographically concentrated in a limited number of countries such as South Africa and Russia, resulting in substantial geopolitical risks. In contrast, the raw metals used for non-precious metal catalysts—such as nickel (Ni), iron (Fe), and manganese (Mn)—are more evenly distributed worldwide, enabling carbon-neutral nations to establish independent and resilient energy supply chains based on domestic technologies.

System design flexibility represents another important advantage. Non-precious metal catalysts operate stably primarily under alkaline conditions, such as in alkaline electrolysis cells (AEC) and anion exchange membrane water electrolysis (AEMWE). This allows the use of cost-effective materials—including stainless steel, nickel alloys, and polymer-based components—in place of expensive titanium bipolar plates that are required to withstand corrosion under acidic conditions. As a result, reductions in catalyst cost propagate through the system, leading to cascading decreases in overall hardware costs at the stack and system levels.

3-2. Technological Dilemmas and Limitations

Despite their significant potential in terms of cost efficiency and scalability, non-precious metal catalysts have not yet been able to fully replace noble-metal catalysts in water electrolysis systems.

The fundamental reason lies in deficiencies across three critical aspects of catalyst design—namely catalytic activity and chemical and electrochemical stability—which remain significantly inferior to those of noble-metal catalysts. In general, enhancing catalytic activity requires inducing highly reactive and often metastable surface states; however, such states tend to promote catalyst dissolution or structural degradation during long-term operation. This trade-off between activity and durability constitutes a central technological dilemma that hinders the commercialization of non-precious metal catalysts (Seh et al., 2017).

3-3-1. Thermodynamic Instability under Acidic Conditions

The most critical limitation of non-precious metal catalysts lies in their chemical stability. State-of-the-art PEM water electrolysis systems operate under strongly acidic conditions (pH ¡ 1), in which most non-precious metals—such as Ni, Fe, and Mn—are thermodynamically unstable and readily dissolve into ionic species through leaching. This process drastically shortens catalyst lifetime and introduces a further dilemma, as the dissolved metal ions can contaminate the ion-exchange membrane, leading to degradation of overall system performance.

Importantly, this phenomenon extends beyond a simple loss of catalyst mass. Leached metal ions can migrate into the polymer electrolyte membrane, where they interfere with ion-exchange functionality or cause membrane fouling. These effects result in increased cell resistance, higher operating voltages, and reduced energy efficiency. Consequently, the instability of non-precious metal catalysts represents not merely a materials-level challenge, but a system-wide threat that compromises both the durability and reliability of water electrolysis devices.

3-3-2 High Overpotential of the Oxygen Evolution Reaction (OER)

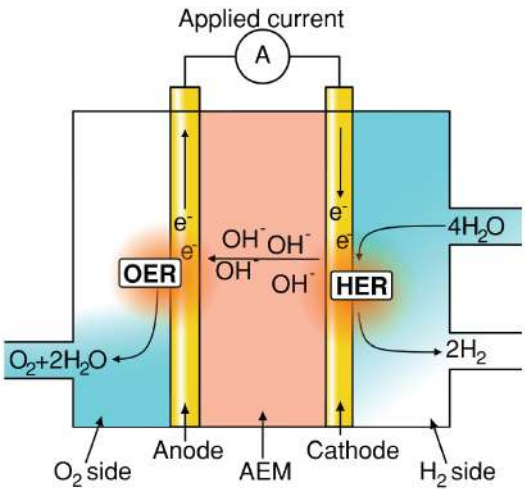


Figure 2. OER and HER in AEM

Water electrolysis is a technology that splits water(H_2O) into hydrogen(H_2) and oxygen(O_2) using electricity. This process proceeds through two half-reactions: the hydrogen evolution reaction (HER), in which hydrogen is produced from water at one electrode, and the oxygen evolution reaction (OER), in which oxygen is generated from water at the other electrode.

The hydrogen evolution reaction (HER) involves relatively simple electron-transfer steps and therefore proceeds more readily. In contrast, the oxygen evolution reaction (OER) is a much more complex process, requiring the sequential transfer of four electrons and protons to form a single oxygen molecule. As a result, OER demands significantly higher energy input and serves as the key rate-limiting step that ultimately governs the overall efficiency of water electrolysis.

Consequently, the oxygen evolution reaction (OER) is the critical process that demands particular attention, as it ultimately determines the overall efficiency of water electrolysis. While non-precious metal catalysts can exhibit performance comparable to that of noble-metal catalysts for the hydrogen evolution reaction (HER), they suffer from substantially higher energy barriers—manifested as larger overpotentials—during OER due

to its complex four-electron transfer pathway. This results in a greater fraction of the supplied electrical energy being dissipated as heat, thereby reducing the overall energy efficiency of the electrolysis system, even when evaluated on a lower heating value (LHV) basis.

The reaction that ultimately determines the overall energy efficiency of a water electrolysis system is the oxygen evolution reaction (OER). During OER, oxygen is generated from water molecules through a complex reaction pathway involving the sequential transfer of four electrons and protons, which intrinsically requires a high activation energy.

Under alkaline conditions, non-precious metal catalysts often exhibit hydrogen evolution reaction (HER) performance comparable to that of noble-metal catalysts. However, during the oxygen evolution reaction (OER), these catalysts require significantly higher overpotentials because they struggle to simultaneously stabilize the key reaction intermediates (M–OH, M–O, and M–OOH).

This overpotential indicates that a substantial portion of the electrical energy applied to the electrode is not converted into useful chemical energy but is instead dissipated as heat, ultimately leading to a reduction in the overall efficiency of the water electrolysis system when evaluated on a lower heating value (LHV) basis (Seh et al., 2017).

3-3-4 Correlation Between Electrical Conductivity and Durability

Metal oxides and hydroxides, which are commonly employed as non-precious metal catalysts, inherently exhibit much lower electrical conductivity than their metallic counterparts. Low conductivity increases electron transport resistance within the electrode, thereby limiting reaction kinetics. To compensate for this drawback, conductive carbon materials—such as carbon black or carbon nanotubes (CNTs)—are often used as catalyst supports.

However, this strategy introduces an additional challenge. Under the high anodic potentials required for the oxygen evolution reaction (typically 1.5 V vs. RHE), carbon materials become thermodynamically unstable and undergo oxidation to carbon dioxide (CO_2), a phenomenon known as carbon corrosion. As a result, active catalytic sites can be lost and the electrode structure may collapse, leading to a severe degradation in durability during long-term operation.

3-3-5 Socioeconomic Limitations

Even when technological performance is promising at the laboratory scale, industrial implementation requires careful consideration of manufacturing processes, long-term reliability, and regulatory and standardization frameworks. Non-precious metal catalysts are often sensitive to compositional variations, making it challenging to achieve uniform performance when scaled to large-area electrodes. In addition, the lack of extensive long-term operational data compared to noble-metal catalysts leads to higher perceived investment risk from financial and policy perspectives.

4. Materials Science Breakthroughs: Emerging Research Trends to Overcome Current Limitations

To overcome these limitations, recent research has progressed beyond simply modifying material composition toward the deliberate design of atomic and electronic structures. The integration of nanotechnology with computational science presents new opportunities to address the fundamental weaknesses of non-precious metal catalysts and to unlock previously inaccessible performance regimes.

4-1.Triple Phase Boundary (TPB) and Catalyst Structure

Catalytic reactions occur at surfaces. When examined at the nanometer (nm) scale, it becomes evident that these reactions take place only at the interfaces where three key elements converge (Bard and Faulkner, 2001).
Solid catalyst: Provides the active sites where reactions occur
Electrolyte (ions): Serves as the transport pathway for ions (e.g., OH^-)
Gas/liquid phase (water): Supplies reactants and enables the removal of reaction products
The interface at which these three components intersect is referred to as the triple phase boundary (TPB). Regardless of the total number of catalyst particles present, “isolated catalysts” that lack access to ionic transport or electron supply cannot participate in electrochemical reactions. Consequently, maximizing the TPB area through nanoscale structural control represents a key strategy for enhancing catalytic performance.

4-2. Maximizing Catalytic Activity through Electronic Structure Engineering

Doping and alloying strategies involve the introduction of trace amounts of elements such as iron or manganese into the nickel lattice to modulate the electronic density and the occupancy of metal d orbitals (Trotochaud et al., 2014). This approach aims to realize the Sabatier principle by optimizing the binding energies between reaction intermediates (M–O, M–OH, and M–OOH) and the catalyst surface. By achieving an ideal interaction that is neither too strong nor too weak, the reaction energy barrier for the oxygen evolution reaction can be effectively reduced(Nørskov et al., 2005). In addition, single-atom catalysts (SACs) anchor individual metal atoms onto a support, enabling every metal atom to function as an active site. This approach maximizes metal utilization efficiency while demonstrating the potential to achieve noble-metal-like reactivity using non-precious metals (Seh et al., 2017).

4-3. Enhancing Surface Area through Nanostructural Engineering

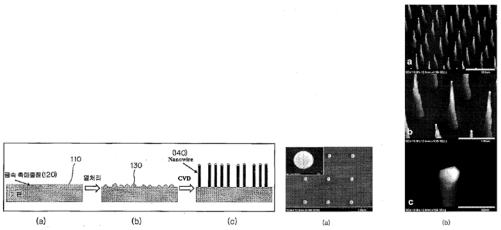


Figure 3. Growth Process of Nanowires

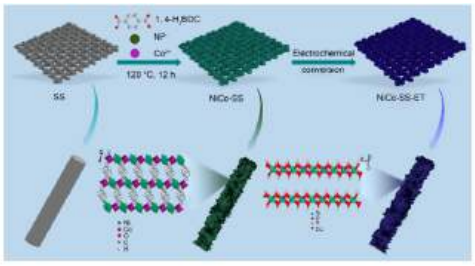


Figure 4. Growth Process of Nanosheets

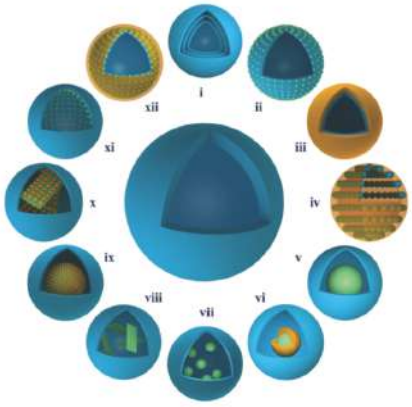


Figure 5. Basic Hollow Structures and Their Various Modified Forms

Fabricating catalysts in the form of nanowires, nanosheets, or hollow structures maximizes the effective surface area available for electrochemical reactions. Even when the same amount of non-precious metal is used, such nanostructuring significantly increases the number of accessible active sites, thereby enabling current densities to be enhanced by several-fold.

4-3. Layered Double Hydroxides (LDHs) and Hybrid Materials

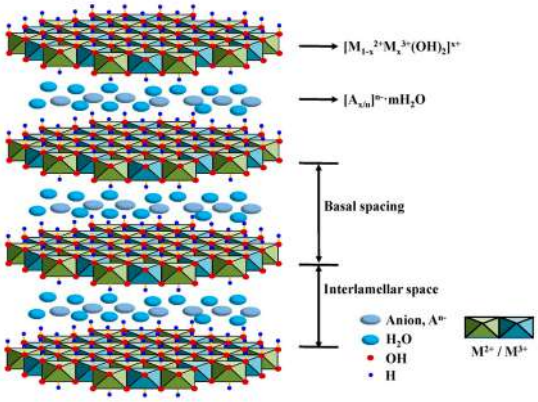


Figure 6. Structure of Metal Layered Double Hydroxides (Metal LDHs)

Recently, Ni–Fe layered double hydroxides (LDHs) have attracted considerable attention due to their layered structures, which facilitate ion transport while structurally stabilizing active sites. Several studies have reported that, under alkaline water electrolysis conditions, Ni–Fe LDHs can exhibit oxygen evolution reaction (OER) performance that rivals or even surpasses that of noble-metal-based catalysts(Gong et al., 2013). In addition, hybridization with graphene or conductive polymers can act as a protective layer that physically suppresses catalyst corrosion, thereby enhancing structural stability during electrochemical operation.

5. Conclusion

Non-precious metal catalyst technologies represent the final puzzle piece in achieving the economic viability of sustainable hydrogen production systems. Beyond the pursuit of low-cost materials, current efforts focus on simultaneously attaining noble-metal-level activity and durability through atomic-scale electronic structure engineering and nanoscale materials design.

However, for laboratory-scale achievements to translate into large-scale industrial deployment, these advances must be accompanied by rigorous long-term stability validation through accelerated stress tests (ASTs) and the development of scalable, large-area catalyst coating processes. If non-precious-metal-based AEM water electrolysis technologies are successfully commercialized, hydrogen will no longer remain an expensive and limited energy carrier, but will instead become a widely accessible and practical energy resource.

Ultimately, the future of green hydrogen will be determined on the surfaces of nanoscale catalysts, where the convergence of materials science and chem

Example Questions

Research on a freely chosen topic is permitted, including the specified question.

Q1. The theoretical minimum electricity required to produce 1 kg of hydrogen is 33.3 kWh. However, due to overpotentials, practical PEM water electrolysis systems consume approximately 50 kWh per kilogram of hydrogen. If a newly developed non-precious Ni–Fe catalyst improves voltage efficiency by 5%, how much electricity cost can be saved annually at a plant producing 1,000 tons of hydrogen per year? (Assume an electricity price of 150 KRW per kWh.) Research Guide: Analyze how efficiency improvements extend beyond scientific achievement to generate tangible economic value (return on investment, ROI) in real industrial applications.

Q2. Non-precious metal catalysts can exhibit high activity under alkaline conditions, but they often face durability issues: after more than 1,000 hours of operation, surface restructuring can occur, leading to performance degradation. Design an Accelerated Stress Test (AST) protocol to evaluate this degradation behavior. Research Guide: Using chemical principles, explain how you would test whether catalyst particles undergo leaching when (i) the cell voltage is rapidly cycled or (ii) the operating temperature is increased.

Q3. Match the most suitable water electrolysis technology to each of the three regions below, and explain your choices by linking them to the relevant materials-related and environmental characteristics of each region.

- Region A: A large-scale wind power complex is present, resulting in highly fluctuating electricity supply.
- Region B: A large nuclear power plant is located nearby, providing an abundant supply of low-cost high-temperature waste heat.
- Region C: The area has a high concentration of livestock farms, and biogas plants are already in operation.

Research Guide: Develop an optimal system scenario for each region by considering the dynamic response characteristics and operating temperature requirements of each electrolysis technology (PEMWE, SOEC, and AEC).

References

Bard, A. J., and L. R. Faulkner. 2001. Electrochemical Methods: Fundamentals and Applications. 2nd edition. Wiley.

Carmo, M., D. L. Fritz, J. Mergel, and D. Stolten. 2013. A comprehensive review on PEM water electrolysis. International Journal of Hydrogen Energy **38**:4901–4934.

Dekel, D. R. 2018. Review of cell performance in anion exchange membrane fuel cells. Journal of Power Sources **375**:158–169.

Gong, M., Y. Li, H. Wang, Y. Liang, J. Z. Wu, J. Zhou, J. Wang, T. Regier, F. Wei, and H. Dai. 2013. An Advanced Ni–Fe Layered Double Hydroxide Electrocatalyst for Water Oxidation. Journal of the American Chemical Society **135**:8452–8455.

International Energy Agency, 2019. The Future of Hydrogen.

Merle, G., M. Wessling, and K. Nijmeijer. 2011. Anion exchange membranes for alkaline fuel cells. Journal of Membrane Science **377**:1–35.

Nørskov, J. K., T. Bligaard, A. Logadottir, J. R. Kitchin, J. Chen, S. Pandelov, and U. Stimming. 2005. Trends in the Exchange Current for Hydrogen Evolution. Journal of The Electrochemical Society **152**:J23–J26.

O'Hayre, R., S.-W. Cha, W. Colella, and F. B. Prinz. 2016. Fuel Cell Fundamentals. 3rd edition. Wiley.

Seh, Z. W., J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov, and T. F. Jaramillo. 2017. Combining theory and experiment in electrocatalysis: Insights into materials design. Science **355**:eaad4998. https://www.science.org/doi/abs/10.1126/science.aad4998.

Trotochaud, L., S. L. Young, J. K. Ranney, and S. W. Boettcher. 2014. Nickel–Iron Oxyhydroxide Oxygen-Evolution Electrocatalysts: The Role of Intentional and Incidental Iron Incorporation. Journal of the American Chemical Society **136**:6744–6753.

Zeng, K., and D. Zhang. 2010. Recent progress in alkaline water electrolysis for hydrogen production. Progress in Energy and Combustion Science **36**:307–326.

Zou, X., and Y. Zhang. 2015. Noble metal-free hydrogen evolution catalysts for water splitting. Chemical Society Reviews **44**:5148–5180.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Changhee Kim

M. chkim@kentech.ac.kr

Prof. Prof. Chang-Hee Kim received his Ph.D. degree in the Department of Materials Science andEngineering from the Korea Advanced Institute of Science and Technology (KAIST), Korea. Since joining the Hydrogen Research Department at the Korea Institute of Energy Research (KIER) in 2004, he has been actively involved in multidisciplinary research projects on green hydrogen production technologies and also served as the Head of the Hydrogen Research Department, leading multiple national flagship projects. He made a significant contribution to advancing the technological capabilities of Korean companies in the field of green hydrogen production and achieved technology transfer to Korean industry with a cumulative value of approximately 3.8 billion KRW. He was honored with the Presidential Citation Ribbon from the Korean government and the Minister's Citation from the Ministry of Science and ICT in 2021 for his significant contributions in this field. He is currently a full professor in the Hydrogen Track of the School of Energy Technology at the Korea Institute of Energy Technology (KENTECH), and actively serves as the Chair of the POSCO Advisory Council, the Korean representative member of the IEA Water Electrolysis Task, a committee member of the R&D Advisory Committee of Korea Hydro & Nuclear Power (KHNP), and as Vice President of the Korean Hydrogen and New Energy Society as well as the Korean Society of Environmental Engineers.

KENTECH
Korea Institute of Energy Technology

V

전력반도체 :

친환경 에너지 시대로 향하는 길

p. 098 - 123

Vol. 2 (2026)

전력반도체: 친환경 에너지 시대로 향하는 길

Abstract 본 탐구 자료는 신재생에너지 전력변환 시스템의 핵심 기술로 주목받는 전력반도체의 물리적 원리와 구조적 특성, 그리고 실제 에너지 시스템에서의 적용 사례를 종합적으로 분석한다. 반도체의 에너지 밴드 이론과 도핑을 기반으로 PN 접합의 형성과 전류 제어 원리를 살펴보고, 고전압·고전류 환경에 적합한 PIN 접합 구조의 특성과 전압 차단 메커니즘을 분석한다. 이어서 전력반도체 내부 구조와 동작 원리를 비교하고, 각 소자의 스위칭 특성과 기생 요소가 전력 손실 및 시스템 효율에 미치는 영향을 고찰한다. 나아가 태양광 및 풍력 발전 시스템에서 전력반도체가 수행하는 역할을 전력 흐름과 동적 스위칭 관점에서 분석함으로써, 전력반도체의 성능이 신재생에너지 시스템의 효율과 신뢰성을 결정하는 핵심 요소임을 확인한다. 본 탐구 자료를 통해 전력반도체는 단순한 전력 제어 소자를 넘어, 친환경 에너지 사회 구현을 가능하게 하는 기반 기술임을 제시하고자 한다.

Introduction

현대 사회는 탄소 중립이라는 거대한 변화의 흐름 속에서 화석 연료 중심의 에너지 사회에서 친환경 에너지 중심의 사회로 급격히 전환되고 있다. 생산되는 전력을 효율적으로 제어하고 변환하는 기술은 국가 산업 경쟁력을 좌우하는 척도이다. 이러한 에너지 전환 시대의 중심에는 전력을 각 기기에 맞게 분배하고 변환하는 전력 반도체(power semiconductor)가 있다. 전력 반도체는 고전압·고전류의 극한 환경을 견디며 전력을 변환, 변압, 분배하는 핵심 소자이다. 데이터 센터의 전력 효율화부터 전기차(EV)의 주행 거리 향상, 그리고 스마트 그리드 구축에 이르기까지 전력 반도체의 수요는 폭발적으로 증가하고 있다. 특히 전력 손실을 최소화하여 에너지 효율을 극대화하는 것은 미래 산업의 핵심 과제이며, 이를 실현하기 위해서는 소자의 동작 원리에 대한 근본적인 이해가 필수적이다. 이를 위해 본 탐구 자료에서는 전력 반도체의 기반이 되는 반도체 이론과 구조적 특징을 살펴보고자 한다. 이를 바탕으로 현대 전력 제어의 작동 원리를 알아보고 미래 에너지 산업에서의 적용 사례를 소개하고자 한다. 본 탐구자료는 단순한 소자의 기능을 넘어, 미래 친환경 에너지 사회를 지탱하는 기반 기술로서 전력 반도체의 가치를 알아볼 수 있는 기초를 제공한다.

1. 배경지식: 반도체의 물성 및 PN 접합

1-1. 반도체의 정의

우리 주변의 물질들은 구리와 알루미늄처럼 전기가 잘 흐르는 도체, 고무, 나무처럼 전기가 잘 흐르지 않는 부도체, 도체와 부도체의 중간 성질을 갖는 반도체로 나눌 수 있다. 반도체는 빛을 쪼이거나 전압을 걸어주는 것처럼 특정 조건을 만족하면 전기가 흐르는 성질을 가진 물질이다. 반도체는 이러한 특성을 이용하여 전류의 흐름을 조절하는데 사용하는 핵심 소자이다.

1-2. 원자 구조로 이해하는 반도체 : 에너지 밴드 이론

물질의 전기적 성질은 원자 내부의 전자들의 에너지 상태에 의해 결정된다. 이때 전자들이 존재하는 특정한 에너지 범위를 밴드(band)라고 한다. 밴드(band)의 종류에는 원자가 띠(valence band)와 전도 띠(conduction band) 두 가지가 있고, 두 밴드(band) 사이의 간격을 띠 틈(band gap)이라 한다. 원자가 띠

(valence band)는 원자핵의 영향을 강하게 받는 전자들이 존재하는 영역으로, 이 영역의 전자들은 원자에 속박되어 있다. 전도 띠(conduction band)는 자유롭게 움직이며 전류를 만들 수 있는 자유전자들이 존재하는 에너지 영역이다. 원자가 띠(valence band)의 전자가 띠 틈(band gap) 이상의 에너지를 받으면 전도 띠(conduction band)로 이동하게 된다. 띠 틈(band gap)의 크기는 보통 0.1eV ~ 4eV 정도이고, 반도체 제작에 가장 많이 사용되는 소재인 실리콘의 경우는 1.12eV이다.

1-3. 실리콘의 결정 구조와 도핑

실리콘은 반도체 제작에 가장 많이 사용되는 물질이다. 최외각 전자가 4개로 주변의 실리콘 원자 4개와 공유 결합을 통한 결정을 형성한다. 순수한 실리콘 결정 즉, 순수 반도체(intrinsic semiconductor)는 모든 전자가 결합 형성에 참여하여 자유 전자가 없기 때문에 상온에서 전류가 잘 흐르지 않는다. 그렇기 때문에 순수 반도체에 불순물을 첨가한 불순물 반도체(extrinsic semiconductor)를 만듦으로써 전류가 흐를 수 있게 한다. 불순물의 종류는 인(P)이나 비소(As) 같은 15족 원소와 붕소(B)와 갈륨(Ga) 같은 13족 원소가 있다. 15족 원소를 첨가하여 만들어진 N형 반도체는 결합에 참여하지 않는 전자가 자유롭게 이동하며 전류가 흐를 수 있다. 이와 달리 순수한 반도체에 13족 원소를 첨가하여 만들어지는 P형 반도체는 양공(hole)이라 불리는 빈 공간을 통해 전류가 흐를 수 있게 된다.

1-4. PN접합의 형성 및 전류의 흐름

P형 반도체와 N형 반도체 각각은 전류의 흐름을 조절 할 수 없기 때문에 두 종류의 반도체를 접합시켜 사용한다. 각 불순물 반도체의 다수 캐리어(majority carrier)가 다르기 때문에 접합면 부근에서 확산 및 재결합(recombination)이 일어난다. 이러한 과정 속에서 만들어진 공핍층(depletion region) 내부의 고정된 전하들에 의하여 형성된 전기장이 양공과 전자가 확산되는 것을 막아서 전류가 흐르지 못한다. PN 접합에 순방향 바이어스(P형 반도체에 +극, N형 반도체에 -극)를 가하면 정전위 장벽인 확산 포텐셜이 해소되며 전류가 흐르게 된다. 반대로 역방향 바이어스(P형 반도체에 -극, N형 반도체에 +극)를 가하면 양공(hole)은 음극 쪽으로, 자유전자(free electron)는 양극 쪽으로 몰리며 공핍층이 두꺼워지므로 전류가 흐르지 않게 된다. 이렇게 PN접합 다이오드가 가해진 바이어스에 따라 공핍층의 두께가 변하는 특성 때문에 정류 작용을 할 수 있다.

2. 전력 반도체의 기본 구조 : PIN 접합

2-1. PIN 다이오드란 무엇인가?

전력 반도체(power MOSFET)는 고전압과 고전류 환경에서 작동하는 반도체이므로 기존 PN접합 구조가 아닌 다른 구조가 필요하다. PIN 접합은 기존 PN 접합 구조에서 P형 반도체와 N형 반도체 사이에 도핑 농도가 약하여 순수 반도체에 가까운 반도체(intrinsic semiconductor)를 놓은 구조이다. 이때 보통 intrinsic 영역은 약하게 도핑된 N형 반도체가 위치한다.

2-2. PIN 접합 구조가 높은 전압을 견딜 수 있는 원리

PN 접합의 전기장을 그래프로 그려보면 경계면을 기준으로 대칭인 삼각형 모양으로 만들어진다. 버틸 수 있는 최대 전압의 크기는 V=Ed 식에 의해서 그래프 내부의 면적이다. 한편, 전기장 그래프의 기울기는 반도체의 도핑 농도에 비례 하므로 PIN 접합에서의 전기장은 사다리꼴로 나타나게 된다. 기존 PN 접합보다 intrinsic 영역의 넓이 만큼 버틸 수 있는 전압의 크기가 증가하므로 PIN 접합이 보다 높은 전압까지 버틸 수 있게 되는 것이다.

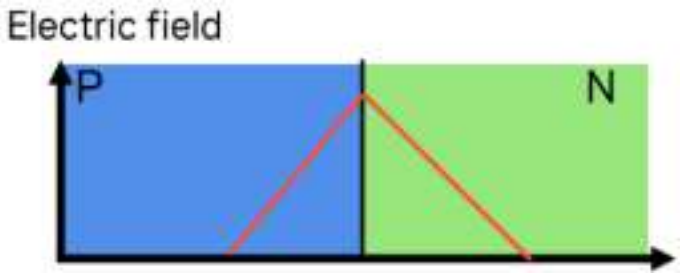


Figure 1. PN 접합에서의 전기장

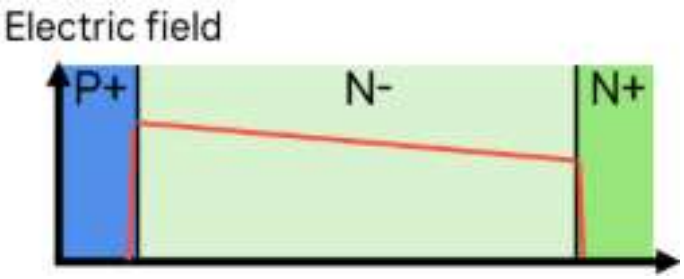


Figure 2. PIN 접합에서의 전기장

2-3. PIN 접합에서 관찰되는 현상 : Reverse Blocking

PIN 접합에서 역방향 바이어스가 걸리면 공핍층이 두꺼워지며 누설 전류(leakage current)가 흐르게 된다. 역방향 바이어스가 걸리게 되면 intrinsic 영역 전체가 공핍 영역(depletion region)이 된다. 이때, 통계적으로 매우 작은 양의 electron-hole pair가 generation되게 된다. 생성된 전자와 양공은 강한 전기장에 의해 양 끝으로 이동하면서 leakage current를 만든다. 이 현상을 reverse blocking이라 한다. drift region이 길어지고, 온도가 올라가고, 띠 틈(band gap)이 작아지고, lifetime이 감소한다면 leakage current가 강해진다.

2-4. PIN 접합에서 관찰되는 현상 : Reverse Recovery

PIN 접합에 순방향 바이어스를 걸어 놓은 상태에서 갑자기 역방향 바이어스를 걸었을 때 발생하는 현상으로 reverse blocking 현상처럼 drift region에서 통계적으로 generation되는 electron-hole pair가 단시간에 사라지지 않기 때문에 발생한다. 역방향 바이어스로 바꾸었을 때 스위치가 오프되며 순방향 전류가 급격하게 감소한다. 이때, drift region에 recombination 되지 않은 전자와 양공에 의해 역방향 전류가 흐르다가 carrier가 제거(recombination) 되며 depletion 상태로 회복된다. 역방향 전류가 흐르기 시작한 시간부터 depletion 회복이 완료될 때까지 걸리는 시간을 reverse recovery time이라 한다.

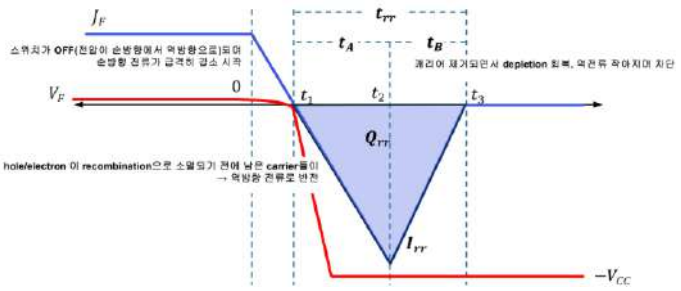


Figure 3. reverse recovery가 발생할 때 시간에 따른 전압과 전류의 관계

3. 금속과 반도체의 만남 : Schottky Junction과 Ohmic Contact

3-1. Schottky junction

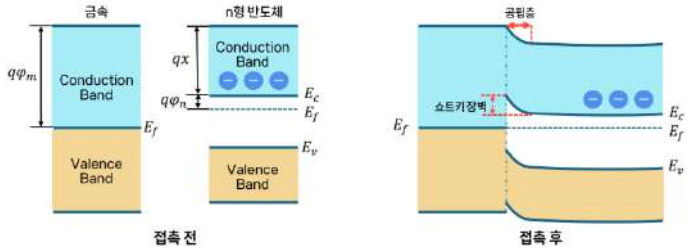


Figure 4. n형 반도체와 금속이 접촉해서 생기는 Schottky junction

금속과 반도체가 접촉할 때 일함수 차이에 의해 에너지 밴드가 휘어지며 쇼트키 장벽(Schottky barrier)이 형성되는 접합이다. carrier들이 이동하면서 반도체에 depletion이 형성된다. 평형 상태는 페르미 준위가 align된 상태로 쇼트키 장벽에 의해 전류가 흐르지 않는다. 순방향 전압을 걸어주게 된다면 쇼트키 장벽이 해소되며 전류가 흐르게 된다. 반대로 역방향 바이어스를 걸어주면 쇼트키 장벽이 높아지며 전류가 흐르지 못한다. 한편, 역방향 바이어스의 세기를 높여준다면 반도체와 금속 사이의 전기장이 강해지고, 추가적인 multiplication 효과들에 의해 breakdown이 발생하게 된다.

3-2. Ohmic contact

Schottky junction과 달리 반도체와 금속 접촉 경계면에서 에너지 장벽이 거의 없는 접촉이다. 예를 들어 금속과 N형 반도체가 Ohmic contact을 했다면 금속 표면에 전자가 축적(accumulation)되며 전류가 자유롭게 흐를 수 있게 된다. 펠티어 효과(Peltier effect)는 Ohmic contact에 의해 일어나는 대표적인 현상이다. 전자가 금속에서 반도체로 이동할 때는 에너지를 흡수하며 냉각되고, 반도체에서 금속으로 이동할 때는 에너지를 방출하며 발열한다.

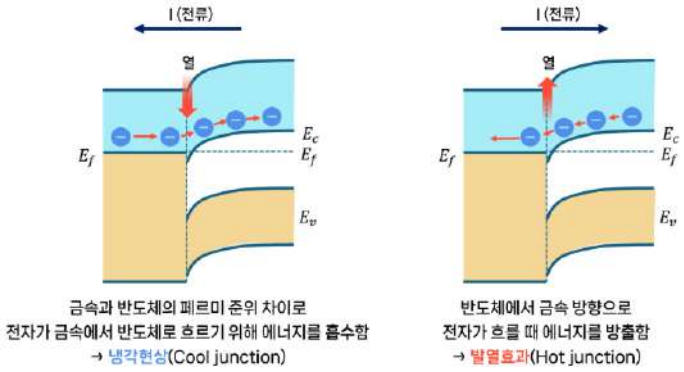


Figure 5. Peltier effect

4.Power MOSFET란 무엇일까?

4-1-1. MOS Capacitor

MOS는 Metal Oxide Semiconductor의 약자로 금속, 산화막, 반도체를 접합시킨 구조이다. 이 구조에 전압을 걸어주면 축전기(capacitor)처럼 동작하게 된다. 이를 MOS capacitor라고 부른다. 예를 들어 P형

반도체를 사용한 MOS capcitor의 경우 역방향 전압을 걸어줬을 때 P형 반도체와 oxide 접합면에 hole이 축적된다. 반대로 순방향 전압을 걸었을 때 P형 반도체의 hole이 밀려나게 되고, 고정된 이온만 남아서 depletion이 형성 된다. 더 강한 순방향 바이어스를 가하게 되면 전자가 P형 반도체 표면에 축적되므로 inversion이 발생하게 된다. 한편, 반도체의 도핑 농도가 높을 수록 같은 전압에서 depletion의 너비가 감소한다.

4-1-2. 반도체 소자의 기본 구조 : MOSFET

FET는 Field Effect Transistor의 약자로 전계 효과 트랜지스터라는 의미이다. 이는 전압을 제어하여 전류 흐름을 조절하는 스위치 증폭 기능을 한다. MOSFET은 기판의 종류에 따라 두 종류로 나뉘게 된다. 기판이 P형 반도체이면 N 채널 MOSFET(NMOS), 기판이 N형 반도체이면 P 채널 MOSFET(PMOS)이다. MOSFET을 동작시키기 위해서는 Gate에 전압을 걸어주는 것이 중요하다. gate 전압을 걸어주게 되면 앞서 설명한 MOS capacitor처럼 동작하며 source와 drain 사이에 채널(channel)이 연결된다.

4-2. Power MOSFET의 구조

큰 전력을 다루기 위해 만들어진 power MOSFET은 크게 3가지 구조가 있다. 가장 단순한 구조인 V MOSFET은 gate가 역삼각형 형태로 박혀있는 구조이다. 가장 단순하지만 전계가 꼭짓점에 집중되어 열화 문제가 발생한다. 이를 해결하기 위해 고안된 구조가 planar gate MOSFET과 trench gate MOSFET이다. 그 중 trench gate MOSFET이 낮은 저항과 높은 전류 밀도를 가지고 있지만 공정이 조금 더 복잡하고, planar gate MOSFET은 구조 공정이 단순하다는 특징이 있다. gate 전압을 인가하게 되면 p-base에는 inversion이 발생하고, n- drift 영역에는 accumulation이 발생하며 drain에서 source로 전류가 흐르게 된다.

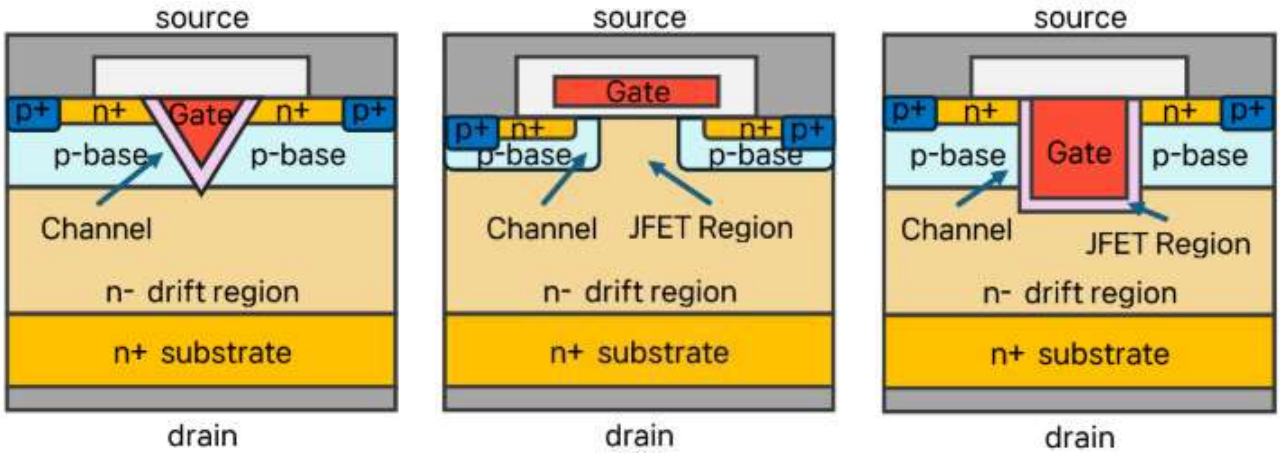


Figure 6. 왼쪽에서 부터 VMOSFET, planar gate MOSFET, trench gate MOSFET

4-2-1. JFET이란 무엇일까?

JFET은 Junction Field Effect Transistor의 약자로 P-N-P 영역으로 이루어진 트랜지스터이다. 아래 그림에서 p-base에 전압을 걸어주게 되면 평형상태에서 n-영역의 depletion이 해소 되고 전류가 흐르게 된다. power MOSFET에서도 JFET영역이 존재하게 되고, 이 영역의 저항이 존재하기 때문에 depletion 영역이 전류가 흐를 수 있는 통로의 너비를 제한시킨다.

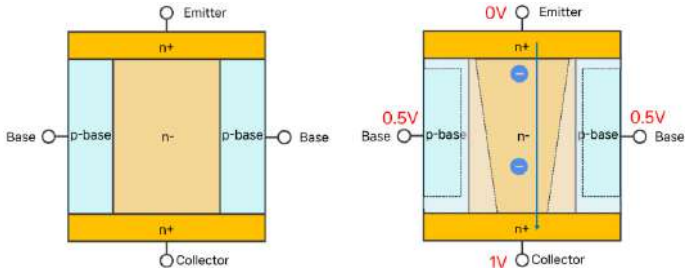


Figure 7. JFET에서 전류가 흐르는 원리

4-3. Power MOSFET에 존재하는 기생 커패시턴스(Parasitic Capacitance)

전력 반도체에는 다양한 불순물 반도체 금속들이 접합하고 있는 구조이므로 전압을 걸었을 때 내부에서 다양한 위치에서 커패시턴스 처럼 작동하게 된다. 기생 커패시턴스에는 크게 3가지 종류가 있다. gate와 source에 의해 발생하는 capacitance는 충전 방전에 직접 영향을 줘서 스위칭 속도를 결정한다. gate와 drain에 의해 발생하는 capacitance는 스위칭 할 때 전압/전류의 변화를 느리게 만든다. 마지막으로 drain과 source에 의해 발생하는 capacitance는 출력 capacitance 성분으로 동작한다.

5. IGBT

IGBT는 절연 게이트 양극성 트랜지스터(Insulated Gate Bipolar Transistor)의 약자로, MOSFET과 PIN 다이오드를 결합한 대표적인 전력반도체 소자이다. IGBT는 대전류를 처리할 수 있어 고전력 전자 시스템에서 핵심적인 역할을 한다. 특히 태양광·풍력 발전 인버터와 같이 직류 전력을 교류로 변환하는 재생에너지 시스템에서 널리 사용된다. IGBT는 Gate, Collector, Emitter의 세 단자로 구성된다. 내부적으로는 PNPN의 4층 반도체 구조를 가지며, 입력부는 MOS 구조, 출력부는 BJT(bipolar junction transistor)와 유사한 특성을 갖는다. 이러한 구조는 높은 전압 차단 능력과 낮은 도통 손실을 동시에 가능하게 한다. Gate-Emitter 전압이 문턱 전압 이상으로 인가되면 반전 채널이 형성되어 전자가 이동하고, 이에 따라 Collector 측 P 영역에서 정공이 N-drift 영역으로 주입된다. 이로 인해 원래 저항이 크던 N-drift 영역의 전하 밀도가 증가하여 저항이 감소하는 전도도 변조(Conductivity modulation)가 발생하며, 결과적으로 도통 손실이 줄어든다. 반대로 게이트 전압을 제거하면 채널이 사라지며 소자는 OFF 상태가 된다. IGBT는 입력 임피던스가 높아 게이트 구동 전력이 작고, 수백 수천 볼트의 고전압과 대전류를 처리할 수 있다. 다만 MOSFET에 비해 스위칭 속도가 느리고, 의도치 않은 전류 경로가 발생하여 회로가 제어 불가능한 상태로 빠지는 Latch-up 현상이 발생할 수 있어 회로 설계 시 주의가 필요하다. 장점으로는 높은 전압·전류 처리 능력, 낮은 구동 손실, 비교적 단순한 제어 방식이 있다. 단점으로는 스위칭 손실이 크고 고주파 동작에 부적합하다는 점이 있다. IGBT는 태양광 인버터, 풍력 발전 컨버터, 전기차 구동 시스템, 산업용 모터 제어, 무정전 전원 공급 장치(UPS) 등 다양한 고출력 전력 변환 장치에 사용된다. 특히 재생에너지 분야에서는 높은 효율과 안정적인 전력 변환을 가능하게 하는 핵심 소자로 활용된다. (Gtake, n.d.)

6. Dynamic Switching

스위칭은 전력 반도체 소자를 켜진 상태와 꺼진 상태 사이에서 전환하는 동작이다. 반도체 소자 내에서 초당 수천에서 수백만 번 반복하며 전류의 흐름을 제어한다. 전력 변환의 효율은 스위칭에 밀접한 관련이 있다. 전압/주파수 조절, 소형화와 경량화, 스위칭 손실 감소 등이 반도체 소자의 효율에 큰 영향을 미친다. 스위칭의 종류에는 크게 2종류가 있다. hard switching은 켜지거나 꺼질 때에도 전압이 높아 고주파에서 동작할 때 손실이 크다. soft switching은 스위치 타이밍을 조절하여 전압 또는 전류가 0일 때 스위칭하여 손실을 최소화 한다.

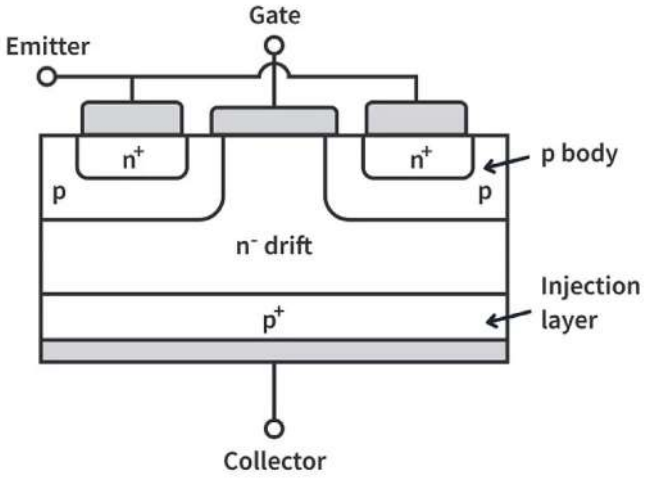


Figure 8. IGBT의 구조

6-1. Hard Switching(Inductive Switching)

hard switching은 MOSFET이 켜져있을 때와 꺼져있을 때 각각 다르게 동작한다. MOSFET이 켜진 상황에서는 전류가 전원, 부하 인덕터, MOSFET 순서로 흐르며 에너지를 축적한다. MOSFET이 꺼진 상황에서는 인덕터가 전류를 유지하려는 성질에 의해 다이오드를 통하는 환류가 발생한다.

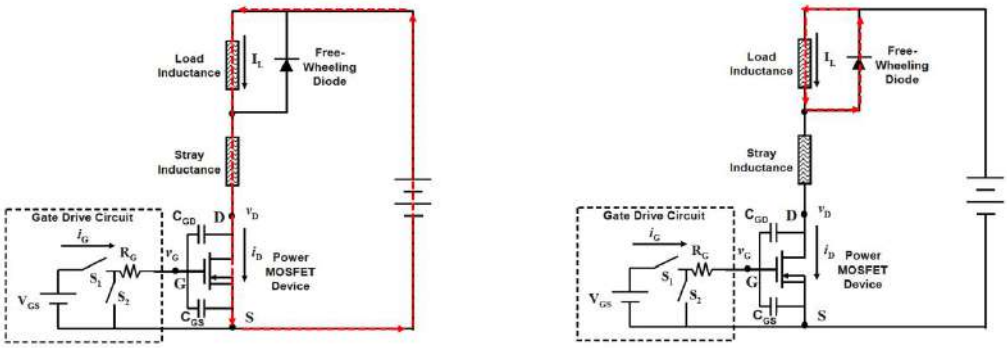


Figure 9. MOSFET의 다이나믹 스위칭 시 전류 경로 변화

7. 신재생에너지와 전력반도체

7-1. 태양광 발전 시스템의 전력 흐름

태양광 발전 시스템은 태양전지가 태양광을 받아 전기 에너지로 변환하지만, 이 출력 전력은 직류(DC) 형태이기 때문에 그대로 전력망이나 가정용 전기로 사용할 수 없다. 이를 위해 태양광 발전 시스템에는 태양광 인버터가 핵심적으로 포함된다. 인버터는 DC 전력을 교류(AC)로 변환함으로써 전력망과 동기화시켜 사용할 수 있도록 한다. 태양광 인버터의 전력 변환 과정은 크게 두 단계로 이루어진다. 전환하는 과정에서 에너지가 스위칭, 도통, 필터, 배선 등을 통해 열로 빠지게 되므로 효율이 100%가 아니다. 그렇기 때문에 최대한 발전량을 늘리기 위해 MPPT 알고리즘이 사용된다. MPPT 알고리즘이란, 일조량과 기온 변화에 따라 태양광 패널에 출력 전압과 전류를 실시간으로 샘플링 하여 항상 최대 전력을 생산하도록 제어하는 알고리즘이다. 태양관 인버터의 전력 변환의 첫번째 과정은 DC/DC 변환 단계로 여기서

MPPT(Maximum Power Point Tracking) 알고리즘을 적용해 태양전지가 최대 효율로 전력을 생산하도록 전압을 조정한다. 그 다음 MPPT로 최적화된 DC 전력을 DC/AC 변환 단계를 통해 교류로 전환한다. 이 과정을 통해 태양광 발전 시스템은 발전 효율을 높이고 에너지 손실을 줄일 수 있다. 이처럼 인버터는 단순 전력 변환기를 넘어 에너지 최적화 장치이며, 그 중심에는 전력반도체가 있다.

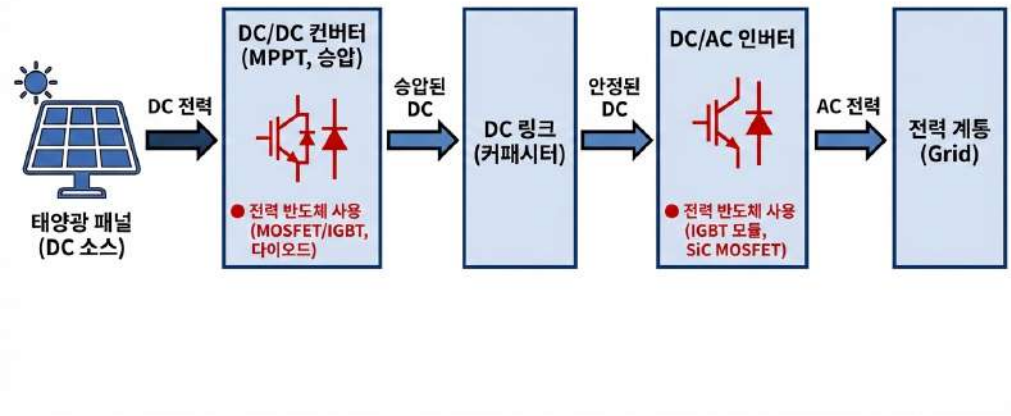


Figure 10. 태양광 발전 시스템에서 전력반도체 기반 전력 변환 흐름

7-2. 태양광 발전과 전력반도체

태양광 인버터 내부에서 전력반도체는 직류(DC) 전력을 교류(AC)로 변환하고, 계통과 연계 가능한 전력 품질을 확보하는 핵심적인 역할을 수행한다. 이 과정에서 전력반도체 소자들은 고전압·대전류 조건에서 반복적인 스위칭 동작을 수행하며, 이러한 스위칭 특성은 전력 변환 효율과 시스템 신뢰성에 직접적인 영향을 미친다. 태양광 인버터에는 주로 IGBT, MOSFET, BJT, 다이오드와 같은 전력반도체 소자가 사용되며, 각 소자는 전력 변환 단계와 제어 목적에 따라 역할이 분담된다. IGBT는 고전압·대전류 환경에서 낮은 도통 손실과 안정적인 스위칭 특성을 제공하여 MPPT 과정과 DC-AC 변환 단계에서 핵심적으로 활용된다. MOSFET은 높은 스위칭 속도와 낮은 게이트 구동 전력을 바탕으로 저전력 또는 고주파 전력 변환 구간에서 효율 향상에 기여한다. BJT는 주로 신호 제어용으로 사용되어 회로의 안정성을 확보하고 저전력 신호의 스위칭 및 증폭을 담당한다. 또한 다이오드는 전류의 단방향 흐름을 보장함으로써 역전류로 인한 소자 손상과 비정상 동작을 방지하여 시스템의 안정성을 높인다. 이러한 전력반도체 소자들의 적용은 태양광 발전 시스템의 효율과 신뢰성을 동시에 향상시키며, 분산형 에너지 시스템, 스마트 그리드, 에너지 저장 시스템(ESS)과의 통합을 촉진함으로써 재생에너지 확대와 탄소 저감 목표 달성에 중요한 역할을 한다. Figure 11은 태양광 인버터 시스템의 주요 구성 요소와 전력 흐름을 나타낸 블록 다이어그램으로, 전력반도체를 이용한 DC/DC 및 DC/AC 변환, 에너지 저장, 제어 및 통신 기능이 통합된 구조를 보여준다. (Nexperia, n.d.)

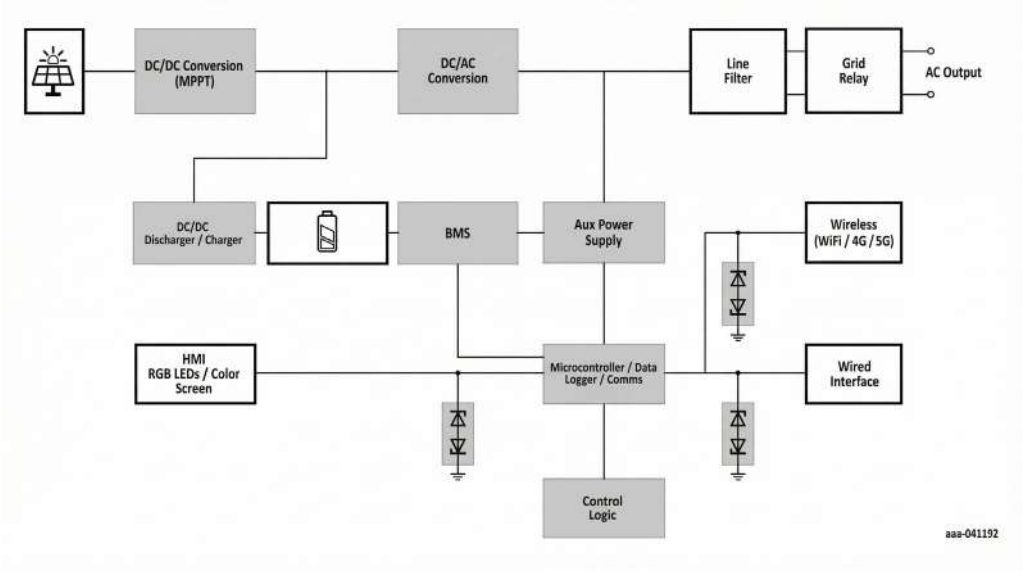


Figure 11. 태양광 인버터 시스템의 구조도

7-3. 풍력발전의 출력 특성

풍력발전 시스템의 가장 큰 특징은 출력이 일정하지 않다는 점이다. 풍속은 시간에 따라 지속적으로 변화하며, 이에 따라 발전기의 출력 전압과 전류 역시 비선형적이고 불확실하게 변동한다. 최근 대형 풍력 터빈은 6 MW 이상, 일부는 10 MW급까지 확대되고 있으며, 이러한 대용량 시스템에서는 출력 변동이 전력변환기에 큰 전기적 부하를 유발한다. 특히 풍력발전에서는 대규모 신호 변동이 빈번하게 발생한다. 이는 발전기와 계통 사이에 위치한 전력변환기가 넓은 안정 영역을 확보해야 함을 의미한다. 만약 전력반도체의 특성과 제어 전략이 적절하지 않다면, 스위칭 손실 증가와 발열로 인해 변환기의 수명이 급격히 감소할 수 있다. 실제로 풍속이 일정 수준 이상으로 증가할 경우, 전력변환기의 수명이 크게 감소할 수 있다는 연구 결과도 보고되고 있다. 이러한 특성 때문에 풍력발전 시스템에서는 단순한 정적 특성보다 전력반도체의 동적 스위칭 특성, 즉 다이나믹 스위칭 거동이 매우 중요하게 작용한다.

7-4. 풍력발전과 전력반도체

풍력발전 시스템에서 전력반도체는 발전기에서 생성된 전력을 계통에 적합한 형태로 변환하는 핵심 역할을 한다. 풍력 발전은 바람의 세기, 방향 등에 따라 전기가 일정하지 않게 만들어진다. 하지만, 현재 사용되는 전력망은 전압 및 주파수가 일정한 전기를 요구한다. 그렇기 때문에 풍력 발전 시스템에 있는 컨버터 내부의 전력반도체가 고전압·대전류 조건에서 반복적인 스위칭 동작을 수행한다. 이때 발생하는 다이나믹 스위칭 특성이 시스템 전체의 효율과 신뢰성을 결정한다. 풍력발전용 전력변환기에 주로 사용되는 소자로는 SiC MOSFET과 Si IGBT가 있다. Si IGBT는 기술이 많이 발전되어 있기 때문에 가격이 싸고 내구성이 검증되었다는 장점이 있지만, 턴 오프 신호 후에도 전류가 흐르는 테일 전류로 인한 열손실이 있다는 단점이 있다. SiC MOSFET은 높은 효율, 낮은 발열, 가벼운 무게, 빠른 스위칭 속도의 장점을 가지고 있지만, 아직 산화막의 신뢰도가 낮고 빠른 스위칭으로 인해 발생하는 노이즈가 절연 파괴를 일으킬 수 있는 문제가 있다. 풍력발전 환경에서는 다이나믹 스위칭 손실이 단순한 효율 저하를 넘어 제어 안정성 문제로 이어질 수 있다. 스위칭 과정에서 발생하는 전압 오버슈트와 손실은 전력변환기의 열적 부하를 증가시키며, 이는 시스템의 신뢰성과 수명을 제한하는 요인이 된다. 따라서 최근 풍력발전용 전력변환기에서는 전력반도체의 물리적 스위칭 특성을 고려한 적응 제어와 전력전자 빌딩 블록(PEBB) 구조가 적용되고 있다. (Yang et al., 2013)

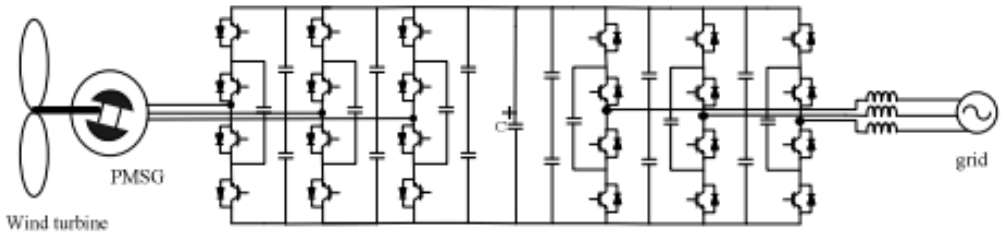


Figure 12. SiC MOSFET 컨버터를 적용한 풍력 터빈

Conclusions

본 탐구 자료에서는 신재생에너지 시스템의 핵심 기술로서 전력반도체의 역할을 반도체 물성의 기초 이론부터 실제 에너지 시스템 적용 사례까지 단계적으로 살펴보았다. 반도체의 에너지 밴드 구조와 도핑을 기반으로 한 PN 접합, 그리고 이를 확장한 PIN 접합 구조는 고전압·고전류 환경에서 안정적인 전력 제어를 가능하게 하는 전력반도체의 물리적 근간이다. 특히 PIN 접합에서 나타나는 reverse blocking과 reverse recovery 현상은 전력반도체가 단순한 스위칭 소자가 아니라, 전력 손실과 시스템 신뢰성에 직접적인 영향을 미치는 정교한 물리 시스템임을 보여준다. 또한 MOSFET과 IGBT와 같은 대표적인 전력반도체 소자들은 각각의 구조적 특성과 동작 원리에 따라 신재생에너지 시스템에서 서로 다른 역할을 수행한다. 태양광 발전에서는 높은 효율의 DC-AC 변환과 MPPT 구현을 위해 전력반도체가 핵심적으로 사용되며, 풍력발전에서는 출력 변동성이 큰 환경 속에서 전력반도체의 다이나믹 스위칭 특성이 시스템의 안정성과 수명을 결정짓는 중요한 요소로 작용한다. 이를 통해 전력반도체의 성능은 단순한 소자 수준을 넘어, 전체 에너지 시스템의 효율과 신뢰성을 좌우하는 핵심 요인임을 알 수 있다. 전력반도체는 미래 친환경 에너지 사회를 지탱하는 핵심 기술이며, 전력 손실을 최소화하고 에너지 효율을 극대화함으로써 신재생에너지의 안정적 활용과 지속 가능한 전력 시스템 구축을 가능하게 할 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

1. **전력 반도체 소자의 종류에는 *IGBT*와 *MOSFET*이 있는데 각 소자가 사용되기 적합한 미래 기술에는 무엇이 있을까*?***

2. **전력 반도체의 효율 향상이 가져올 사회와 산업에 가져올 영향은 구체적으로 어떤 것이 있을까*?***

3. **반도체의 소재로 널리 사용되는 *Si*의 한계와 이를 해결하기 위한 소재로는 어떤 것들이 있을까*?***

References

Gtake, n.d. What is IGBT? IGBT working principle. Gtake Industry News. https://www.gtake.com/ko/industry-news/what-is-igbt-igbt-working-principle/.

Nexperia, n.d. Harnessing the Sun: Semiconductors in Solar Inverters. Efficiency Wins. https://efficiencywins.nexperia.com/innovation/harnessing-the-sun-semiconductors-in-solar-inverters.

Yang, J., X. Fu, and S. Xue. 2013. Semiconductor Devices of Wind Power Converter and Its Control. International Journal of Smart Grid and Clean Energy **2**:208–214.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Hyemin Kang

M. H.KANG@kentech.ac.kr

Hyemin Kang received the Ph.D. degree in Electrical Engineering from University of Cambridge in 2020. He is currently an Associate Professor in the School of Energy Engineering at Korea Institute of Energy Technology (KENTECH) and serves as a Senior Research Collaborator at the University of Cambridge. Prior to joining academia, he worked in the power semiconductor industry at Fairchild Semiconductor and currently acts as an advisor to Magnachip Semiconductor. His research interests include power semiconductor device design and physics, focusing on Power MOSFETs, IGBTs, superjunction devices, and ultra-high-power semiconductor technologies.

KENTECH
Korea Institute of Energy Technology

Vol. 2 (2026)

Power Semiconductors: The Path Toward a Sustainable Energy Era

Abstract This study provides a comprehensive analysis of power semiconductors, which are emerging as a core technology in renewable energy power conversion systems. It examines the physical principles and structural characteristics of power semiconductors, along with their practical applications in modern energy systems. Beginning with semiconductor energy band theory and doping mechanisms, the formation of PN junctions and the principles of current control are explored. The characteristics of PIN junction structures, which are suitable for high-voltage and high-current environments, and their voltage-blocking mechanisms are then analyzed. Subsequently, the internal structures and operating principles of various power semiconductor devices are compared, and the influence of switching characteristics and parasitic elements on power loss and system efficiency is discussed. Furthermore, the role of power semiconductors in photovoltaic and wind power systems is analyzed from the perspectives of power flow and dynamic switching behavior. Through this analysis, it is demonstrated that the performance of power semiconductors is a critical factor determining the efficiency and reliability of renewable energy systems. This study aims to highlight power semiconductors not merely as power control components, but as foundational technologies enabling the realization of a sustainable energy society.

Introduction

Modern society is undergoing a major transition toward carbon neutrality, shifting rapidly from a fossil-fuel-based energy structure to one centered on renewable energy sources. Technologies that efficiently control and convert generated power have become a key indicator of national industrial competitiveness. At the heart of this energy transition are power semiconductors, which distribute and convert electrical power to meet the requirements of various systems and devices. Power semiconductors are essential components capable of withstanding extreme environments involving high voltage and high current while performing power conversion, transformation, and distribution.

Demand for power semiconductors has increased dramatically across a wide range of applications, from improving power efficiency in data centers to extending driving range in electric vehicles (EVs) and enabling smart grid infrastructures. In particular, minimizing power loss and maximizing energy efficiency are critical challenges for future industries, and achieving these goals requires a fundamental understanding of device operating principles. Accordingly, this study examines the semiconductor theories and structural characteristics that underpin power semiconductor devices. Based on this foundation, the operating principles of modern power control systems and their applications in future energy industries are explored. This work aims to provide a foundational understanding of power semiconductors not only as functional devices, but also as key enabling technologies supporting a sustainable energy society.

1. Background: Semiconductor Properties and PN Junctions

1-1. Definition of Semiconductors

Materials in our surroundings can be classified into conductors, such as copper and aluminum that allow electricity to flow easily; insulators, such as rubber and wood that impede electrical flow; and semiconductors,

which exhibit electrical properties intermediate between conductors and insulators. Semiconductors conduct electricity only under specific conditions, such as when exposed to light or an applied voltage. These characteristics make semiconductors essential components for controlling the flow of electric current.

1-2. Understanding Semiconductors Through Atomic Structure: Energy Band Theory

The electrical properties of materials are determined by the energy states of electrons within atoms. These energy ranges are referred to as energy bands. Two primary bands are the valence band and the conduction band, separated by an energy gap known as the band gap. Electrons in the valence band are strongly bound to atomic nuclei, while electrons in the conduction band can move freely and contribute to electrical current. When electrons in the valence band receive energy greater than the band gap, they transition to the conduction band. The band gap typically ranges from 0.1 eV to 4 eV, and silicon, the most widely used semiconductor material, has a band gap of approximately 1.12 eV.

1-3. Crystal Structure and Doping of Silicon

Silicon is the most commonly used material in semiconductor fabrication. Each silicon atom has four valence electrons and forms covalent bonds with four neighboring atoms, creating a crystal lattice. Pure silicon, known as an intrinsic semiconductor, has no free charge carriers at room temperature, resulting in poor electrical conductivity. To enable current flow, impurities are intentionally introduced, creating extrinsic semiconductors. Group V elements such as phosphorus (P) and arsenic (As) produce n-type semiconductors with free electrons, while group III elements such as boron (B) and gallium (Ga) create p-type semiconductors in which current flows via holes.

1-4. Formation of PN Junctions and Current Flow

Neither p-type nor n-type semiconductors alone can effectively control current flow, so they are joined to form PN junctions. Due to differences in majority carriers, diffusion and recombination occur near the junction interface, forming a depletion region. The electric field generated by fixed charges in the depletion region prevents further diffusion of carriers. When a forward bias is applied (positive voltage to the p-side and negative voltage to the n-side), the built-in potential barrier is reduced, allowing current to flow. Conversely, under reverse bias, the depletion region widens, preventing current flow. This bias-dependent behavior enables rectification.

2. Fundamental Structure of Power Semiconductors: PIN Junctions

2-1. What Is a PIN Diode?

Power MOSFETs operate in high-voltage and high-current environments, requiring structures beyond conventional PN junctions. A PIN junction inserts a lightly doped or nearly intrinsic region between the p-type and n-type regions. In practice, this intrinsic region is often a lightly doped n-type layer.

2-2. Why PIN Junctions Can Withstand High Voltage

The electric field distribution in a PN junction forms a triangular shape, and the maximum sustainable voltage corresponds to the area under the electric field curve. In a PIN junction, the lightly doped intrinsic region produces a trapezoidal electric field distribution, increasing the voltage-blocking capability in proportion to the intrinsic layer width. As a result, PIN junctions can withstand significantly higher voltages than conventional PN junctions.

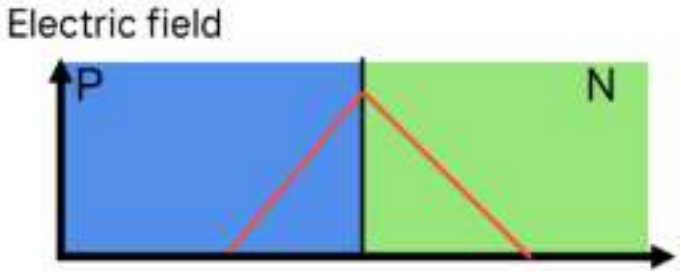


Figure 1. Electric field distribution in a PN junction

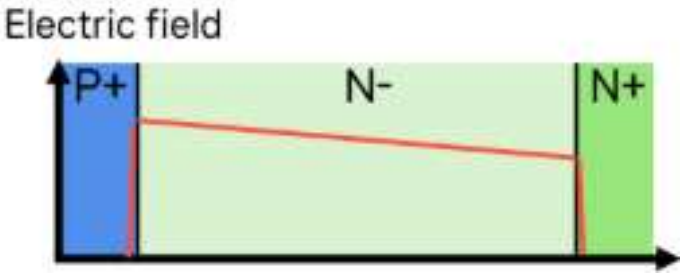


Figure 2. Electric field distribution in a PIN junction

2-3. Reverse Blocking in PIN Junctions

Under reverse bias, the depletion region in a PIN junction expands across the intrinsic layer, allowing a small leakage current to flow due to thermally generated electron–hole pairs. These carriers are swept by the strong electric field, producing leakage current known as reverse blocking. Leakage current increases with longer drift regions, higher temperatures, smaller band gaps, and shorter carrier lifetimes.

2-4. Reverse Recovery in PIN Junctions

Reverse recovery occurs when a PIN junction transitions abruptly from forward bias to reverse bias. Stored carriers in the drift region do not recombine instantly, causing a transient reverse current to flow. The time required for the device to return to the fully depleted state is known as the reverse recovery time.

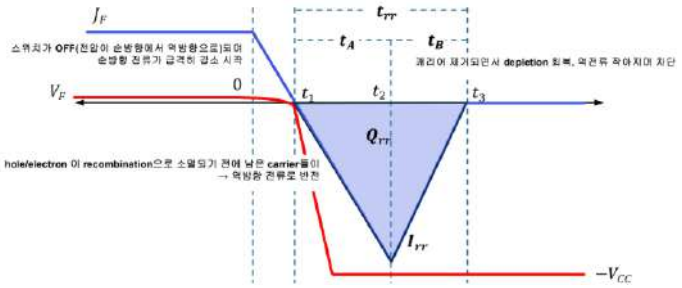


Figure 3. Temporal relationship between voltage and current during the reverse recovery process

3. Metal–Semiconductor Interfaces: Schottky Junctions and Ohmic Contacts

3-1. Schottky junction

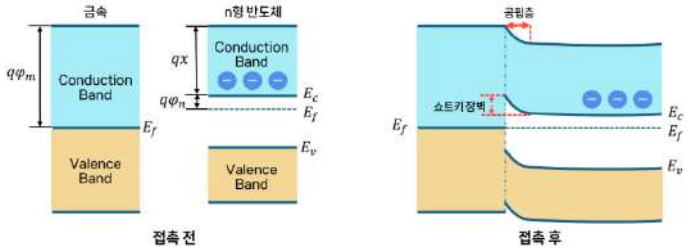


Figure 4. Schottky junction formed when an n-type semiconductor and a metal come into contact

When a metal contacts a semiconductor, differences in work function cause energy band bending, forming a Schottky barrier. In equilibrium, aligned Fermi levels prevent current flow. Forward bias lowers the barrier, allowing current conduction, while reverse bias raises the barrier, blocking current. At sufficiently high reverse bias, breakdown can occur due to strong electric fields and carrier multiplication effects.

3-2. Ohmic contact

Unlike Schottky junctions, Ohmic contacts exhibit negligible energy barriers, allowing free current flow. For example, when a metal forms an Ohmic contact with an n-type semiconductor, electrons accumulate at the interface. The Peltier effect is a representative phenomenon associated with Ohmic contacts, in which heat is absorbed or released depending on carrier direction.

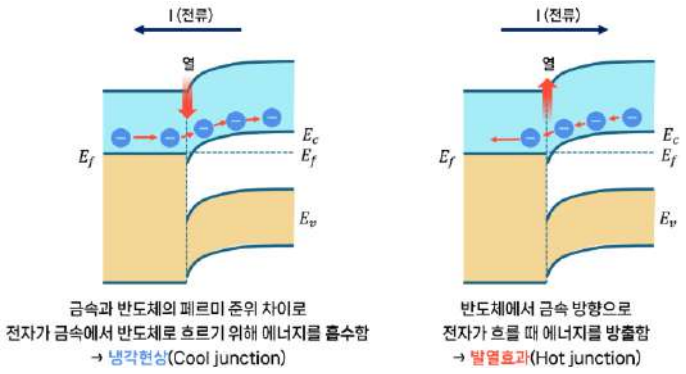


Figure 5. Peltier effect

4.What Is a Power MOSFET?

4-1-1. MOS Capacitor

MOS stands for Metal–Oxide–Semiconductor. When voltage is applied, the structure behaves like a capacitor. Depending on the applied voltage polarity, accumulation, depletion, or inversion occurs at the semiconductor–oxide interface.

4-1-2. Structure of MOSFET

FET stands for Field Effect Transistor. MOSFETs control current using an electric field applied to the gate terminal. Depending on the substrate type, they are classified as NMOS or PMOS devices. Applying a gate voltage forms a conductive channel between the source and drain.

4-2. Power MOSFET Structures

Power MOSFETs designed to handle high power can be broadly classified into three structural types. The simplest structure is the V-groove MOSFET (V-MOSFET), in which the gate is formed in an inverted triangular shape. Although this structure is simple, the electric field tends to concentrate at the vertex of the gate, leading to reliability degradation. To overcome this limitation, planar gate MOSFETs and trench gate MOSFETs were developed. Among them, the trench gate MOSFET offers lower on-resistance and higher current density, but requires a more complex fabrication process. In contrast, the planar gate MOSFET features a simpler device structure and manufacturing process. When a gate voltage is applied, an inversion layer is formed in the p-base region, while carrier accumulation occurs in the n drift region, allowing current to flow from the drain to the source.

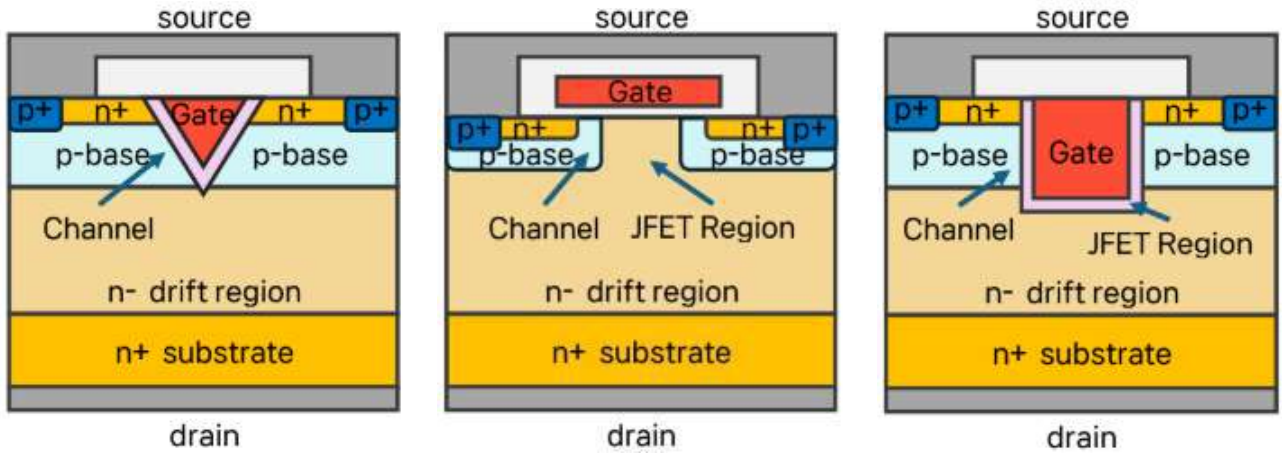


Figure 6. From the left VMOSFET, planar gate MOSFET, trench gate MOSFET

4-2-1. Whay is JFET?

A JFET (Junction Field Effect Transistor) is a transistor composed of a P–N–P region. As shown in the figure below, when a voltage is applied to the p-base, the depletion region in the n region is reduced from its equilibrium state, allowing current to flow.

In power MOSFETs, a JFET region also exists, and the resistance of this region plays an important role. The depletion region in the JFET area limits the effective width of the current conduction path, thereby affecting the overall on-resistance of the device.

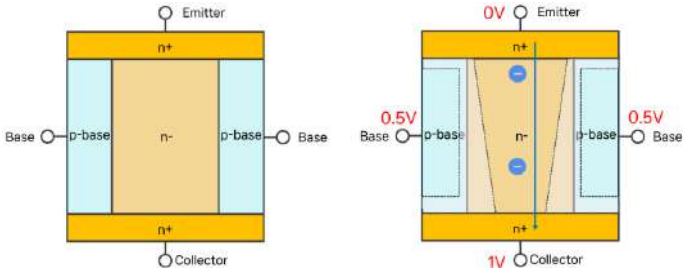


Figure 7. Principle of current flow in JFET

4-3. Parasitic Capacitances in Power MOSFETs

Power semiconductor devices consist of structures in which various semiconductor regions and metal layers are joined together. As a result, when a voltage is applied, multiple internal locations behave like capacitors. There are three major types of parasitic capacitances. The capacitance formed between the gate and the source directly affects the charging and discharging behavior, thereby determining the switching speed. The capacitance between the gate and the drain slows down the changes in voltage and current during switching. Finally, the capacitance between the drain and the source acts as an output capacitance component.

5. IGBT

An IGBT (Insulated Gate Bipolar Transistor) is a representative power semiconductor device that combines a MOSFET and a PIN diode. IGBTs are capable of handling large currents and therefore play a key role in high-power electronic systems. They are widely used in renewable energy systems such as solar and wind power inverters, where DC power is converted into AC power. An IGBT has three terminals: Gate, Collector, and Emitter. Internally, it has a four-layer PNPN semiconductor structure, where the input stage is based on a MOS structure and the output stage exhibits characteristics similar to a BJT. This structure enables both high voltage blocking capability and low conduction loss. When the Gate–Emitter voltage exceeds the threshold voltage, an inversion channel is formed and electrons begin to flow. As a result, holes are injected from the P-type collector region into the N-drift region. This increases the carrier concentration in the originally high-resistance N-drift region, reducing its resistance through conductivity modulation, which leads to lower conduction loss. When the gate voltage is removed, the channel disappears and the device turns off. IGBTs have high input impedance and therefore require low gate driving power, while being capable of handling high voltages ranging from hundreds to thousands of volts and large currents. However, compared to MOSFETs, IGBTs have slower switching speeds and may suffer from latch-up phenomena, where unintended current paths cause loss of control. The advantages of IGBTs include high voltage and current handling capability, low gate driving loss, and relatively simple control. The disadvantages include higher switching losses and unsuitability for high-frequency operation. IGBTs are widely used in solar inverters, wind power converters, electric vehicle drive systems, industrial motor control, and uninterruptible power supplies (UPS). In particular, they serve as key devices in renewable energy systems by enabling high efficiency and stable power conversion. (Gtake, n.d.)

6. Dynamic Switching

Switching refers to the operation of turning power semiconductor devices on and off. This process is repeated thousands to millions of times per second within semiconductor devices to control the flow of current. The efficiency of power conversion is closely related to switching behavior. Voltage and frequency regulation, miniaturization, weight reduction, and reduction of switching losses all have a significant impact on device efficiency.

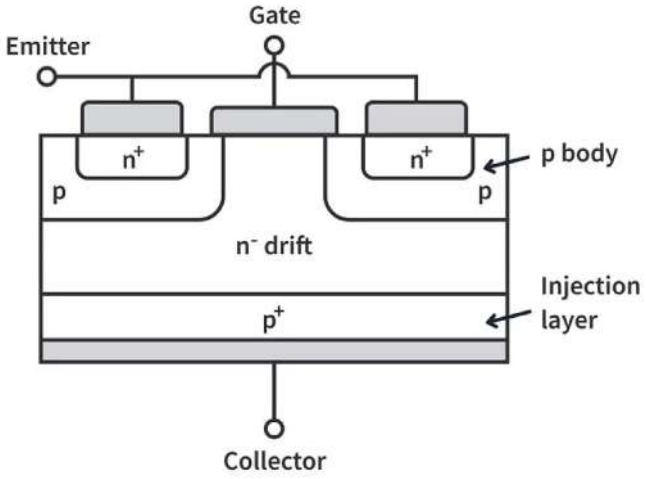


Figure 8. IGBT structure

There are two main types of switching: hard switching and soft switching. In hard switching, the voltage remains high during turn-on and turn-off, resulting in large losses at high-frequency operation. Soft switching minimizes losses by controlling the switching timing so that switching occurs when either the voltage or the current is zero.

6-1. Hard Switching(Inductive Switching)

Hard switching exhibits different behavior depending on whether the MOSFET is on or off. When the MOSFET is turned on, current flows through the power supply, load inductor, and MOSFET, storing energy in the inductor. When the MOSFET is turned off, the inductor attempts to maintain the current flow, causing a freewheeling current through the diode.

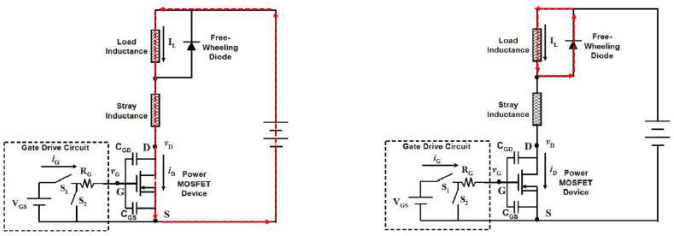


Figure 9. Current path change during dynamic switching of MOSFET

7. Renewable Energy and Power Semiconductors

7-1. Power Flow in Photovoltaic Systems

In photovoltaic power generation systems, solar cells convert sunlight into electrical energy. However, the generated power is in the form of direct current (DC), which cannot be directly used by the power grid or household appliances. Therefore, photovoltaic systems require solar inverters as a core component.

Inverters convert DC power into alternating current (AC), enabling synchronization with the power grid. The power conversion process of a solar inverter consists of two main stages. During conversion, energy is lost

as heat through switching, conduction, filtering, and wiring, so the efficiency is not 100%. To maximize power generation, Maximum Power Point Tracking (MPPT) algorithms are used.

MPPT algorithms continuously sample the output voltage and current of solar panels in real time according to changes in solar irradiance and temperature, controlling the system to always operate at the maximum power point.

First, in the DC/DC conversion stage, the MPPT algorithm adjusts the voltage so that the solar cells generate power at maximum efficiency. Second, the optimized DC power is converted into AC power through the DC/AC conversion stage. Through this process, photovoltaic systems improve generation efficiency and reduce energy losses. Thus, the inverter is not merely a power converter but an energy optimization device, with power semiconductors at its core.

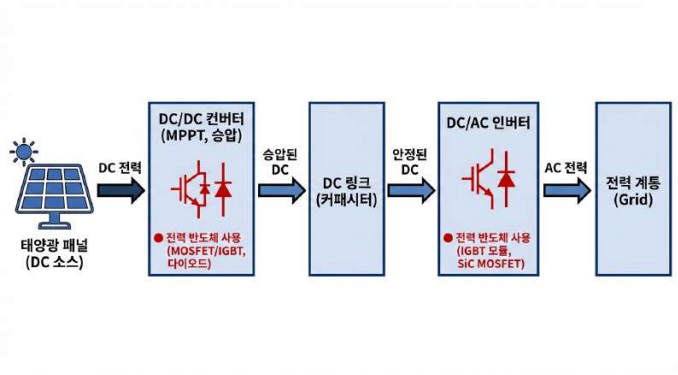


Figure 10. Power semiconductor-based power conversion flow in solar power generation systems

7-2. Photovoltaic Power Generation and Power Semiconductors

Inside a solar inverter, power semiconductor devices play a key role in converting DC power into AC power and ensuring grid-compatible power quality. During this process, power semiconductor devices repeatedly perform switching operations under high-voltage and high-current conditions, and these switching characteristics directly affect power conversion efficiency and system reliability. Solar inverters mainly use power semiconductor devices such as IGBTs, MOSFETs, BJTs, and diodes, each serving different roles depending on the power conversion stage and control objectives. IGBTs provide low conduction loss and stable switching characteristics under high-voltage and high-current conditions and are therefore essential in MPPT processes and DC–AC conversion stages. MOSFETs contribute to efficiency improvement in low-power or high-frequency conversion stages due to their fast switching speed and low gate driving power. BJTs are mainly used for signal control to ensure circuit stability and handle low-power signal switching and amplification. Diodes ensure unidirectional current flow, preventing device damage and abnormal operation caused by reverse current, thereby improving system stability. The application of these power semiconductor devices enhances both efficiency and reliability in photovoltaic systems and plays an important role in expanding renewable energy by facilitating integration with distributed energy systems, smart grids, and energy storage systems (ESS). (Nexperia, n.d.)

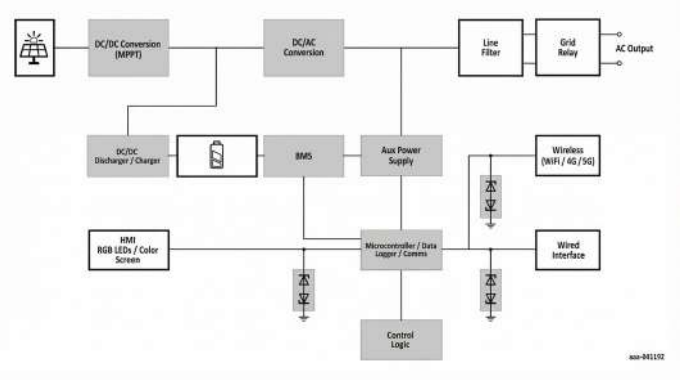


Figure 11. Structure of a solar inverter system

7-3. Output Characteristics of Wind Power Generation

The most significant characteristic of wind power generation systems is that their output is not constant. Wind speed continuously varies over time, causing nonlinear and uncertain fluctuations in the generator output voltage and current. Recently, large-scale wind turbines have expanded to capacities exceeding 6 MW, with some reaching 10 MW, and in such high-capacity systems, output fluctuations impose significant electrical stress on power converters. In wind power generation, large signal variations occur frequently, which requires the power converter located between the generator and the grid to maintain a wide stable operating region. If the characteristics of power semiconductor devices and control strategies are not properly designed, increased switching losses and heat generation can significantly reduce the converter's lifespan. Studies have reported that when wind speed exceeds a certain level, the lifetime of power converters can decrease substantially. Due to these characteristics, dynamic switching behavior of power semiconductor devices plays a far more important role in wind power systems than simple static characteristics.

7-4. Wind Power Generation and Power Semiconductors

In wind power generation systems, power semiconductor devices play a critical role in converting the electrical power generated by the turbine into a form suitable for grid connection. Wind power generation produces electricity with varying voltage and frequency depending on wind speed and direction, whereas modern power grids require stable voltage and frequency. Therefore, power semiconductor devices inside wind power converters perform repetitive switching operations under high-voltage and high-current conditions, and the resulting dynamic switching characteristics determine the overall efficiency and reliability of the system. The main devices used in wind power converters are SiC MOSFETs and Si IGBTs. Si IGBTs are relatively inexpensive and have proven durability due to their mature technology, but they suffer from thermal losses caused by tail current that persists after turn-off. SiC MOSFETs offer high efficiency, low heat generation, light weight, and fast switching speed, but they still face challenges such as limited oxide reliability and insulation breakdown caused by noise generated during fast switching. In wind power environments, dynamic switching losses can lead not only to efficiency degradation but also to control stability issues. Voltage overshoot and losses generated during switching increase the thermal load on power converters, limiting system reliability and lifetime. As a result, recent wind power converters employ adaptive control strategies that consider the physical switching characteristics of power semiconductor devices, as well as Power Electronics Building Block (PEBB) architectures.(Yang et al., 2013)

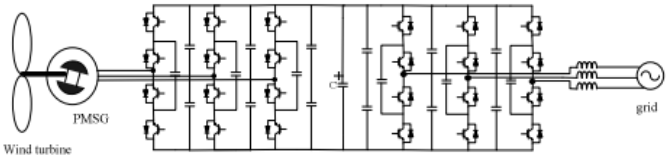


Figure 12. Wind turbines using SiC MOSFET converters

Conclusions

This study examined the role of power semiconductor devices as a core technology in renewable energy systems, progressing from fundamental semiconductor physics to practical energy system applications. PN junctions based on energy band structures and doping, as well as extended PIN junction structures, form the physical foundation of power semiconductor devices that enable stable power control in high-voltage and high-current environments. In particular, reverse blocking and reverse recovery phenomena in PIN junctions demonstrate that power semiconductor devices are not merely switching elements but sophisticated physical systems that directly affect power loss and system reliability.Representative power semiconductor devices such as MOSFETs and IGBTs perform different roles in renewable energy systems according to their structural characteristics and operating principles. In photovoltaic systems, power semiconductors are essential for high-efficiency DC–AC conversion and MPPT implementation, while in wind power systems, the dynamic switching characteristics of power semiconductors play a decisive role in determining system stability and lifetime under highly fluctuating output conditions. These observations show that the performance of power semiconductor devices goes beyond the device level and is a key factor that governs the efficiency and reliability of entire energy systems. Power semiconductors are a fundamental technology supporting future eco-friendly energy societies, enabling stable utilization of renewable energy and sustainable power system construction by minimizing power loss and maximizing energy efficiency.

Example Questions

You may research any topic other than the relevant question.

1. Among power semiconductor devices such as IGBTs and MOSFETs, which devices are suitable for future technologies, and why?

2. What specific impacts will improvements in power semiconductor efficiency have on society and industry?

3. What are the limitations of silicon (Si), which is widely used as a semiconductor material, and what alternative materials can overcome these limitations?

References

Gtake, n.d. What is IGBT? IGBT working principle. Gtake Industry News. https://www.gtake.com/ko/industry-news/what-is-igbt-igbt-working-principle/.

Nexperia, n.d. Harnessing the Sun: Semiconductors in Solar Inverters. Efficiency Wins. https://efficiencywins.nexperia.com/innovation/harnessing-the-sun-semiconductors-in-solar-inverters.

Yang, J., X. Fu, and S. Xue. 2013. Semiconductor Devices of Wind Power Converter and Its Control. International Journal of Smart Grid and Clean Energy **2**:208–214.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Hyemin Kang

M. H.KANG@kentech.ac.kr

Hyemin Kang received the Ph.D. degree in Electrical Engineering from University of Cambridge in 2020. He is currently an Associate Professor in the School of Energy Engineering at Korea Institute of Energy Technology (KENTECH) and serves as a Senior Research Collaborator at the University of Cambridge. Prior to joining academia, he worked in the power semiconductor industry at Fairchild Semiconductor and currently acts as an advisor to Magnachip Semiconductor. His research interests include power semiconductor device design and physics, focusing on Power MOSFETs, IGBTs, superjunction devices, and ultra-high-power semiconductor technologies.

KENTECH
Korea Institute of Energy Technology

VI

이미지에서 패턴을 학습하다 -

CNN의 개념과 활용

p. 126 - 146

Vol.3 (2026)

이미지에서 패턴을 학습하다: CNN의 개념과 활용

Abstract

Convolutional Neural Network(CNN)는 이미지와 같은 고차원 데이터를 효율적으로 처리하기 위해 설계된 딥러닝 모델이다. 본 자료는 CNN의 기본 구조와 핵심 개념을 정리하고, 이러한 구조적 특징이 어떻게 계산 복잡성을 줄이면서도 강력한 특징 추출을 가능하게 하는지 살펴본다. 또한 재료 설계, 에너지 시스템 관리, 재생에너지 예측, 전력망 안정성 평가 등 에너지 공학 분야에서의 활용 사례를 통해 CNN이 복잡한 문제 해결에 어떻게 적용되고 있는지를 살펴본다.

Introduction

인공 신경망(Artificial Neural Network, ANN)은 생물학적 신경계에서 영감을 받아 상호 연결된 계산 노드(n뉴런)들로 구성되며, 입력 데이터에서 학습하여 출력을 최적화한다. 입력 계층(input layer), 은닉 계층(hidden layer), 출력 계층(output layer)으로 이루어진 기본 구조를 가지며, 여러 은닉 계층을 쌓는 것을 딥러닝이라 한다.

전통적인 ANN은 이미지와 같이 차원이 높고 복잡한 정보를 처리할 때 한계를 드러낸다. 입력 데이터의 차원이 커지면 단일 뉴런이 처리해야 할 가중치의 수가 기하급수적으로 증가하기 때문이다. 이러한 복잡성 문제는 계산 비용을 크게 증가시킬 뿐만 아니라, 모델이 훈련 데이터에만 지나치게 최적화되어 실제 새로운 데이터에 대한 예측 성능이 떨어지는 과적합(overfitting)의 위험을 증가시킨다.

Convolutional Neural Network(CNN)는 학습을 통해 스스로 최적화되는 뉴런으로 구성된다는 점에서 전통적인 ANN과 유사하지만, 이미지 내 패턴 인식에 특화되어 있다. CNN은 뉴런들이 3차원(높이, 너비, 깊이)로 구성되어 각 계층의 뉴런들이 이전 계층의 작은 영역에만 연결되는 것이 특징이다. 이러한 특징은 이미지를 더 효율적으로 처리하도록 하여 모델 설정에 필요한 파라미터를 추가로 줄인다. 이러한 특성 때문에 CNN은 이미지 인식 및 패턴 분석 분야에서 핵심적인 딥러닝 모델로 활용되고 있다.(O'Shea and Nash, 2015)

What Does a CNN Do?

CNN은 생물학적 영감을 받아 이미지의 추상적인 특징을 학습하고 객체를 효율적으로 식별할 수 있는 모델이다. CNN의 핵심 특징은 이미지를 입력으로 받아 그 안에 존재하는 특징을 단계적으로 추출한다는 점이다.

예를 들어, 초기 합성곱 계층에서는 특정 방향의 모서리나 색상 변화와 같은 단순한 시각적 특징에 반응하는 필터를 학습한다. 상위 계층에서는 점점 더 복잡한 모양, 예를 들어 벌집 모양이나 바퀴 모양의 패턴과 같은 전체적인 시각적 특징을 감지하면 활성화되는 필터를 학습한다.

이러한 계층적 특징 추출 구조는 모델이 이미지의 핵심 정보를 포착할 수 있게 하여 이미지 분류, 객체 탐지, 얼굴 표정 인식 등 다양한 영역에서 뛰어난 성능을 보이는 중요한 이유 중 하나이다. (Indolia et al., 2018; Li et al., 2024)

Fundamentals of Convolutional Neural Network

CNN은 주로 convolutional layer, pooling layer, fully connected layer, activation function으로 구성된다.(Indolia et al., 2018)

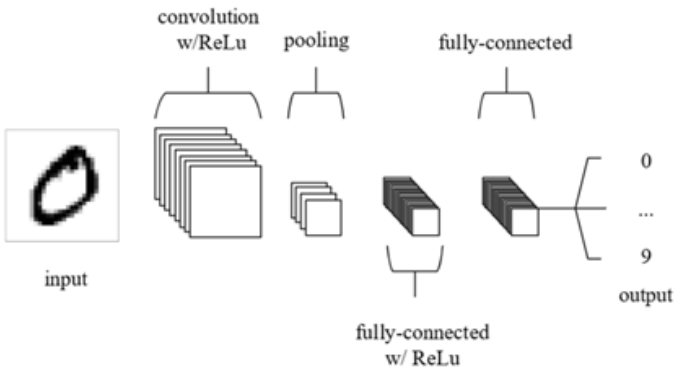


Figure 1. 간단한 CNN 구조 (O'Shea and Nash, 2015)

Convolutional Layer

합성곱 계층(Convolutional layer)은 CNN의 핵심 구성 요소로, 이미지와 같은 입력 데이터에서 특징을 효과적으로 추출하는 역할을 한다. 이 계층에서는 필터 또는 커널이라고 불리는 가중치 벡터가 입력 데이터를 위를 슬라이딩하며 특징 맵(feature map)을 추출한다. 이때 다음 계층의 각 뉴런은 이전 계층의 일부 뉴런과 연결되며, 이러한 국소적 연결 구조를 수용장(receptive field)이라고 한다. 즉, 한 뉴런이 입력에서 영향을 받는 공간적 범위를 의미하며, 입력 이미지의 국소적인 특징은 이 수용장을 통해 추출된다.

앞서 언급했듯이, 이미지와 같은 고차원 입력을 일반적인 ANN에 그대로 적용하면 매우 많은 가중치가 필요하게 된다. 예를 들어 64x64x3 크기의 RGB 이미지를 완전 연결(Fully-connected) 방식으로 처리하면 하나의 뉴런이 12288개의 가중치를 가져야 한다. 그러나 합성곱 계층에서는 수용장 크기를 6x6으로 설정한 경우 각 뉴런은 6x6x3, 즉 108개의 가중치만을 갖게 되어 모델의 복잡도를 크게 줄일 수 있다.

또한, 합성곱 계층에서는 특정 영역에서 계산하기 유용한 특징이 있다면 다른 영역에서도 유용할 가능성이 높다는 가정에 기반해 파라미터 공유(Parameter sharing)를 한다. 각 특징 맵이 동일한 가중치와 편향을 공유하게 함으로써 합성곱 계층에서 생성되는 파라미터 수가 크게 감소한다.(Indolia et al., 2018; O'Shea and Nash, 2015)

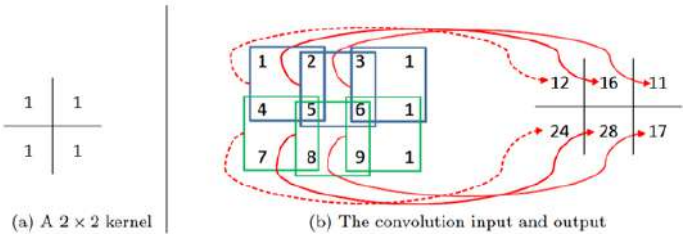


Figure 2. 합성곱 계층 작동 방식 (Wu, 2017)

Optimization of Convolutional Layer

합성곱 계층의 출력 최적화는 다음과 같은 하이퍼파라미터를 통해 이루어진다.(O'Shea and Nash, 2015; Wu, 2017)

- 깊이(Depth): 합성곱 계층의 출력 볼륨의 깊이를 수동으로 설정할 수 있다. 깊이를 줄이면 네트워크의 총 뉴런 수를 크게 최소화할 수 있지만, 모델의 패턴 인식 능력도 크게 감소시킬 수 있다.
- 스트라이드(Stride): 수용장이 입력 데이터 위를 이동하는 간격을 의미한다. 스트라이드를 1로 설정하면 수용장이 크게 겹쳐져 매우 촘촘한 출력이 생성된다. 반면 스트라이드 값을 크게 설정하면 이동 간격이 넓어져 겹침이 줄어들고, 공간적 차원이 낮은 출력을 생성한다.
- 제로-패딩(Zero-Padding): 입력 데이터의 가장자리에 0을 추가하여 출력의 공간적 크기를 조절하는 방법이다. 이를 통해 합성곱 후에도 공간적 크기를 유지하거나 원하는 크기로 조절할 수 있다.

합성곱 계층의 출력 크기는 다음과 같이 계산할 수 있다.

$$O = \frac{(V - R) + 2Z}{S} + 1$$

여기서 V는 입력 크기, R은 수용장 또는 필터의 크기, Z는 제로 패딩의 양, S는 스트라이드를 의미한다.

Table 1. 하이퍼파라미터 설정에 따른 출력 크기(O) 변화 예시

입력(V)	수용장(R)	패딩(Z)	스트라이드(S)	출력(O)	비고
32	3×3	0	1	**30**	패딩 미사용 시 크기 감소
32	3×3	1	1	**32**	입력 크기 보존 (Same Padding)
32	5×5	2	2	**16**	스트라이드에 의한 차원 축소

Pooling Layer

합성곱 계층에서 특징이 추출된 이후에는 보통 풀링 계층(Pooling layer)이 이어진다. 특징이 한번 감지되면 그 정확한 위치는 덜 중요해지기 때문이다. 풀링 계층은 파라미터의 수를 현저히 줄이고 모델이 입력 위치의 미세한 변화에 크게 영향을 받지 않도록 돕는다(Translation invariance).

풀링은 특정 크기의 영역(window)을 정해 그 영역 내에 있는 입력 요소들에 풀링 함수를 적용하여 하나의 출력 값을 생성한다. 예를 들어, 2x2 크기의 영역을 선택해 하나의 값으로 줄일 수 있다. 대표적인 방법으로는 해당 영역의 평균값을 사용하는 평균 풀링(average pooling)과 가장 큰 값을 선택하는 맥스 풀링(max pooling)이 있다. 이 중 맥스 풀링이 가장 일반적으로 사용되는 기술로, 맵 크기를 크게 줄일 수 있다.(Indolia et al., 2018; Wu, 2017)

Fully Connected Layer

완전 연결 계층(Fully connected layer)은 인접한 두 계층의 뉴런들과 직접 연결되는 뉴런들로 구성된다. 이는 전통적인 ANN의 뉴런 배열 방식과 유사하며, 최종 출력을 얻기 위해 가중치 벡터와 입력 벡터의 내적으로 계산된다.(Indolia et al., 2018)

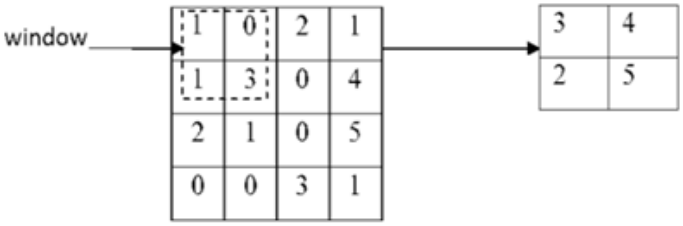


Figure 3. 2x2 window를 사용한 풀링 연산 (Indolia et al., 2018)

Activation Function

기존의 인공신경망에서는 시그모이드(sigmoid) 활성화 함수가 널리 사용되어 왔다. 그러나 이후 제안된 Rectified Linear Unit(ReLU)은 계산 효율성과 학습 안정성 측면에서 장점을 보여 널리 활용되고 있다.

ReLU는 계산이 단순하여 미분 계산이 쉬워 학습 과정에서 연산 효율이 높다. 시그모이드 함수는 $f(x) = \frac{1}{1+e^{-x}}$로 표현되며, 입력값이 매우 크거나 작아질 경우 출력이 포화 상태에 가까워지는데, 이로 인해 기울기가 매우 작아져 학습 속도가 느려질 수 있다. 반면 $f(x) = \max(0, x)$로 표현되는 ReLU는 입력이 0보다 클 때 선형적으로 증가하는 비포화 특성을 가지므로, 상대적으로 빠른 학습이 가능하다. 또한, ReLU는 양의 영역에서 기울기가 일정하게 유지되므로 기울기 소실(vanishing gradient) 문제를 완화하는 데 도움이 된다.

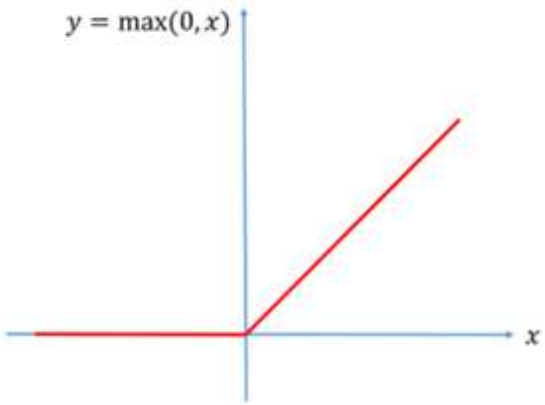


Figure 4. ReLU 함수의 시각화 (Wu, 2017)

그러나 ReLU는 학습 과정에서 일부 뉴런이 항상 0만 출력하게 되어 더 이상 활성화되지 않는 '죽은 ReLU(Dying ReLU)' 문제가 발생할 수 있다. 이 문제는 Leaky ReLU를 통해 해결할 수 있는데, 입력이 양수일 때는 기존 ReLU와 동일하게 작동하지만, 음수 영역에서도 작은 기울기를 유지하도록 하여 뉴런이 완전히 비활성화되는 현상을 줄여준다.(Indolia et al., 2018)

Applications of CNN in Energy Engineering

Material Design

CNN의 핵심 아이디어인 주변 이웃의 정보를 활용하여 각 위치의 특징을 업데이트하는 방식은 격자 구조를 가지는 이미지 데이터에만 국한되지 않는다. Crystal Graph Convolutional Neural Network(CGCNN)는 이러한 합성곱 연산을 원자 연결 그래프로 확장한 사례이다.

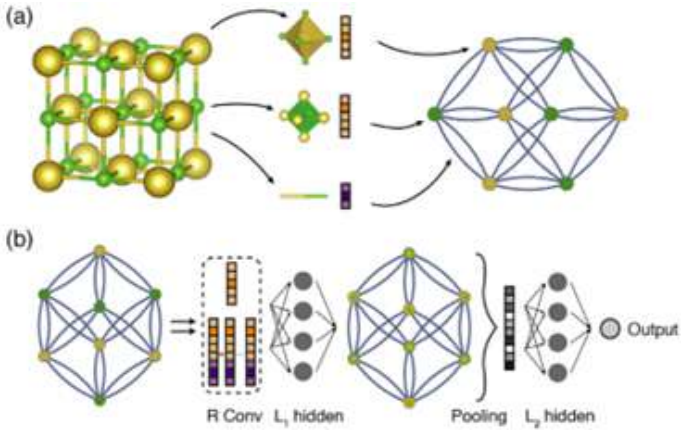


Figure 5. CGCNN 구조 (Xie and Grossman, 2018)

계산화학 기반 재료 설계 분야에서는 새로운 소재의 물성을 예측하기 위해 밀도 범함수 이론(DFT)과 같은 계산 방법을 사용한다. 그러나 DFT 계산은 정확하지만 많은 시간이 소요된다는 한계가 있다. 이를 해결하기 위해 기계학습을 활용하여 신소재 설계를 가속화하려는 시도가 이루어지고 있으며, 대표적인 예로는 CGCNN이 있다. 이 모델은 각 원자를 노드(node)로, 원자 간의 결합을 엣지(edge)로 정의하여 결정 구조를 그래프로 변환한 뒤, CNN의 합성곱 연산을 그래프 위에서 수행하여 형성 에너지, 밴드갭, 전단 계수 등 다양한 물성을 예측한다.

CGCNN은 다양한 구조와 조성을 가진 결정에 대해 DFT 계산과 유사한 높은 정확도를 보이며 매우 빠른 속도로 정확하게 예측할 수 있다. 이 접근법의 특징은 사람이 미리 설계한 복잡한 특징 벡터 없이, 결정 내 원자 연결성으로부터 직접 재료의 특성을 학습한다는 점이다. 또한, 국소 화학적 환경의 기여도를 추출하여 재료 설계를 위한 추가적인 화학적 통찰을 얻을 수 있다.(Xie and Grossman, 2018)

Energy System Management

에너지 시스템 관리 분야에서는 건물, 산업 설비, 스마트 그리드 등에서 발생하는 에너지 소비 패턴을 정확히 예측하고 효율적으로 제어하는 것이 중요하다. 최근 한 연구는 건물의 전 생애주기(계획, 설계, 건설, 운영 및 유지보수) 관점에서 에너지 소비를 관리할 수 있는 건물 정보 시스템을 이용한 CNN 기반 모델을 제안했다.

이 모델은 건물 메타 정보(용도, 시운전 시간, 층수, 건물 위치), 계량기 데이터(전력, 냉방, 증기, 난방), 기상 정보를 통합하여 학습하며, 범주형 데이터를 효과적으로 처리하기 위해 임베딩 계층을 도입하고, 시계열 특성을 반영하기 위해 1차원 합성곱을 활용하였다.(Zhou, 2022)

Renewable Energy Prediction

태양광 발전은 재생에너지 발전의 중요한 요소로, 풍부하고 환경친화적이며, 전 세계적으로 태양광 발전 시스템 설치가 가속화되고 있다. 그러나 태양광 발전으로 생성되는 전력은 건설 위치, 시간, 패널 성능 및 크기에 따라 달라질 수 있으며, 간헐적인 변동성을 가지고 있다. 기존의 태양광 발전량 예측은 일사량, 온도, 습도, 운량, 풍속 등 기상 데이터에 의존했으나, 국내 PV 발전소 대부분은 비용 문제로 환경 센서가 설치되어 있지 않아 해당 데이터를 확보하기 어렵다는 문제점이 있다. 이를 해결하기 위해 과거 발전량 시계열 데이터를 분석하여 LSTM과 CNN을 결합한 하이브리드 모델이 제안되었다.

이 모델에서 CNN은 태양광 발전량 데이터 패턴을 분석하여 현재 날씨 상태를 분류하는 역할을 수행한다. CNN에 의해 분류된 날씨 상태에 따라 LSTM은 해당 날씨에 특화된 발전량 패턴을 학습하고 예측

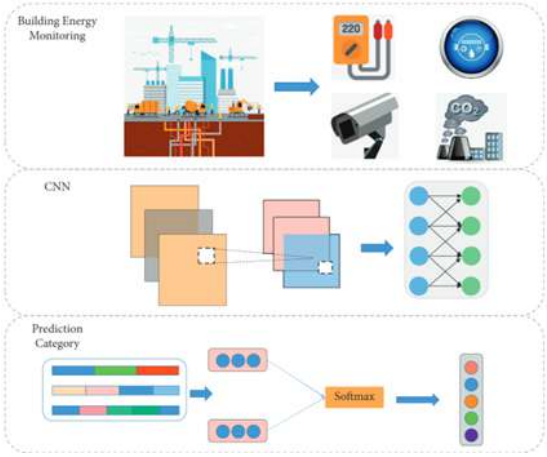


Figure 6. 건물 정보 시스템을 이용한 CNN 기반 모델 흐름도 (Zhou, 2022)

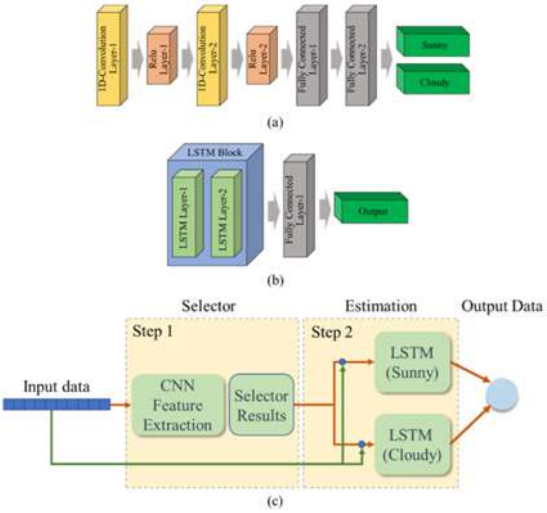


Figure 7. CNN-LSTM 하이브리드 모델 구조 (Lim et al., 2022)

한다. 이 모델은 서로 다른 신경망 구조의 장점을 결합하여 단일 모델의 한계를 극복한 사례이다.(Lim et al., 2022)

Power Grid Safety Assessment

복잡해지는 전력 시스템에서 신속한 전력 복구를 위해 정밀한 고장 분류 및 위치 파악 기술의 중요성이 증가하고 있다. 기존의 고장 검출 방법은 부하 및 네트워크 토폴로지, 조건, 설치된 장비로 인해 어려움을 겪고 있다. 최근에는 전압 파형 분석을 활용하여 전력 시스템의 고장을 감지, 분류하는 새로운 방법을 제시했다.

이 방법에서는 노이즈 감소를 위한 PFPT(Piecewise Function Put Together) 알고리즘과 Park's Transform을 사용하여 전압 파형을 회전 좌표계로 바꾸어 이미지로 변환한다. 변환된 이미지에는 고장 유형별로 고유한 기하학적 패턴이 나타나는데, CNN은 이 패턴을 학습하여 정상 상태와 고장 상태를 구분한 뒤 고장 유형을 식별한다.(Shafei et al., 2024)

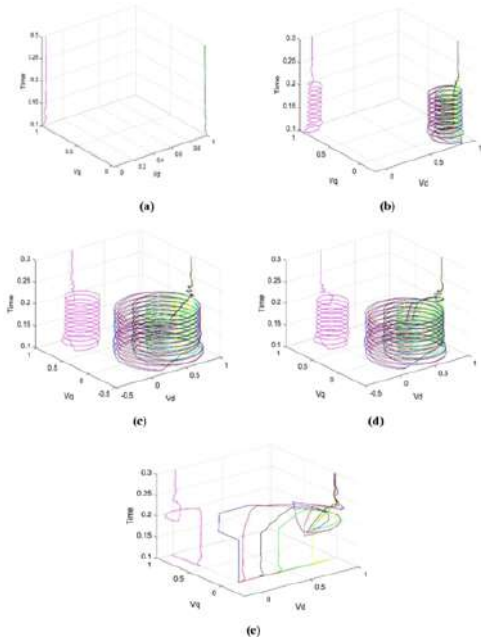


Figure 8. Park's Transform을 이용한 파형의 시각화 (Shafei et al., 2024)

Conclusion

CNN은 전통적인 ANN과 비슷하지만, 데이터의 공간적 구조를 효과적으로 활용하기 위해 설계된 구조의 신경망이다. 국소 연결과 파라미터 공유를 통해 모델의 복잡도를 크게 줄이면서도, 입력 데이터의 특징을 단계적으로 추출할 수 있다는 점이 핵심적인 강점이다. 합성곱, 풀링, 활성화 함수로 이루어지는 구조는 고차원 데이터에서 중요한 특징을 선택적으로 추출하고 일반화하는 과정을 체계적으로 구현한다.

에너지 공학 분야에서 CNN은 재료 특성 예측, 에너지 소비 관리, 재생에너지 발전량 예측, 전력망 안정성 평가 등 다양한 영역에서 활용되고 있다. 이는 CNN이 단순한 이미지 분류 모델을 넘어, 복잡한 공학 시스템의 패턴을 이해하고 예측하는 도구로 확장되고 있음을 보여준다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

Q. CNN은 이미지와 같은 고차원 데이터를 효율적으로 처리하기 위해 설계된 구조로, 합성곱 계층, 풀링 계층, 완전 연결 계층, 그리고 활성화 함수로 구성된 딥러닝 모델이다. 추가자료로 제시된 파이썬 예제 코드를 실행해 결과를 관찰한 뒤, 다음 단계를 따라 분석하고 응용해 보자.

1. **커널 크기를 *3x3, 5x5*로 변경해 학습을 수행한 뒤, 학습 곡선, 테스트 정확도, *heatmap*[1] 결과를 비교하여 커널 크기에 따른 수렴 정도를 분석한다.**

2. **합성곱 계층의 개수를 늘린 뒤 성능 변화를 비교하여 깊이에 따른 성능 향상 정도와 과적합 여부를 분석한다.**

3. **스트라이드를 변경하거나, 제로-패딩을 추가한 뒤 출력 크기와 성능을 비교하여 파라미터 값에 따른 정보 보존 정도를 분석한다.**

4. **학습 데이터의 *100%, 50%, 20%*만 사용해 학습을 수행한 뒤 성능을 비교하여 데이터의 양이 성능과 과적합 여부에 미치는 영향을 분석한다.**

5. **본문에 제시된 *CNN*의 적용 사례*(*재료 설계, 에너지 소비 예측, 재생에너지 발전량 예측, 전력망 안정성 평가*)* 중 하나를 선택하여 최적화 목표를 설정하고 *CNN* 모델을 설계한다. 이때 *CNN*을 적용했을 때 입력 데이터의 구조적 특성과 기대되는 장점, 한계점을 이유와 함께 분석한다.**

[1] 합성곱 계층의 연산 결과인 특징 맵의 활성화 수치를 색상으로 변환하여 원본 이미지 위에 겹쳐서 나타낸 것이다. 추가자료의 예제 코드에서는 밝은 노란색이나 붉은색으로 표시된 영역은 해당 필터가 찾고자 하는 특징이 강하게 나타난 지점을 의미한다.

References

Indolia, S., A. K. Goswami, S. Mishra, and P. Asopa. 2018. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. Procedia Computer Science **132**:679–688.

Li, F.-F., A. Karpathy, and J. Johnson, 2024. CS231n: Convolutional Neural Networks for Visual Recognition. URL https://cs231n.github.io/convolutional-networks/.

Lim, S.-C., J.-H. Huh, S.-H. Hong, C.-Y. Park, and J.-C. Kim. 2022. Solar Power Forecasting Using CNN-LSTM Hybrid Model. Energies **15**:8233.

O'Shea, K., and R. Nash, 2015. An Introduction to Convolutional Neural Networks.

Shafei, A. P., J. F. A. Silva, and J. Monteiro. 2024. Convolutional neural network approach for fault detection and characterization in medium voltage distribution networks. e-Prime - Advances in Electrical Engineering, Electronics and Energy **10**:100820.

Wu, J., 2017. Introduction to Convolutional Neural Networks. Lecture Notes, National Key Lab for Novel Software Technology, Nanjing University. URL https://lamda.nju.edu.cn/wujx/teaching/CNN.pdf.

Xie, T., and J. C. Grossman. 2018. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters **120**:145301.

Zhou, J. 2022. Convolutional Neural Network Based Energy Consumption Management Model for the Full Life Cycle of Buildings and Information System Design. Mobile Information Systems **2022**:1–9.

Contributors



Taegyeong Kim

M. ktg0829@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Energy AI and Environmental Climate Technology. Belonging Meta Lab, as an undergraduate researcher, I am working on catalyst optimization and novel catalyst development using machine learning. I aim to save energy and contribute to the development of a catalyst for removing harmful substances by greatly reducing the experimental cost and time required to develop a catalyst.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Seunghyun Yoon

M. syoon@kentech.ac.kr

Professor Yoon Seunghyun is affiliated with the Energy AI Track at KENTECH. His research areas include Deep Reinforcement Learning (DRL) and Multi-Agent DRL-based Artificial Intelligence (AI), as well as Cybersecurity encompassing Network Monitoring, Traffic Sampling & Inspection, Intrusion Detection, Moving Target Defense (MTD), Deception Defense, and AI-based Cybersecurity. Additionally, he is involved in Advanced Networking Technologies.

KENTECH
Korea Institute of Energy Technology

Learning Patterns from Images: Concepts and Applications of CNNs

Abstract

Convolutional Neural Networks (CNNs) are deep learning models designed to efficiently process high-dimensional data such as images. This paper outlines the basic structure and core concepts of CNNs and examines how these structural features enable powerful feature extraction while reducing computational complexity. Furthermore, this paper explores how CNNs are applied to complex problems through use cases in energy engineering, including material design, energy system management, renewable energy forecasting, and power grid stability assessment.

Introduction

Artificial Neural Networks (ANNs) are inspired by biological neural systems and consist of interconnected computational nodes (neurons). They learn from input data to optimize their output. They have a basic structure comprising an input layer, hidden layers, and an output layer. Stacking multiple hidden layers is called deep learning.

Traditional ANNs reveal limitations when processing high-dimensional and complex information, such as images. This is because, as the dimension of input data increases, the number of weights a single neuron must process grows exponentially. This complexity issue not only significantly increases computational cost but also raises the risk of overfitting, where the model becomes overly optimized for the training data, leading to poor predictive performance on new, unseen data.

Convolutional Neural Networks (CNNs) are similar to traditional ANNs in that they consist of neurons that self-optimize through learning, but they are specialized for pattern recognition within images. A key feature of CNNs is that their neurons are arranged in three dimensions (height, width, depth), with neurons in each layer connected only to small regions of the previous layer. This characteristic enables more efficient image processing, further reducing the number of parameters required for model configuration. Due to these properties, CNNs are utilized as core deep learning models in the fields of image recognition and pattern analysis.(O'Shea and Nash, 2015)

What Does a CNN Do?

CNNs are biologically inspired models that learn abstract features from images and can efficiently identify objects. The core feature of CNNs is their ability to take images as input and extract features from them step by step.

For instance, the initial convolutional layer learns filters that respond to simple visual features, such as edges in specific directions or color changes. In higher layers, filters are trained to activate when detecting increasingly complex shapes, such as overall visual features like honeycomb or wheel-shaped patterns.

This hierarchical feature extraction structure enables the model to capture key information within images, which is crucial to its outstanding performance in diverse domains such as image classification, object detection, and facial expression recognition. (Indolia et al., 2018; Li et al., 2024)

Fundamentals of Convolutional Neural Network

CNN is mainly composed of a convolutional layer, a pooling layer, a fully connected layer, and an activation function.(Indolia et al., 2018)

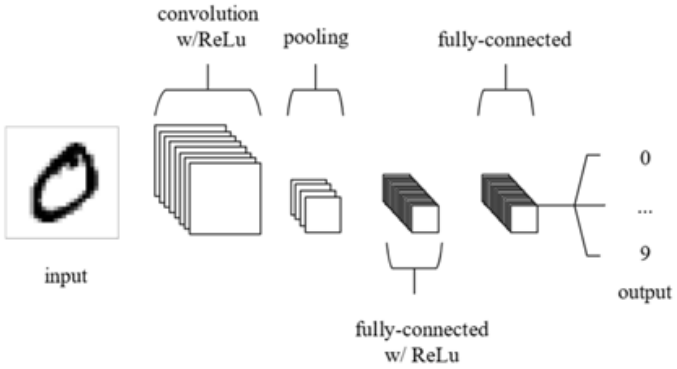


Figure 1. A simple CNN architecture (O'Shea and Nash, 2015)

Convolutional Layer

The convolutional layer is a core component of CNNs, serving to effectively extract features from input data such as images. In this layer, a weighted vector called a filter or kernel slides over the input data to extract feature maps. Each neuron in the next layer connects to a subset of neurons in the previous layer. This localized connection structure is called the receptive field. It defines the spatial range of the input that influences a single neuron, enabling the extraction of local features from the input image through this receptive field.

As mentioned earlier, applying high-dimensional inputs like images directly to a standard ANN requires an extremely large number of weights. For example, processing a 64x64x3 RGB image using a fully-connected approach would require each neuron to have 12,288 weights. However, in a convolutional layer, setting the receptive field size to 6x6 means each neuron only has 6x6x3, or 108 weights, significantly reducing the model's complexity.

Furthermore, convolutional layers employ parameter sharing based on the assumption that features useful for computation in one region are likely useful in others as well. By having each feature map share the same weights and biases, the number of parameters generated in the convolutional layer is greatly reduced.(Indolia et al., 2018; O'Shea and Nash, 2015)

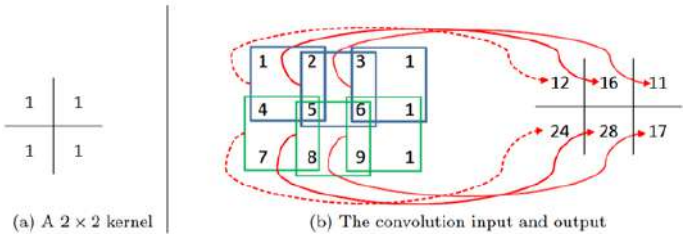


Figure 2. Convolution layer operation mechanism (Wu, 2017)

Optimization of Convolutional Layer

Output optimization of the convolutional layer is achieved through the following hyperparameters.(O'Shea and Nash, 2015; Wu, 2017)

- Depth: The depth of the output volume in the convolutional layer can be manually set. While reducing the depth can significantly minimize the total number of neurons in the network, it can also greatly diminish the model's pattern recognition capabilities.
- Stride: The stride is the interval at which the receptive field moves over the input data. Setting the stride to 1 causes the receptive fields to overlap extensively, resulting in a dense output. Conversely, increasing the stride value causes the receptive fields to move farther apart, reducing overlap and producing an output with a lower spatial dimension.
- Zero-Padding: This method adjusts the spatial size of the output by adding zeros to the edges of the input data. This allows the spatial size to be maintained or adjusted to the desired size, even after convolution.

The output size of the convolution layer can be calculated as follows:

$$O = \frac{(V - R) + 2Z}{S} + 1$$

Here, V denotes the input size, R denotes the receptive field or filter size, Z denotes the amount of zero padding, and S denotes the stride.

Table 1. Example of output size (O) change according to hyperparameter settings

Input(V)	**Receptive field**(R)	**Padding**(Z)	**Stride**(S)	**Output**(O)	**Note**
32	3×3	0	1	**30**	Size reduction when padding is not used
32	3×3	1	1	**32**	Input size preservation (Same Padding)
32	5×5	2	2	**16**	Dimension reduction by stride

Pooling Layer

After features are extracted in the convolutional layer, a pooling layer usually follows. This is because, once a feature is detected, its exact location becomes less important. The pooling layer

significantly reduces the number of parameters, helping the model become less sensitive to minor changes in input position (i.e., translation invariance).

A pooling layer defines a region of a specific size, called a window, and applies a pooling function to the input elements within that region to generate a single output value. For instance, a 2x2 region can be selected and condensed into a single value. Common methods include average pooling, which uses the mean value of the region, and max pooling, which selects the maximum value. Max pooling is the most widely used technique, as it can drastically reduce the map size.(Indolia et al., 2018; Wu, 2017)

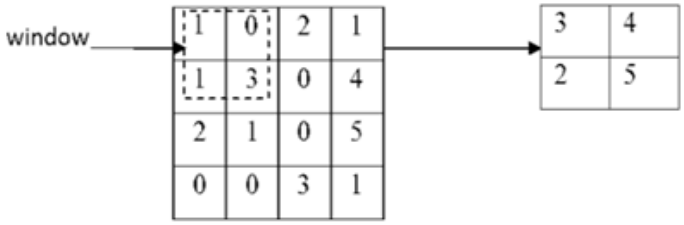


Figure 3. Pooling operation using a 2x2 window (Indolia et al., 2018)

Fully Connected Layer

A fully connected layer consists of neurons that are directly connected to neurons in the two adjacent layers. This arrangement resembles the traditional neuron layout in ANNs, and the final output is computed by the inner product of the weight vector and the input vector.(Indolia et al., 2018)

Activation Function

In traditional artificial neural networks, the sigmoid activation function has been widely used. However, the Rectified Linear Unit (ReLU), which was subsequently proposed, has demonstrated advantages in computational efficiency and learning stability, leading to its widespread adoption.

ReLU is computationally simple, which makes differentiation easy and results in high computational efficiency during training. The sigmoid function is expressed as $f(x) = \frac{1}{1+e^{-x}}$. When the input is very large or very small, the output approaches saturation, which causes the gradient to become very small and potentially slows down learning. In contrast, ReLU, expressed as $f(x) = \max(0, x)$, increases linearly when the input is greater than zero, enabling relatively fast learning due to its non-saturating characteristic. Additionally, ReLU maintains a constant gradient in the positive region, which helps mitigate the vanishing gradient problem.

However, ReLU can suffer from the “dying ReLU” problem during training, where some neurons permanently output only zero and cease to activate. Leaky ReLU can address this issue. It operates identically to standard ReLU for positive inputs but maintains a small slope in the negative region. This reduces the phenomenon of neurons becoming completely inactive.(Indolia et al., 2018)

Applications of CNN in Energy Engineering

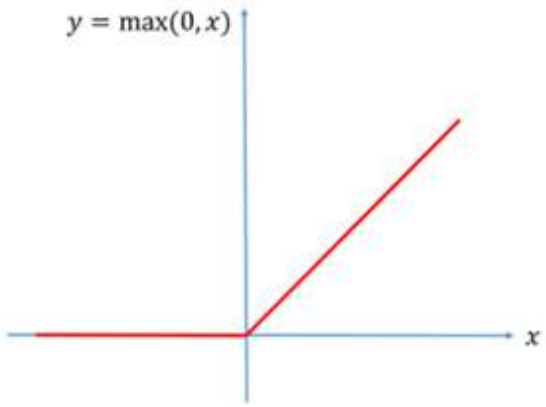


Figure 4. Visualization of the ReLU function (Wu, 2017)

Material Design

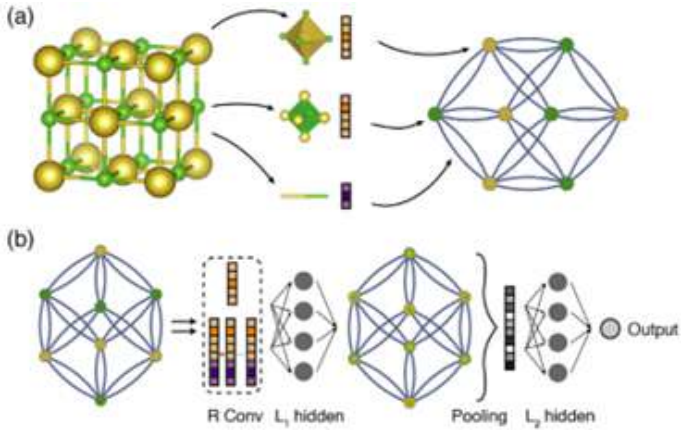


Figure 5. Illustration of the CGCNN (Xie and Grossman, 2018)

The core idea of CNNs—updating features at each location using information from nearby neighbors—is not limited to grid-structured image data. The Crystal Graph Convolutional Neural Network (CGCNN), for example, extends these operations to atomic connection graphs.

In computational chemistry-based materials design, methods such as density functional theory (DFT) are employed to predict the properties of new materials. However, while DFT calculations are accurate, they are computationally expensive. To address this issue, researchers are attempting to accelerate new material design by applying machine learning; CGCNN is a prime example of this approach. This model converts crystal structures into graphs by defining each atom as a node and each atomic bond as an edge. The model then performs CNN convolution operations on the graph to predict properties such as formation energy, bandgap, and shear modulus.

CGCNN achieves accuracy comparable to DFT calculations for crystals with diverse structures and compositions, enabling rapid, precise predictions. A key feature of this approach is that it learns material properties directly from atomic connectivity within the crystal without relying on complex, predefined human-generated feature vectors. Furthermore, CGCNN extracts contributions from the local chemical environment, providing additional chemical insights for material de-

sign.(Xie and Grossman, 2018)

Energy System Management

Accurately predicting and efficiently controlling energy consumption patterns generated by buildings, industrial facilities, and smart grids is crucial in the field of energy system management. A recent study proposed a convolutional neural network (CNN)-based model that utilizes a building information system to manage energy consumption from a building's entire lifecycle (planning, design, construction, operation, and maintenance).

The model incorporates building metadata (e.g., purpose, commissioning time, number of floors, and building location), meter data (e.g., electricity, cooling, steam, and heating), and weather information for training. The model incorporates an embedding layer to handle categorical data effectively and utilizes one-dimensional convolutions to reflect time-series characteristics.(Zhou, 2022)

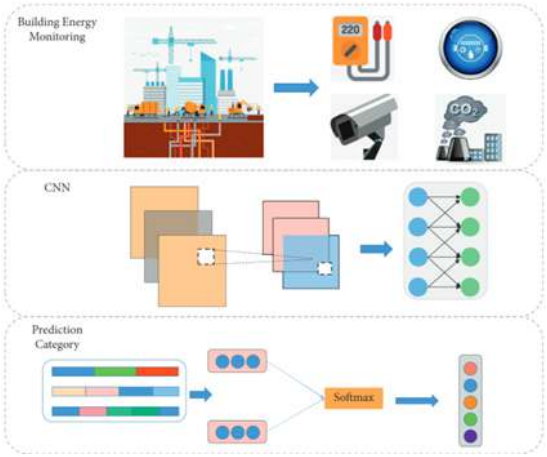


Figure 6. CNN-based model flow diagram using building information systems(Zhou, 2022)

Renewable Energy Prediction

Solar power generation is a crucial element of renewable energy production, being abundant and environmentally friendly, and the installation of solar power systems is accelerating worldwide. However, the electricity generated by solar power can vary depending on the construction location, time, panel performance, and size, and it possesses intermittent variability. Traditional solar power generation forecasting relied on meteorological data such as solar radiation, temperature, humidity, cloud cover, and wind speed. However, most domestic PV power plants lack environmental sensors due to cost constraints, making it difficult to obtain this data. To address this, a hybrid model combining LSTM and CNN was proposed by analyzing historical power generation time series data.

In this model, the CNN analyzes patterns in solar power generation data to classify the current weather state. Based on the weather state classified by the CNN, the LSTM learns and predicts generation patterns specific to that weather condition. This model represents an example of overcoming the limitations of a single model by combining the strengths of different neural network architectures.(Lim et al., 2022)

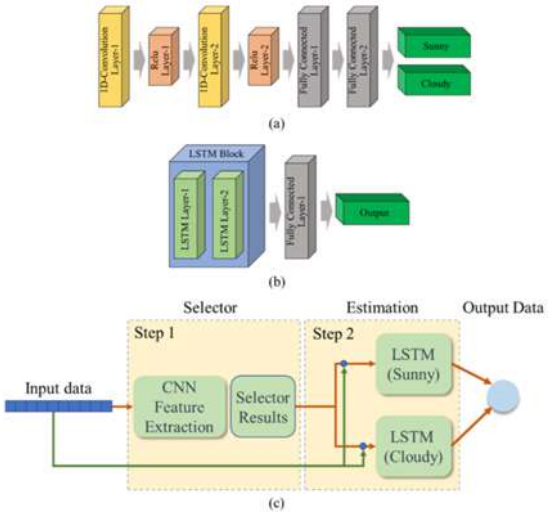


Figure 7. CNN-LSTM hybrid model architecture (Lim et al., 2022)

Power Grid Safety Assessment

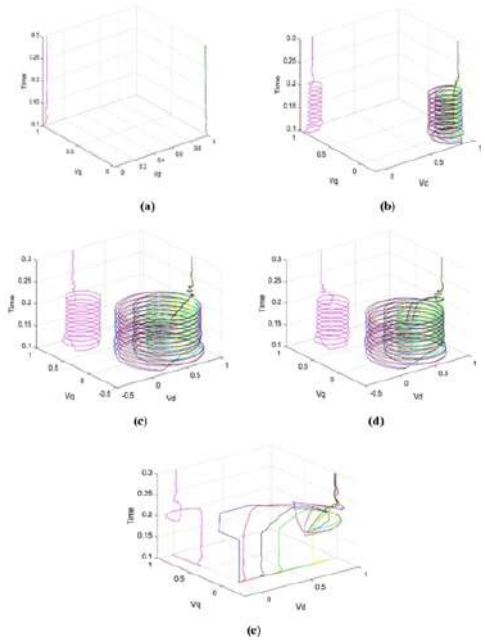


Figure 8. Visualization of waveforms using Park's transform (Shafei et al., 2024)

Precise fault classification and location detection technology is becoming increasingly important for rapidly restoring power in complex power systems. Current methods of fault detection face challenges due to load, network topology conditions, and installed equipment. Recently, a new method utilizing voltage waveform analysis to detect and classify faults in power systems has been proposed.

This method uses the PFPT (Piecewise Function Put Together) algorithm to reduce noise and Park's transform to convert voltage waveforms into images in a rotated coordinate system. The transformed image exhibits unique geometric patterns that are specific to each fault type. A

CNN learns these patterns to distinguish between normal and fault states and identify the fault type.(Shafei et al., 2024)

Conclusion

CNNs are neural networks designed to effectively utilize the spatial structure of data in a manner similar to that of traditional ANNs. They are particularly strong at significantly reducing model complexity through local connections and parameter sharing while enabling stepwise feature extraction from input data. Their structure, composed of convolution, pooling, and activation functions, systematically extracts and generalizes important features from high-dimensional data.

In energy engineering, CNNs are used in various fields, including material property prediction, energy consumption management, renewable energy generation forecasting, and power grid stability assessment. This shows that CNNs are expanding beyond simple image classification models to become tools for understanding and predicting patterns within complex engineering systems.

Example Questions

You may also conduct research on a topic of your choice other than the following question.

Q. CNN is a deep learning model designed to efficiently process high-dimensional data such as images, consisting of convolutional, pooling, fully connected layers, and activation functions. After running the Python example code provided in the supplementary materials and observing the results, follow the next steps to analyze and apply it.

1. After performing training with kernel sizes of 3x3 and 5x5, analyze the convergence rate based on kernel size by comparing the learning curve, test accuracy, and heatmap[1] results.

2. After increasing the number of convolutional layers, compare the performance changes to analyze the degree of performance improvement according to depth and whether there is overfitting.

3. Analyze the degree of information preservation according to parameter values by changing the stride or adding zero-padding and comparing the output size and performance.

4. Analyze the effect of the amount of data on performance and overfitting by performing training using only 100%, 50%, and 20% of the training data and comparing the performance.

5. Choose one of the CNN application cases presented in the text, such as material design, energy consumption prediction, renewable energy generation prediction, or power grid stability assessment. Then, set an optimization objective and design a CNN model. Analyze the structural characteristics of the input data when applying the CNN model, along with its expected advantages and limitations, providing reasons for each.

[1] This represents the activation values of the feature map, the result of the convolution layer's operation converted into colors and overlaid on the original image. In the example code from the supplementary materials, areas displayed in bright yellow or red indicate points where the desired feature is strongly present.

References

Indolia, S., A. K. Goswami, S. Mishra, and P. Asopa. 2018. Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. Procedia Computer Science **132**:679–688.

Li, F.-F., A. Karpathy, and J. Johnson, 2024. CS231n: Convolutional Neural Networks for Visual Recognition. URL https://cs231n.github.io/convolutional-networks/.

Lim, S.-C., J.-H. Huh, S.-H. Hong, C.-Y. Park, and J.-C. Kim. 2022. Solar Power Forecasting Using CNN-LSTM Hybrid Model. Energies **15**:8233.

O'Shea, K., and R. Nash, 2015. An Introduction to Convolutional Neural Networks.

Shafei, A. P., J. F. A. Silva, and J. Monteiro. 2024. Convolutional neural network approach for fault detection and characterization in medium voltage distribution networks. e-Prime - Advances in Electrical Engineering, Electronics and Energy **10**:100820.

Wu, J., 2017. Introduction to Convolutional Neural Networks. Lecture Notes, National Key Lab for Novel Software Technology, Nanjing University. URL https://lamda.nju.edu.cn/wujx/teaching/CNN.pdf.

Xie, T., and J. C. Grossman. 2018. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters **120**:145301.

Zhou, J. 2022. Convolutional Neural Network Based Energy Consumption Management Model for the Full Life Cycle of Buildings and Information System Design. Mobile Information Systems **2022**:1–9.

Contributors



Taegyeong Kim

M. ktg0829@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Energy AI and Environmental Climate Technology. Belonging Meta Lab, as an undergraduate researcher, I am working on catalyst optimization and novel catalyst development using machine learning. I aim to save energy and contribute to the development of a catalyst for removing harmful substances by greatly reducing the experimental cost and time required to develop a catalyst.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Seunghyun Yoon

M. syoon@kentech.ac.kr

Professor Yoon Seunghyun is affiliated with the Energy AI Track at KENTECH. His research areas include Deep Reinforcement Learning (DRL) and Multi-Agent DRL-based Artificial Intelligence (AI), as well as Cybersecurity encompassing Network Monitoring, Traffic Sampling & Inspection, Intrusion Detection, Moving Target Defense (MTD), Deception Defense, and AI-based Cybersecurity. Additionally, he is involved in Advanced Networking Technologies.

KENTECH
Korea Institute of Energy Technology

VII

수소 경제 시대를 대비하기 위한 수소 저장 및 운송 기술

p. 148 - 166

Vol. 4 (2026)

수소 경제 시대를 대비하기 위한 수소 저장 및 운송 기술

Abstract 기후 변화 대응과 탄소중립에 대한 관심이 높아지면서 수소 에너지는 미래 친환경 에너지원의 하나로서 주목받고 있다. 수소는 상온에서 기체 상태로 존재하며 부피 에너지 밀도가 낮아 저장과 운송에 어려움이 있어 기술적, 경제적인 단점을 가지고 있으며, 이러한 문제를 해결하기 위해 수소를 효과적으로 저장하고 운송하는 기술이 다양한 방식으로 개발되고 있으며, 대표적인 방법으로 수소 액화, 암모니아 합성, 그리고 액상 유기 수소 화합물(LOHC) 기반의 기술이 연구되고 있다.
본 탐구자료에서는 수소의 저장과 운송에서 고려되는 세 가지주요 기술들의 원리와 특징을 고찰하고, 각 기술이 가지는 장점, 단점, 그리고 한계점을 고찰하여 향후 기술개발의 방향성을 살펴보고자 한다. 특히 액화 수소의 극저온 저장 방식, 암모니아를 이용한 수소 운반 및 분해 과정, 그리고 LOHC의 수소화·탈수소화 반응의 가역성 확보를 중심으로 기술별 작동 원리를 논의한다. 또한 기술별로 실제 활용사례를 살펴봄으로서 이론적인 특징과 별개로 실용화 수준에 대해서도 고찰하고자한다.
이를 통해 수소 에너지의 저장과 운송이 수소 경제 실현을 위한 핵심 기술임을 확인하고, 효율적인 수소 운송 시스템을 구축하기 위해서는 기술적 특성과 활용 환경을 고려한 적절한 방식의 선택이 필요함을 전달한다. 본 탐구자료는 수소 에너지 기술의 과학적 원리와 공학적 과제를 이해하고, 지속 가능한 에너지 사회를 위한 미래 기술의 방향을 생각해 보는 계기를 제공한다.

Introduction

다양한 에너지 분야(Sector)에서 주로 사용되는 화석 연료는 필연적으로 이산화탄소를 배출하고 있으며, 지구 온난화와 기후변화를 가속시키고 있다. 이러한 문제의 근본적인 해결을 위해 이산화탄소 배출이 없거나, 아주 작은 수준으로 배출하는 대체 에너지원에 대한 관심이 높아지고 있으며, 그중 하나로 수소 에너지가 주목받고 있다. 수소는 사용 후 부산물로 물만 배출되기에 친환경적인 에너지원으로 주목받고 있다. 하지만, 수소는 상온에서 기체 상태로 존재하며 부피 에너지 밀도가 낮아 그대로 운송할 경우 효율이 떨어진다는 한계가 있다. 수소의 효과적인 운송을 위해서는 에너지 저장 밀도를 높이는 것이 필요하고, 이를 위해 물리적, 화학적 특성을 고려한 다양한 수소 저장 및 운송 기술이 제안되고 있다. 본 탐구자료에서는 다양한 수소 운송 방식 중 대표적인 세 가지 방법인 액화 수소 운송 방식, 암모니아 기반 수소 운송 방식, 그리고 LOHC(Liquid Organic Hydrogen Carrier) 기반 수소 운송 방식을 대상으로 그 원리와 장단점을 비교하고 실용화 사례를 분석하고, 이를 기반으로 기술개발의 방향성에 대해서 논의해보고자 한다.

1. 배경지식

1-1. 수소의 특징

수소는 주기율표 상 원자 번호 1번으로 매우 가볍고 우주 질량의 약 75% 를 차지하는 매우 풍부한 원소이다. 무색, 무취, 무미의 기체이기에 누출되었을 때 감지하기는 매우 어렵다. 끓는점은 –253 ℃, 어는점은 –259 ℃로 액체 상태로 유지하기 위해서는 극저온 환경이 필요하다. 질량기준으로 수소의 에너지 밀도는 매우 높아 가솔린의 3배 수준이지만, 분자량이 매우 작기 때문에 부피당 에너지 밀도는 매우 낮은 편이다. 수소는 탄소원자를 포함하지 않기 때문에 에너지원으로 사용될 경우 부산물로 물만 배출하는 매우

친환경적인 에너지원이라 할 수 있다. (The Engineering ToolBox, 2008)

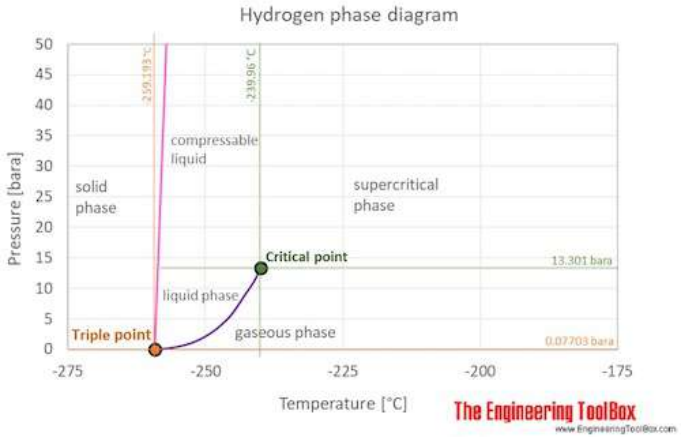


Figure 1. Hydrogen phase diagram

1-2. 친환경 에너지원으로서 수소의 가치

앞서 논의한 바와 같이 탄소중립이 시대적인 과제로 부상하면서 화석연료 중심의 에너지 경제 체제를 새로운 에너지 패러다임으로 수소 경제가 급부상하고 있으며 수소가 미래 친환경 에너지의 핵심으로 평가받는 이유는 아래 세 가지 정도로 정리할 수 있다.

- **유해물질 배출이 없는 친환경 에너지**: 석탄이나 석유 같은 화석연료는 연소 과정에서 다량의 이산화탄소와 미세먼지를 배출한다. 반면, 수소는 연료전지 내에서 산소(O_2)와 화학적, 또는 전기화학적으로 반응할 때, 오직 순수한 물(H_2O)만을 배출한다. 에너지를 변환하는 과정에서 지구 온난화의 주범인 온실가스나 유해 물질을 발생시키지 않는 청정 에너지원이라 할 수 있다.
- **에너지를 저장하는 저장고**: 태양광 에너지, 풍력에너지와 같은 재생에너지 발전은 날씨나 계절에 따라 발전량이 일정하자 않으며, 수요에 따라 공급량을 조절하기 어렵다는 근본적인 단점을 가지고 있다. 전기 에너지는 전기 자체로 저장이 불가능하기에 수요에 비해 공급이 과잉되면 전기에너지를 버려야 하며, 공급이 수요를 따르지 못하면 대규모 정전과 같은 사고가 발생하게 된다. 공급이 수요를 앞지르는 상황에서 잉여 전력을 활용하여 물을 전기 분해하고, 수소로 저장(그린수소라 통칭)하면, 장시간, 대용량으로 에너지를 저장하고 원거리로 운송도 가능하다.
- **광범위한 활용분야**: 수소는 대형 모빌리티, 철강과 화학산업 등 다양한 분야에서 활용이 가능하다. 배터리 기반 전기추진 적용이 극히 제한적인 장거리 화물트럭, 대형 선박, 항공기 등에서는 수소 연료의 사용, 특히 연료전지를 이용한 추진방식이 매루 유리하다. 그리고 수소 환원제철, 수소기반 석유 화학 및 정유공정 등 막대한 에너지원이 사용되는 분야에서 수소는 화석 연료를 대체할 수 있는 대안으로 평가받고 있다.

2. 액화 수소

액화 수소는 수소의 온도를 –253℃ 정도의 극저온(수소의 끓는점)으로 냉각시킨 액체상태의 수로를 일컫는다.

2-1. 액화 수소의 장점

액화 수소는 기체 수소 대비 부피가 1/800 수준으로 줄어들기 때문에 대용량 수송에 유리하다. 수소를 액화시키는 과정에서 불순물 제거도 가능하다는 부차적인 장점도 있다. 또한 액체 상태이기 때문에 저장용기간 이송도 빠르게 진행할 수 있더는 장점이 있다.

2-2. 액화 수소의 단점

액화과정에서 많은 전력이 소요되기에 수소가 가진 에너지(발열량 기준)의 30% 정도가 사용될 정도로 에너지 효율이 좋지 않다. 또한, -253℃의 극저온 상태로 존재하기에 아무리 단열이 잘 된 저장탱크라 할지라도 외부로부터 미세한 열침입으로 인해 기체상태로 바뀌어 외부로 배출할 수 밖에 없는 수소가 발생하며 이는 저장 관점에서 손실이 된다. 또한 극저온 수소 저장을 위해 고가의 소재를 사용해야 한다는 단점도 존재한다.

3. 암모니아 기반 수소 운반 기술

흔히 화학 비료의 원료로 알려진 암모니아가 최근 수소 경제 시대에서 유망한 저장, 이송 수단으로 각광받고 있다. 그 이유는 다루기 어려운 기체 상태의 수소를 다른 원소(질소)와 결합시켜 운반하기 쉬운 화합물의 형태로 전환하였기 때문이다.

3-1. 암모니아의 특성 : 수소와 결합한 질소

암모니아의 화학식은 NH3로 질소 원자 1개와 수소 원자 3개가 결합하여 형성된 분자이다. 암모니아가 수소 운반체로서 주목받는 가장 큰 이유는 액화가 수소 자체보다 용이하기 때문이다. 수소를 액화시키기 위해서는 –253℃ 라는 매우 낮은 온도가 필요한 반면, 암모니아는 –33℃ 수준의 온도나, 대기압의 8배 정도의 기압환경에서 쉽게 액화된다. 이러한 이유로 저장을 안정적으로 할 수 있고, 운송 비용을 줄일 수 있다.

3-2. 하버-보슈법을 이용한 암모니아 합성 과정

암모니아 합성을 위해서는 수소와 질소를 화학적으로 반응시키며, 20세기 초에 개발된 하버-보슈법(Haber-Bosch process)이 통상적으로 널리 사용된다. 합성효율을 올리기 위해 촉매를 이용하고, 고온-고압 환경에서 질소와 수소를 합성한다.

$$N_2 + 3H_2 \rightarrow 2NH_3$$

기존에는 천연가스 개질을 통해 생산된 수소(그레이 수소)가 사용되었으나, 최근에는 태양에너지, 풍력에너지로부터 생산된 수소(그린 수소)를 사용하는 시스템이 연구되고 있다.

하버-보슈법

20세기 초 프리츠 하버와 카를 보슈가 개발한 암모니아 합성법으로 공기 중에서 분리된 질소와 수소를 반응시켜 암모니아를 합성하는 방법이다. 400 500℃ 내외의 고온, 200 기압 이상의 고압 환경이 요구되며, 촉매를 사용하여 대량 생산이 가능한 시스템을 구현하였다. 이 기술을 통해 비료로 사용되는 암모니아의 대량 생산이 가능하게 되었고, 인류의 식량 문제를 해결하는데 크게 기여하기도 했었다(Valera-Medina, 2018).

암모니아에서 수소를 추출하는 과정 : 암모니아 크래킹*(Cracking)*

생산지로부터 수요지로 수송된 암모니아로부터 다시 수소를 회수하여 사용하기 위해서는 암모니아에 열을 공급하여 질소와 수소의 결합을 끊어야하며, 이 과정을 통상 암모니아 크래킹(분해)으로 지칭한다. 분해된 기체 중 질소는 공기 중으로 날려 보내고, 수소만 정제하여 다양한 수요활용처(연료전지, 터빈, 산업공정 등)에서 사용할 수 있게 된다. (Smith, 2020)

3-3. 암모니아를 사용한 수소 수송의 장점

앞서 언급한 바와 같이, 암모니아는 가장 널리 사용되는 비료이다. 따라서 이미 전 세계적으로 다양한 지역에서 생산되고, 다양한 지역으로 이송되고 있다. 다시 말해 운반선, 저장 탱크, 하역작업, 파이프라인 등 기존의 인프라를 그대로 활용 할 수 있다는 장점이 있기에 암모니아 형태로 수소를 이송하는 것은 많은 장점을 가진다. 또한 같은 부피의 액화 수소보다 1.5배 정도의 수소를 더 이송할 수 있고, 화합물에 탄소 원소가 포함되어 있지 않기 때문에 사용단계에서 원천적으로 이산화 탄소를 발생시키지 않는다는 장점도 가진다.

3-4. 암모니아를 사용한 수소 수송의 단점

암모니아는 독성물질이므로 취급에 있어 주의를 요하며, 누설시 인체와 접촉하지 않도록 방지대책을 수립하여야 한다. 또한 수소를 암모니아로 합성하는 과정과 암모니아를 분해하여 수소롤 전환하는 과정에 많은 에너지가 투입되기에 전주기 효율이 떨어진다는 단점을 가진다.

4. 액상 유기 수소 화합물(LOHC, Liquid Organic Hydrogen Carrier)

수소를 안전하게 운반하기 위한 방은의 하나로 독성물질인 암모니아나 극저온 단열이 필요한 액화수소 대신 석유화학 물질과 비슷한 성질을 가진 유기확학물이 주목받고 있고, 이러한 물질을 통칭하여 액상 유기 수소 화합물(LOHC)이라 부른다. 수소를 화학적으로 머금었다가 필요할 때 분리해 내는 물질이 LOHC이며, 기존에 익숙하게 사용되는 석유류 제품과 유사한 물리적, 화학적 특성을 가지므로 취급이 상대적으로 용이하는 장접을 가진다.

4-1. LOHC 기술이란?

LOHC(Liquid Organic Hydrogen Carrier) 기술은 상온과 대기압(또는 상대적으로 낮은 압력)에서 액체 상태를 유지하는 유기 화합물을 이용해 수소를 저장하고 운송하는 방싱을 일컫는다. 수소를 포함한 화합물(수소화 물질)을 수소 사용처로로 운송한 뒤 다시 수소를 분리하여 사용처로 공급하고, 수소가 빠져나간 빈 화합물은 다시 수소 생산처로 가져가 재사용하는 방식이다.

4-2. 수소 저장 및 추출의 원리

LOHC 기술은 화학 반응이 양방향으로 일어날 수 있는 가역 반응(Reversible Reaction)을 핵심 원리로 사용한다. 대표적인 LOHC 물질은 톨루엔(toluene)과 메틸사이클로헥산(methylcyclohexane)으로 두 물질이 가역 반응을 통해 수소를 저장(수소화) 및 추출(탈수소화)할 수 있다. (Dong et al., 2021)

4-2-1. 수소화 반응*(Hydrogenation)*을 통한 수소 저장

수소가 비어있는 불포화 화합물(톨루엔)에 수소 기체를 반응시켜, 수소를 포화 화합물(메틸사이클로헥산)로 생성한다. 이 과정은 발열반응이므로, 촉배를 이용하여 반응 효율을 향상시킨다. 또한, 작동온도를 유지하기 위해 열을 외부로 배출시켜야 한다.

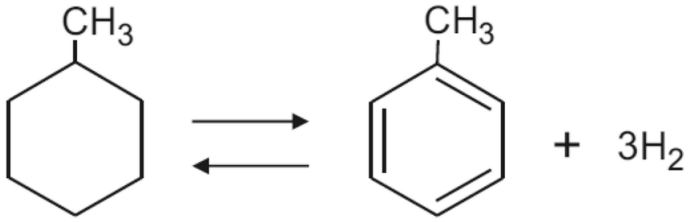


Figure 2. The hydrogenation and dehydrogenation of toluene to methylcyclohexane

4-2-2. 탈수소화 반응*(Dehydrogenation)*을 통한 수소 추출

목적지에 도착한 후 메틸사이클로헥산을 250℃ 이상으로 가열하여 열을 공급하면, 수소를 붙잡고 있는 화학 결합이 끊어지면서 머금고 있던 수소가 외부로 빠져나오게 된다. 아 반응은 열을 흡수해야만 반응이 진행되는 흡열 반응이기에 외부로부터 열을 공급해주어야 한다. 분리된 수소는 연료로 사용하고, 원래 상태로 돌아온 톨루엔은 다시 수소를 담기 위해 회수된다.

4-3. LOHC 방식의 장점

상온, 대기압 조건에서 석유류와 같이 액체 상태로 존재하기에 보관이 용이하다. 또한, 암모니아처럼 강한 독성을 갖지 않기에 안정성도 우수하다. LOHC 물질은 일반적인 석유류와 유사한 특성을 가지기에 기존에 석유류를 이송하던, 유조선, 탱크로리 트럿, 주유소의 파이프라인과 저장 탱크를 큰 변경 없이 사용할 수 있다는 장점을 가지며, 초기 인프라 구축 비용을 획기적으로 줄일 수 있다. 또한 저장물질을 재사용 할 수 있기 때문에 경제적으로도 효율적이다.

4-4. LOHC 방식의 단점

수소 추출 시 250℃ 내외의 높은 온도가 요구되기에 외부 열원이 필요하며, 이는 에너지 소모로 연결된다. 수소를 저장하는 화합물 자체의 무게가 무겁기 때문에 장거리 수송에서는 다소 불리하다. 특히 전체 중량대비 수소의 저장비율은 6% 정도로 낮다는 단점이 있다.

5. 수소 운송 기술의 실제 적용과 선택 전략

앞서 살펴본 바와 같이, 액화수소, 암모니아, 그리고 LOHC 기반의 수소 운송은 각각 고유한 장점과 한계를 지닌다. 따라서 운송 거리, 규모, 활용 목적에 따라 적합한 기술을 선택하는 것이 중요하다. 각 수소 운송 기술의 실제 적용 사례를 통해 기술별 활용 조건과 한계를 비교하고, 상황에 적합한 운송 전략의 방향성을 살펴보고자 한다.

5-1. 액화 수소 기반 수소 운송

액화 수소 기반 수소 운송의 대표적인 사례로는 일본의 HySTRA(Hydrogen Energy Supply-chain Technology Research Association) 프로젝트가 있다. 이 프로젝트는 호주에서 생산된 수소를 액화하여 일본으로 해상 운송하는 국제 수소 공급망을 실증하는 것을 목표로 하였다. HySTRA 프로젝트에서는 갈탄을 활용해 수소를 생산한 뒤 이를 정제하고, 수소의 끓는점(253℃)의 극저온으로 액화하여 전용 운반선을 통해 일본으로 운송하였다. 이후 일본에서 액화된 수소를 하역하고 저장하다. 이 프로젝트에서 액화 수소 방식이 채택된 이유는 크게 세 가지로 볼 수 있다. 첫째, 프로젝트를 주도한 Kawasaki Heavy Industries는 극저온 저장 및 운송 분야의 기술적 기반을 보유하고 있어 액화 수소 공급망 구축에 유리했다. 둘째, Liquid Organic Hydrogen Carrier(LOHC)나 암모니아 기반 방식은 수소 추출 이후 운반 매개체의 회수·재순환 과정이 필요해 추가적인 물류 부담과 에너지 손실이 발생할 수 있다. 이에 비해 액화 수소는 이러한 추가

공정 없이 직접 운송이 가능하다는 장점이 있다. 셋째, 일본은 운송된 수소를 연료전지 등 고순도 수소 활용 분야에 적용하고자 했기 때문에, 크래킹이나 탈수소화 과정 없이 바로 사용할 수 있는 액화 수소가 선택되었다. (CO2-free Hydrogen Energy Supply-chain Technology Research Association (HySTRA), 2021)

5-2. 암모니아 기반 수소 운송

암모니아 기반 수소 운송의 대표 사례는 사우디아라비아와 일본 간 블루 암모니아 공급 프로젝트이다. 이 방식에서는 수소를 암모니아(NH_3)로 변환한 뒤 액체 상태로 저장·운송하고, 목적지에서 다시 수소로 분해하여 사용한다. 암모니아는 비교적 온화한 조건(33℃ 또는 낮은 압력)에서 액화가 가능하며, 이미 전 세계적으로 비료 산업을 통해 구축된 운송·저장 인프라를 활용할 수 있다. 또한 암모니아는 수소를 화학적으로 저장한 형태로, 필요 시 촉매 반응을 통해 수소를 다시 추출할 수 있다. 또한, 이 산업에서 암모니아는 직접 연소를 통해 에너지도 제공하고, 수소 캐리어의 역할도 수행한다. 이 과정을 통해 크래킹을 위해 사용되는 에너지를 줄여서 높은 효율을 만들 수 있었다. (가스신문, 2021)

5-3. LOHC 기반 수소 운송

LOHC 기반 수소 운송의 대표적인 사례로는 일본의 Chiyoda Corporation이 개발한 SPERA Hydrogen 프로젝트가 있다. 이 프로젝트는 수소를 톨루엔과 반응시켜 메틸사이클로헥산(MCH) 형태로 저장한 뒤 액체 상태로 운송하고, 목적지에서 다시 수소를 추출하는 방식으로 운영된다. 실제 실증에서는 브루나이에서 생산된 수소를 LOHC 형태로 일본까지 운송하는 데 성공하였으며, 이는 세계 최초의 LOHC 기반 국제 수소 공급망 사례로 평가된다. 이 프로젝트의 목적이 수소를 가솔린과 같이 상압, 상온 환경에서 액체 상태로 만들어 운반하는 것이기 때문에 LOHC가 사용되었다. 또한, 이 회사가 보유하고 있던 거대한 원유 저장 탱크, 화학제품 운반선, 파이프라인을 그대로 활용할 수 있다는 장점을 가지고 있다. LOHC를 활용한다면 오랫동안 보관할 수 있고, 암모니아와 달리 독성이 약해서 일반 거주지에 운반 및 저장하는 데 장점이 있어 사용되었다. (Chiyoda Corporation, 2020)

Conclusions

기후 위기를 극복할 차세대 친환경 에너지원으로 주목받는 수소는 저장과 운송이라는 거대한 기술적 장벽을 가지고 있다. 이를 극복하기 위해 극저온 냉각 방식인 액화수소, 높은 수소 저장 밀도와 기존 인프라 활용이 장점인 암모니아, 그리고 상온에서 일반 액체처럼 안전하게 다룰 수 있는 LOHC(액상 유기 수소 화합물) 기술이 각각의 뚜렷한 장단점을 바탕으로 활발히 연구되고 있다. 상황에 따라 사용되는 기술이 달라지며 운송 거리와 규모, 사용 목적에 따라 여러 기술을 적절히 조합하여 수소 운송 인프라를 구성하는 것이 미래 수소 경제를 앞당기는 가장 현실적인 방안이다. 그리고 이러한 기술의 상용화를 위해 과학자들이 반드시 풀어야 할 숙제도 남아있다. 미래의 수소 경제가 진정한 친환경적 가치를 지니려면 수소를 생산하고, 저장하고, 운송하며, 사용하는 모든 전 주기에서 탄소 배출이 완벽하게 통제되어야 한다. 이번 탐구를 통해 수소 에너지가 가진 무한한 잠재력과 그 이면에 숨겨진 과학적, 공학적 과제들을 이해하고, 지속 가능한 미래를 위한 화학 기술의 역할에 대해 깊이 고민해 볼 수 있기를 기대한다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

1.* 특정 상황에 가장 적합한 수송 방식은 무엇일까*?

운송 거리, 규모, 활용 목적이 다른 상황을 직접 가정하고, 액화수소·암모니아·LOHC 중 가장 적합한 방식을 비교·분석해보자.

2.* 암모니아 크래킹 과정 후 남는 불순물은 어떤 화학적·전기화학적 원리로 연료전지의 성능을 저하시킬까*?

특히 촉매 표면이나 전해질 막에서 어떤 반응이 일어나 손상이 발생할까?

3.* 수소는 생산 방식에 따라 그린 수소, 블루 수소, 그레이 수소 등으로 구분할 수 있다, 각 수소의 생산 방식과 탄소 배출량은 어떻게 다를까*?* 또한 장기적으로 가장 지속 가능한 수소 생산 방식은 무엇일까*?

4.* 본문에서 설명한 톨루엔–메틸사이클로헥산 외에 다른 *LOHC* 물질은 어떤 것이 있으며, 각각의 수소화 *(Hydrogenation)*와 탈수소화*(Dehydrogenation)* 반응 조건은 어떻게 다를까*?

References

Chiyoda Corporation, 2020. The World's First Global Hydrogen Supply Chain Demonstration Project. https://www.chiyodacorp.com/en/service/lowcarbon/hydrogen/demonstration/.

CO2-free Hydrogen Energy Supply-chain Technology Research Association (HySTRA), 2021. Hydrogen Energy Supply Chain Pilot Project between Australia and Japan. https://www.hystra.or.jp/en/project/.

Dong, Z., A. Mukhtar, and H. Lin. 2021. Heterogeneous Catalysis on Liquid Organic Hydrogen Carriers. Topics in Catalysis **64**.

Smith, C. e. a. 2020. Ammonia as a transport vector for hydrogen. Energy Environmental Science .

The Engineering ToolBox, 2008. Hydrogen - Thermophysical Properties. https://www.engineeringtoolbox.com/hydrogen-d_1419.html.

Valera-Medina, A. e. a. 2018. Ammonia for power. Progress in Energy and Combustion Science .

가스신문. 2021. 암모니아 기반 수소 운송 관련 산업 동향. 가스신문 https://www.gasnews.com/news/articleView.html?idxno=94639.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



YoungDuk Lee

M. ydlee@kentech.ac.kr

Prof. Dr. Young Duk Lee received his Ph.D. in Energy Engineering from Technische Universität Berlin in 2015. He is currently an Associate Professor in the Department of Energy Engineering at the Korea Institute of Energy Technology (KENTECH). He serves as an Associate Editor of the Korean Hydrogen & New Energy Society and as a Scientific Committee Member of the ECOS international society. Before joining academia, he worked as a Principal Researcher at the Korea Institute of Machinery & Materials (KIMM) from 2005 to 2021. His research interests include fuel cells, electrolysis, hybrid energy systems, process simulation, exergy analysis, and life cycle assessment, with a particular focus on CO_2-free hydrogen production and fuel-cell-based hybrid power generation.

KENTECH
Korea Institute of Energy Technology

Vol. 4 (2026)

Hydrogen Storage and Transportation Technology to Prepare for the Era of the Hydrogen Economy

Abstract With the intensifying global focus on climate change mitigation and carbon neutrality, hydrogen energy is garnering significant attention as a pivotal eco-friendly energy source for the future. However, hydrogen inherently possesses technical and economic drawbacks; since it exists as a gas at room temperature with low volumetric energy density, it presents substantial challenges for storage and transportation. To address these issues, various technologies for the efficient storage and transport of hydrogen are being developed, with liquefied hydrogen, ammonia synthesis, and Liquid Organic Hydrogen Carrier (LOHC) systems serving as the primary subjects of current research.

This study aims to examine the fundamental principles and characteristics of these three key technologies, exploring the future direction of technological development by analyzing their respective advantages, disadvantages, and limitations. Specifically, the discussion will focus on the technical operating principles of each method: cryogenic storage for liquefied hydrogen, the transportation and cracking processes for ammonia, and the reversibility of hydrogenation and dehydrogenation reactions in LOHC. Furthermore, by reviewing real-world implementation cases, this paper evaluates the current level of commercialization beyond theoretical frameworks.

Through this analysis, this study reaffirms that hydrogen storage and transportation are the cornerstone technologies for realizing a hydrogen economy. It suggests that establishing an efficient transport system requires the strategic selection of methods based on technical characteristics and application environments. Ultimately, this research provides an opportunity to understand the scientific principles and engineering challenges of hydrogen technology and to contemplate the future of technology for a sustainable energy society.

Introduction

Fossil fuels, which are predominantly used across various energy sectors, inevitably emit carbon dioxide (CO_2), thereby accelerating global warming and climate change. To fundamentally address these issues, there is growing interest in alternative energy sources with minimal or zero carbon emissions, with hydrogen energy emerging as a prominent solution. Hydrogen is recognized as an eco-friendly energy source because it produces only water as a byproduct after use.However, hydrogen exists in a gaseous state at room temperature and possesses low volumetric energy density, which poses significant limitations on transportation efficiency. To ensure effective hydrogen transport, it is essential to increase energy storage density; consequently, various storage and transport technologies considering physical and chemical properties have been proposed.This study examines three representative hydrogen transportation methods: Liquefied Hydrogen, Ammonia-based transport, and Liquid Organic Hydrogen Carrier (LOHC) systems. By comparing their principles, advantages, and disadvantages, and analyzing practical implementation cases, this paper aims to discuss the future direction of technological development in the hydrogen sector.

1. Background Knowledge

1-1. Charateristic of Hydrogen

Hydrogen is the first element on the periodic table with an atomic number of 1. It is exceptionally light and an incredibly abundant element, accounting for approximately 75% of the universe's mass. As a colorless, odorless, and tasteless gas, leaks are difficult to detect. It has a boiling point of -253℃ and a freezing point of -259℃, necessitating cryogenic temperatures and high-pressure environments to maintain it in a liquid state. Hydrogen possesses a remarkably high gravimetric energy density (energy density by mass), which is approximately three times that of gasoline. Conversely, its volumetric energy density is extremely low, equating to roughly 1/3000th that of gasoline. Finally, hydrogen is a highly eco-friendly energy source, as the only byproduct generated during its utilization is pure water. (The Engineering ToolBox, 2008)

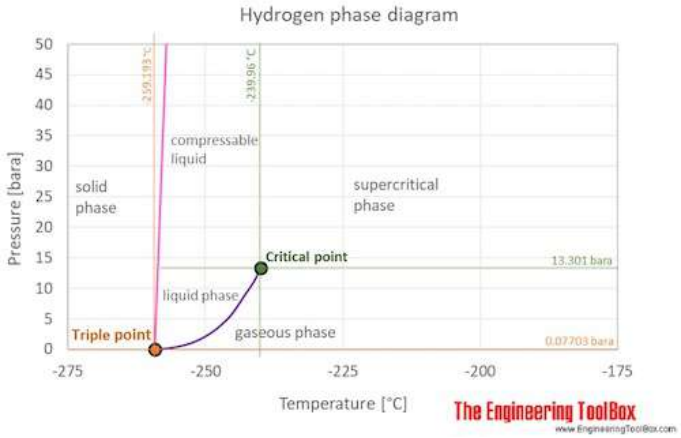


Figure 1. Hydrogen phase diagram

1-2. Value of Hydrogen as a Green Energy

Recently, to combat global climate change, achieving carbon neutrality by reducing fossil fuel consumption and reaching net-zero greenhouse gas emissions has emerged as a critical imperative of our time. Amidst this transition, the hydrogen economy is rapidly rising as a new paradigm to replace the traditional fossil fuel-centric economic system. The reasons why hydrogen is considered the core of future eco-friendly energy can be summarized into the following three points.

- **green energy without gas emission**: Fossil fuels such as coal and petroleum release substantial quantities of carbon dioxide and particulate matter during combustion. In contrast, when hydrogen generates electrical energy through a chemical reaction with oxygen (O_2) within a fuel cell, its sole emission is pure water (H_2O). In essence, it is a zero-carbon energy source that produces neither harmful pollutants nor the greenhouse gases responsible for global warming during the energy generation process.
- **Storage Storing Energy**: Renewable energy sources, such as solar and wind power, possess a critical limitation: their power generation is intermittent, fluctuating depending on weather conditions and seasons. Consequently, any electricity generated but not immediately consumed is ultimately wasted. However, by utilizing this surplus electricity to electrolyze water and produce green hydrogen, the electrical energy can be effectively converted and stored long-term in the form of hydrogen molecules, enabling long-distance transportation.
- **Broad Versatility**: There remain certain sectors that are challenging to address relying solely on electric vehicle (EV) batteries. For heavy-duty transportation that requires moving massive cargo over long

distances—such as heavy trucks, maritime vessels, and aircraft—hydrogen fuel cells present a far more advantageous solution due to the inherent weight constraints of current battery technologies. Furthermore, in heavy industries like steel and chemical manufacturing, which necessitate massive amounts of thermal energy, hydrogen is highly regarded as a viable alternative capable of replacing conventional fossil fuels.

2. Liquid Hydrogen

Liquid hydrogen refers to hydrogen that has been processed into a liquid state by cooling it to a cryogenic temperature of approximately -253℃.

2-1. Advantage of Liquid Hydrogen

Liquid hydrogen undergoes a volume reduction to approximately 1/800th of its gaseous state, which significantly enhances its stability and makes it highly advantageous for large-scale transportation. During the liquefaction process, both nitrogen and oxygen freeze into solid states before hydrogen does, allowing for the easy filtration and removal of these impurities. This inherent purification prevents potential damage to fuel cells during operation, an issue often caused by the residual nitrogen and oxygen present when hydrogen is initially produced. Furthermore, its liquid state allows for significantly faster refueling times compared to hydrogen gas.

2-2. Disdvantage of Liquid Hydrogen

Approximately 30% of hydrogen's inherent energy is consumed during the cooling and liquefaction process, resulting in suboptimal overall efficiency. Additionally, even in highly insulated storage tanks, minute heat exchange with the external environment causes a fraction of the liquid hydrogen to vaporize, leading to continuous losses known as boil-off gas (BOG). Furthermore, hydrogen embrittlement—a phenomenon where hydrogen atoms penetrate the interstitial spaces of conventional metals, thereby weakening their structural integrity—necessitates the use of expensive, specialized materials for storage and transport infrastructures.

3. Ammonia-Based Hydrogen Transport Technology

Ammonia, widely recognized primarily as a feedstock for chemical fertilizers, has recently emerged as the most viable alternative to propel the hydrogen economy forward. This is because it represents the core of hydrogen carrier technology, which involves chemically bonding difficult-to-handle gaseous hydrogen with another element to transform it into an easily transportable compound.

3-1. Characteristics of Ammonia: Nitrogen Bonded with Hydrogen

The chemical formula for ammonia is NH_3, a molecule formed by bonding one nitrogen atom with three hydrogen atoms. The primary reason ammonia is garnering widespread attention as a hydrogen carrier is its ease of liquefaction. While liquefying pure hydrogen necessitates an ultra-low temperature of -253℃, ammonia easily liquefies at approximately -33℃ or under a pressure of roughly 8 atmospheres (atm). Consequently, it boasts high stability and can significantly reduce transportation costs.

3-2. Ammonia Synthesis Process via the Haber-Bosch Process

Ammonia is synthesized by reacting hydrogen gas with nitrogen gas. This synthesis relies on the widely recognized Haber-Bosch process, developed in the early 20th century. It involves the chemical combination of

nitrogen and hydrogen under high-temperature and high-pressure conditions, facilitated by a specific catalyst.

$$N_2 + 3H_2 \rightarrow 2NH_3$$

Recently, research has been actively conducted on technologies utilizing ”green hydrogen,” which is produced by electrolyzing water using eco-friendly energy sources like solar and wind power, rather than being derived from fossil fuels.

Haber-Bosch Process

Developed by Fritz Haber and Carl Bosch in the early 20th century, this method synthesizes ammonia by reacting atmospheric nitrogen with hydrogen. It necessitates a high-temperature environment of 400–500℃ and high-pressure conditions exceeding 200 atm. The incorporation of a catalyst made the mass production of ammonia feasible. Historically, this groundbreaking technology also played a pivotal role in alleviating global food crises by enabling the mass production of ammonia for use in agricultural fertilizers.(Valera-Medina, 2018).

The Process of Extracting Hydrogen from Ammonia: Cracking

This is the process of recovering hydrogen from the transported ammonia for final utilization. Thermal energy is applied to the ammonia, and a catalyst is employed to break its chemical bonds. Following decomposition, the resulting nitrogen gas is safely vented into the atmosphere, while the purified hydrogen is isolated for application in fuel cells or power generation.(Smith, 2020)

3-3. Advantages of Ammonia-Based Hydrogen Transport

As previously noted, ammonia is the world’s most widely utilized fertilizer. Consequently, an extensive global network for its mass production and transportation is already well-established within the agricultural and chemical industries. In other words, existing infrastructure—including transport vessels, storage tanks, and pipelines—can be readily repurposed, enabling rapid commercialization. Furthermore, its volumetric hydrogen density is approximately 1.5 times greater than that of liquid hydrogen. Because the chemical compound itself contains no carbon, zero carbon dioxide is emitted throughout the entire lifecycle of this process.

3-4. Disadvantages of Ammonia-Based Hydrogen Transport

Due to the toxicity of ammonia, extreme caution is required during handling, and rigorous safety protocols must be established to prevent leakage and human exposure. Furthermore, the synthesis of ammonia from hydrogen and the subsequent cracking process to recover hydrogen both require significant energy input. This substantial energy consumption acts as a major drawback, as it leads to a reduction in the overall round-trip efficiency of the entire hydrogen supply chain.

4. Liquid Organic Hydrogen Carriers (LOHC)

As a method for the safe transportation of hydrogen, Liquid Organic Hydrogen Carriers (LOHC) are gaining significant attention as an alternative to toxic ammonia or liquefied hydrogen, which requires stringent cryogenic insulation. LOHCs are organic chemical compounds that possess physical and chemical properties similar to conventional petrochemicals.

Technically, LOHC acts as a medium that chemically bonds with hydrogen (storage) and subsequently releases

it when required. A primary advantage of LOHCs is their relative ease of handling, as they share comparable characteristics with widely used petroleum products, allowing for safer and more convenient management.

4-1. What Is LOHC Technology?

LOHC is a technology that uses organic compounds which remain as stable liquids at room temperature and atmospheric pressure to store and transport hydrogen. Hydrogen is loaded into a carrier compound at the production site, creating a hydrogen-rich liquid. This liquid is then shipped to its destination, where the hydrogen is extracted and used. The empty carrier is sent back to be refilled and reused.

4-2. How Hydrogen Is Stored and Released

The core science behind LOHC is a reversible chemical reaction which is a reaction that can run in both directions. A well known LOHC pair is toluene and methylcyclohexane (MCH). These two molecules cycle back and forth to store and release hydrogen. (Dong et al., 2021)

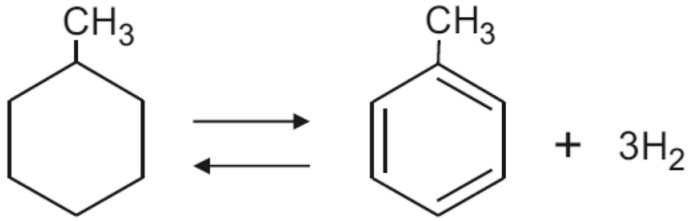


Figure 2. The hydrogenation and dehydrogenation of toluene to methylcyclohexane

4-2-1. Storing Hydrogen: Hydrogenation

In this step, hydrogen gas is reacted with toluene (an unsaturated compound, meaning it has room to bond with more hydrogen). Under the presence of a metal catalyst — such as platinum (Pt), palladium (Pd), or nickel (Ni) — toluene bonds with hydrogen to form methylcyclohexane, a saturated compound that is ”full” of hydrogen. This reaction releases heat, making it an exothermic reaction.

4-2-2. Releasing Hydrogen: Dehydrogenation

At the destination, the methylcyclohexane is heated to above 250℃ with a catalyst. The heat breaks the chemical bonds holding the hydrogen in place, and pure hydrogen gas is released. Because the reaction requires energy input to proceed, it is an endothermic reaction. The released hydrogen is used as fuel, while the toluene that remains is collected and shipped back to be reused.

4-3. Advantages of LOHC

Since LOHC exists as a liquid at room temperature and atmospheric pressure, just like water or oil, it is easy to store. It also has no toxicity unlike ammonia, making it highly safe. LOHC materials have properties similar to general petrochemical products. Therefore, oil tankers, tank trucks, and the pipelines and storage tanks at gas stations that were already used to carry oil can be used as they are without any structural modification, which can dramatically reduce initial infrastructure costs. Since the carrier material can be reused, it is also efficient.

4-4. Disadvantages of LOHC

Extracting hydrogen requires a high temperature of 250℃, which means additional energy must be consumed. The carrier compound holding the hydrogen is heavy, making it disadvantageous for long-distance transport. In particular, the hydrogen content ratio of the total liquid is as low as about 6%. In other words, transporting approximately 100 kg of liquid is required to obtain just 6 kg of hydrogen.

5. Real World Applications and Choosing the Right Strategy

Each of the three hydrogen transport technologies — liquefied hydrogen, ammonia-based transport, and LOHC — has its own strengths and limitations. The right choice depends on factors like transport distance, scale, and end-use requirements. Let's look at real-world examples to understand when and why each approach is used.

5-1. Liquefied Hydrogen Based Hydrogen Transport

A representative example of liquefied hydrogen-based hydrogen transport is Japan's HySTRA (Hydrogen Energy Supply-chain Technology Research Association) project. This project aims to demonstrate an international hydrogen supply chain by liquefying hydrogen produced in Australia and transporting it to Japan by sea. In the HySTRA project, hydrogen is produced using brown coal, then purified and liquefied at an extremely low temperature of approximately 253℃, and transported to Japan via a dedicated carrier vessel. The hydrogen is then offloaded and stored in Japan. There are three main reasons why the liquefied hydrogen method was adopted in this project. First, Kawasaki Heavy Industries, which led the project, had a strong technical foundation in cryogenic storage and transport, making it well-suited for building a liquefied hydrogen supply chain. Second, LOHC- and ammonia-based methods require the recovery and recirculation of the carrier medium after hydrogen extraction, which can result in additional logistical burden and energy loss. In contrast, liquefied hydrogen has the advantage of being transportable directly without such additional processes. Third, since Japan intended to apply the transported hydrogen to high-purity hydrogen applications such as fuel cells, liquefied hydrogen was chosen because it can be used immediately without any cracking or dehydrogenation process.(CO2-free Hydrogen Energy Supply-chain Technology Research Association (HySTRA), 2021)

5-2. Ammonia Based Hydrogen Transport

A representative example of ammonia-based hydrogen transport is the blue ammonia supply project between Saudi Arabia and Japan. In this method, hydrogen is converted into ammonia (NH_3), stored and transported in liquid form, and then decomposed back into hydrogen at the destination for use. Ammonia can be liquefied under relatively mild conditions (33℃ or low pressure), and the transport and storage infrastructure already established worldwide through the fertilizer industry can be utilized. Ammonia is also a form of chemically stored hydrogen, and when needed, hydrogen can be re-extracted through a catalytic reaction. Furthermore, in this industry, ammonia serves both as a direct combustion energy source and as a hydrogen carrier. Through this process, it was possible to reduce the energy used for cracking and achieve high efficiency.(가스신문, 2021)

5-3.LOHC Based Hydrogen Transport

A representative example of LOHC-based hydrogen transport is the SPERA Hydrogen project developed by Japan's Chiyoda Corporation. This project operates by reacting hydrogen with toluene to store it in the form of methylcyclohexane (MCH), transporting it in liquid form, and then extracting the hydrogen again at the destination. In an actual demonstration, hydrogen produced in Brunei was successfully transported to Japan

in LOHC form, and this is regarded as the world's first case of an international hydrogen supply chain based on LOHC. LOHC was used because the goal of this project was to transport hydrogen in liquid form at atmospheric pressure and room temperature, just like gasoline. It also had the advantage of being able to make direct use of the company's existing large crude oil storage tanks, chemical product carrier vessels, and pipelines. LOHC was chosen because it allows for long-term storage, and unlike ammonia, its low toxicity makes it advantageous for transport and storage in general residential areas. (Chiyoda Corporation, 2020)

Conclusions

Hydrogen, which is attracting attention as a next-generation eco-friendly energy source for overcoming the climate crisis, faces a significant technological barrier in storage and transport. To overcome this, liquefied hydrogen using cryogenic cooling, ammonia with its advantages of high hydrogen storage density and utilization of existing infrastructure, and LOHC (Liquid Organic Hydrogen Carrier) technology that can be handled safely like an ordinary liquid at room temperature are all being actively researched, each with their own distinct advantages and limitations. The technology used differs depending on the situation, and the most realistic approach to accelerating the future hydrogen economy is to combine multiple technologies appropriately according to transport distance, scale, and purpose of use. There also remain challenges that scientists must solve for the commercialization of these technologies. For the future hydrogen economy to have true eco-friendly value, carbon emissions must be completely controlled across the entire lifecycle of hydrogen, including production, storage, transport, and use. It is hoped that through this study, one can understand the infinite potential of hydrogen energy and the scientific and engineering challenges hidden beneath it, and reflect deeply on the role of chemical technology for a sustainable future.

Example Questions

You may research any topic other than the relevant question.

1. Which transport method is best for a given situation?

Design a hypothetical scenario with specific transport distance, scale, and end use, and analyze which of the three methods — liquefied hydrogen, ammonia, or LOHC — is most suitable.

2. How do ammonia cracking byproducts damage fuel cells?

After cracking, trace impurities remain. What chemical or electrochemical mechanisms cause degradation at the catalyst surface or electrolyte membrane?

3. Hydrogen can be categorized into green hydrogen, blue hydrogen, and gray hydrogen depending on the production method. How do the production methods and carbon emissions of each type of hydrogen differ? Also, what is the most sustainable hydrogen production method in the long term?

4. What other LOHC materials exist beyond toluene/MCH?

Research alternative carrier molecules and compare their hydrogenation and dehydrogenation conditions.

References

Chiyoda Corporation, 2020. The World's First Global Hydrogen Supply Chain Demonstration Project. https://www.chiyodacorp.com/en/service/lowcarbon/hydrogen/demonstration/.

CO2-free Hydrogen Energy Supply-chain Technology Research Association (HySTRA), 2021. Hydrogen Energy Supply Chain Pilot Project between Australia and Japan. https://www.hystra.or.jp/en/project/.

Dong, Z., A. Mukhtar, and H. Lin. 2021. Heterogeneous Catalysis on Liquid Organic Hydrogen Carriers. Topics in Catalysis **64**.

Smith, C. e. a. 2020. Ammonia as a transport vector for hydrogen. Energy Environmental Science .

The Engineering ToolBox, 2008. Hydrogen - Thermophysical Properties. https://www.engineeringtoolbox.com/hydrogen-d_1419.html.

Valera-Medina, A. e. a. 2018. Ammonia for power. Progress in Energy and Combustion Science .

가스신문. 2021. 암모니아 기반 수소 운송 관련 산업 동향. 가스신문 https://www.gasnews.com/news/articleView.html?idxno=94639.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



YoungDuk Lee

M. ydlee@kentech.ac.kr

Prof. Dr. Young Duk Lee received his Ph.D. in Energy Engineering from Technische Universität Berlin in 2015. He is currently an Associate Professor in the Department of Energy Engineering at the Korea Institute of Energy Technology (KENTECH). He serves as an Associate Editor of the Korean Hydrogen & New Energy Society and as a Scientific Committee Member of the ECOS international society. Before joining academia, he worked as a Principal Researcher at the Korea Institute of Machinery & Materials (KIMM) from 2005 to 2021. His research interests include fuel cells, electrolysis, hybrid energy systems, process simulation, exergy analysis, and life cycle assessment, with a particular focus on CO_2-free hydrogen production and fuel-cell-based hybrid power generation.

KENTECH
Korea Institute of Energy Technology

VIII

다성분계 고용체에서의 격자 왜곡이 전위 이동 억제 및 기계적 물성에 미치는 영향 :

고엔트로피 합금(HEA)의 강화 기전 분석

p. 168 - 199

Vol. 05 (2026)

다성분계 고용체에서의 격자 왜곡이 전위 이동 억제 및 기계적 물성에 미치는 영향 : 고엔트로피 합금(HEA)의 강화 기전 분석

Abstract

고엔트로피 합금(High-Entropy Alloy, HEA)은 5종 이상의 원소를 동등한 비율로 혼합하여 높은 혼합 엔트로피($\Delta S_{mix} \geq 1.5R$)를 통해 단일 고용체 상을 안정화하는 차세대 구조 재료로, 기존 합금 설계의 패러다임을 근본적으로 재정의한다. HEA는 극저온 액체 수소 저장 탱크, 항공기 터빈 블레이드, 원자력 구조재 등 극한 환경 응용 분야에서 기존 합금이 달성하기 어려운 강도–인성 복합 물성을 동시에 발휘한다. 2026년 현재, 수소 경제의 급속한 팽창과 우주·항공 산업의 기술 고도화에 따라 극한 환경 구조 소재에 대한 수요가 전례 없이 증가하고 있으며, HEA는 그 핵심 후보 소재로 부상하고 있다. 본 탐구자료는 HEA의 우수한 기계적 물성의 근원인 격자 왜곡(lattice distortion)이 소성 변형을 담당하는 전위(dislocation)의 이동을 억제하는 물리적 메커니즘을 원자 수준에서 분석하고 이해해본다.

Introduction

현대의 금속 소재 연구는 단순히 더 강하고 가벼운 재료를 찾는 것을 넘어, 원자 수준에서 재료의 구조를 설계하는 방향으로 진화하고 있다. 그 흐름 속에서 기존의 상식을 뒤집는 소재가 등장했다. 바로 고엔트로피 합금(HEA)이다. 전통적인 합금 설계는 철, 알루미늄, 구리처럼 하나의 주원소에 소량의 첨가원소를 넣는 방식이었다. 여러 원소를 다량으로 섞으면 복잡한 금속간 화합물이 형성되어 재료가 취약해진다는 것이 수십 년간의 통념이었다. 그런데 2004년 대만의 Yeh 교수 연구팀은 5종 이상의 원소를 거의 동등한 비율로 혼합했을 때 오히려 단순하고 안정적인 단일 결정 구조의 고용체가 만들어진다는 사실을 발표하며 재료과학계에 큰 충격을 주었다 (Yeh et al., 2004). 이 직관에 반하는 현상이 가능한 이유는 열역학의 핵심 개념인 엔트로피에 있으며, 이것이 바로 이 재료를 고엔트로피 합금이라 부르는 이유다.

HEA의 특성을 지배하는 핵심 원리는 네 가지 효과로 요약된다. 첫 번째는 고엔트로피 효과(High-entropy effect)다. 여러 원소를 혼합하면 혼합 엔트로피(ΔS_{mix})가 커지고, 깁스 자유에너지 관계식 $\Delta G = \Delta H - T\Delta S$에 따라 고온에서 자유에너지가 낮아져 단순한 고용체 상이 열역학적으로 안정해진다. 쉽게 말해, 복잡하게 섞일수록 오히려 구조가 단순해지는 역설이 열역학으로 설명된다.
두 번째는 격자 왜곡 효과(Severe lattice distortion effect)로, 크기와 성질이 다른 원소들이 하나의 결정격자를 무작위로 공유하면서 격자가 국소적으로 뒤틀리는 현상이다. 이것이 HEA가 강한 가장 직접적인 이유이며, 본 탐구자료에서 집중적으로 다루는 주제다.
세 번째는 지연 확산 효과(Sluggish diffusion effect)로, 다종 원소가 혼재하는 복잡한 에너지 지형이 원자의 이동을 느리게 만들어 고온에서도 미세구조가 잘 변하지 않고 크리프(creep) 저항성이 뛰어나다.
네 번째는 칵테일 효과(Cocktail effect)로, 여러 원소의 상호작용에서 비롯되는 시너지 물성을 뜻하며, 단순 혼합 법칙(rule of mixtures)으로는 예측할 수 없는 독특한 강도–인성 동시 향상이 이 효과에 기인한다.

이러한 특성 덕분에 HEA는 2026년 현재 두 가지 핵심 산업에서 특히 주목받고 있다. 수소 경제 분야에서는 액체 수소를 253℃에서 저장·운반하는 탱크와 배관 소재로 CrMnFeCoNi 계열 HEA가 유력하게 연구되고 있다. 기존 금속 재료들은 극저온에서 갑자기 부서지는 취성 파괴(brittle fracture)를 일으키는 반면, 이 합금은 196℃에서도 오히려 인성이 향상되는 독특한 성질을 보인다 (Gludovatz et al., 2014). 이는 액체

수소 저장 용기의 안전성과 직결되는 매우 중요한 특성이다. 우주·항공 산업에서는 항공기 엔진 터빈 블레이드가 1,200℃ 이상의 고온·고압·산화 환경에 수천 시간 노출되어야 한다. 내화성 HEA(Refractory HEA, RHEA)는 MoNbTaVW 계열을 중심으로 기존 초내열합금(superalloy)을 넘어서는 고온 강도를 보이며 차세대 블레이드 소재 후보로 각광받고 있다.

그러나 HEA 연구에는 아직 해결해야 할 과제가 있다. 가장 큰 어려움은 상태도 예측의 복잡성이다. 5종 이상의 원소를 포함하는 다성분계 상태도는 차원이 기하급수적으로 높아져, 어떤 조성이 단일 고용체를 이룰지 금속간 화합물이 생겨날지를 실험 없이 예측하기 매우 어렵다. 이를 해결하기 위해 열역학 계산 기법인 CALPHAD와 머신러닝 기반 조성 예측 모델이 활발하게 개발되고 있다. 또한 순도 높은 다종 원소를 균일하게 혼합하고 대형 부품으로 제조하는 공정은 기존 합금에 비해 비용이 높고 품질 편차를 제어하기 어렵다는 제조 비용과 공정 복잡성 문제도 상용화의 걸림돌로 남아 있다. 이러한 도전 과제들은 HEA의 강화 기전을 더 정밀하게 이해하고 조성-구조-물성 관계를 체계적으로 탐구해야 할 필요성을 역설적으로 보여준다.

1. 배경지식 : 결정 격자와 전위

1-1. 금속의 결정 구조

금속을 구성하는 원자들은 무작위로 흩어져 있는 것이 아니라 규칙적인 3차원 배열, 즉 결정 격자(crystal lattice)를 이룬다. 이 배열의 방식에 따라 면심입방(Face-Centered Cubic, FCC), 체심입방(Body-Centered Cubic, BCC), 조밀육방(Hexagonal Close-Packed, HCP) 구조 등으로 나뉜다. CrMnFeCoNi 계열 HEA는 주로 FCC 구조를 취하는데, 이는 각 단위격자의 모서리 8개와 각 면의 중심 6개에 원자가 위치하는 형태다. FCC 구조가 중요한 이유는 소성 변형이 일어나는 슬립면(slip plane)의 수가 12개로 많아 변형이 여러 방향으로 분산될 수 있기 때문이다. 이것이 CrMnFeCoNi 합금이 극저온에서도 높은 연성을 유지하는 결정 구조적 배경이 된다.

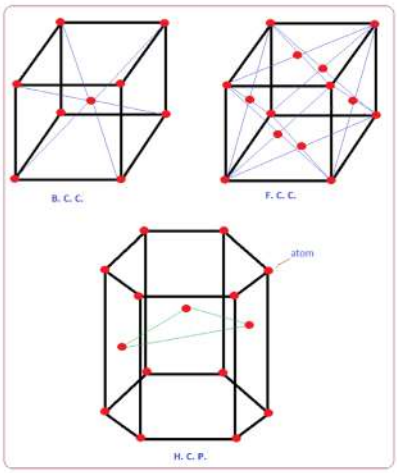


Figure 1. BCC, FCC, HCP 결정 구조의 단위격자 모식도. CrMnFeCoNi HEA는 FCC 구조를 가지며, FCC는 111면을 따라 12개의 독립적인 슬립계(slip system)를 보유하여 BCC(48개이나 임계 전단 응력이 높음) 및 HCP(3개)에 비해 상온에서 소성 변형이 용이함.

1-2. 소성 변형의 주역 : 전위(Dislocation)

금속에 외부에서 힘을 가하면 처음에는 원자 간 거리만 미세하게 변하고, 힘을 제거하면 원래 상태로 돌아오는 탄성 변형(elastic deformation)이 일어난다. 그러나 힘이 임계값인 항복 강도(yield strength)를

초과하면 재료는 영구적으로 모양이 바뀌는 소성 변형(plastic deformation)에 돌입한다. 소성 변형을 단순하게 생각하면 한 층의 원자 전체가 동시에 미끄러지는 것처럼 보인다. 그런데 이 과정에 필요한 이론적 전단 강도를 계산하면 실제로 측정되는 값보다 약 100배에서 1,000배나 크다. 이 거대한 모순을 해결하는 것이 바로 전위(dislocation)라는 선형 결함이다.

전위는 결정 격자 내부에 존재하는 원자 배열의 불완전한 경계선으로, 소성 변형은 격자 전체가 한 번에 이동하는 것이 아니라 전위가 격자 속을 한 원자 간격씩 차례로 이동함으로써 진행된다. 이 과정은 긴 카펫을 통째로 끌지 않고 주름을 한 발씩 앞으로 밀어 이동시키는 것과 본질적으로 동일하다. 훨씬 적은 에너지로 소성 변형이 가능한 이유가 바로 여기에 있다.

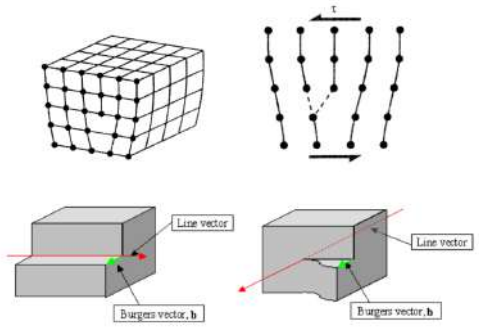


Figure 2. Figure 2. 전위의 형성과 이동 메커니즘 모식도.

(위 왼쪽) 칼날 전위의 구조. (위 오른쪽) 전단 응력 τ 하에서 전위가 한 원자 간격씩 순차적으로 이동하며 소성 변형이 진행된다. (아래) 칼날 전위(왼쪽)에서는 버거스 벡터 b가 전위선에 수직이고, 나사 전위(오른쪽)에서는 평행하다. $|b|$는 전위 이동 시 원자가 실제로 이동하는 거리에 해당한다.

전위의 특성은 버거스 벡터(Burgers vector, b)로 정량화된다. b는 전위 주변의 원자 배열이 완전한 결정 격자로부터 얼마나 어긋나 있는지를 나타내는 벡터로, 그 크기 $|b|$는 전위가 한 번 이동할 때 원자가 실제로 이동하는 거리에 해당한다. 칼날 전위(edge dislocation)는 b와 전위선(dislocation line)이 수직 방향이고, 나사 전위(screw dislocation)는 평행한 방향이다. 전위가 이동하는 데 필요한 임계 분해 전단 응력(Critical Resolved Shear Stress, CRSS)은 $\tau_{CRSS} \propto \mu|b|/L$로 근사된다. 여기서 μ는 전단 탄성계수, L은 전위 간 평균 간격이다. 이동 경로 상에 장애물이 많아질수록 L이 줄어들고 CRSS가 높아져 재료가 강해진다. HEA의 격자 왜곡은 바로 이 장애물 역할을 한다는 것이 이 탐구자료의 핵심 주장이다.

2. 이론과 메커니즘

2-1. HEA에서 격자 왜곡이 발생하는 원인

순수 금속이나 기존 이원계 합금에서는 격자를 구성하는 원자의 종류가 한두 가지이므로 격자 간격이 거의 균일하다. 그러나 HEA에서는 원자 반지름이 서로 다른 5종 이상의 원소가 결정 격자의 자리(lattice site)를 무작위로 점유한다. 대표적인 HEA인 CrMnFeCoNi 합금을 예로 들면, Cr(128 pm), Mn(127 pm), Fe(126 pm), Co(125 pm), Ni(124 pm)의 원자 반지름 차이가 격자 전반에 걸쳐 국소적인 크기 불일치(size mismatch)를 야기한다. 이와 함께 각 원소의 전기음성도 차이(Δχ)도 원자 간 결합력의 방향성을 변화시켜 격자를 추가로 왜곡시킨다. Cr과 Ni의 전기음성도는 각각 1.66과 1.91로 0.25나 차이가 나는데, 이 차이가 격자 내에서 전자 밀도의 불균형을 만들어 원자 간 결합 거리를 변화시킨다.

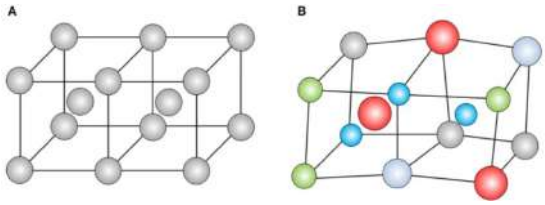


Figure 3. 순수 금속과 HEA의 결정격자 비교 모식도.

(A) 순수 금속에서는 동일한 크기의 원자들이 격자 자리를 균일하게 점유하여 격자가 규칙적이다. (B) HEA에서는 원자 반지름이 서로 다른 5종 이상의 원소(색깔로 구분)가 격자 자리를 무작위로 점유하여 격자가 전반적으로 왜곡된다. 이 왜곡이 내부 응력장을 형성하여 전위 이동을 억제하는 강화 기전의 출발점이 된다. 격자 왜곡의 정도는 원자 크기 불일치 파라미터 δ로 정량화된다. 이 값은 원자 반지름의 가중 평균에 대한 표준편차 개념으로 정의되며, 다음과 같이 계산한다.

$$\delta = \sqrt{\sum x_i \left(1 - \frac{r_i}{\bar{r}}\right)^2} \times 100\,(\%)$$

여기서 x_i는 원소 i의 몰 분율, r_i는 그 원자 반지름, $\bar{r} = \sum x_i r_i$는 전체 평균 원자 반지름이다. δ가 클수록 격자 왜곡이 심하다는 뜻으로, 일반적으로 $\delta \geq 1\%$이면 고용 강화 효과가 유의미하게 나타난다. Zhang 등(2008)은 $\delta < 6.6\%$인 경우 단순 고용체가 안정적으로 형성되고, 이를 초과하면 금속간 화합물이 생성되는 경향이 있음을 보였다. 즉, HEA가 강해지려면 격자 왜곡이 충분히 커야 하지만 동시에 고용체 구조를 유지하기 위해 일정 한계 이하여야 한다는 최적화 조건이 존재한다.

2-2. 격자 왜곡이 형성하는 내부 응력장

크기가 다른 원소들이 격자에 무작위로 들어앉으면 단순히 격자 간격이 늘어나거나 줄어드는 것에서 그치지 않는다. 탄성역학적 관점에서, 주변 원자보다 반지름이 큰 원자가 격자 자리를 차지하면 주변 격자를 바깥쪽으로 밀어내는 압축 응력장(compressive stress field)이 그 주변에 형성된다. 반대로 반지름이 작은 원자 주변에는 격자가 안쪽으로 당겨지는 인장 응력장(tensile stress field)이 형성된다. 순수 금속에서는 모든 원자의 크기가 같으므로 이러한 국소 응력장이 존재하지 않고 결정 전체의 퍼텐셜 에너지 지형이 균일한 주기성을 갖는다. 그러나 HEA에서는 5종 이상의 원소가 무작위로 배치되어 있으므로, 이 국소 응력장들이 결정 전체에 걸쳐 불규칙하고 불균일하게 분포한다. 그 결과 결정의 퍼텐셜 에너지 지형은 고요한 호수 표면이 아니라 크고 작은 언덕과 골짜기가 무질서하게 펼쳐진 험한 산악 지형과 같은 형태를 띠게 된다.

2-3. 격자 왜곡이 전위 이동을 억제하는 원리

전위가 이동하려면 결정 격자가 만드는 퍼텐셜 에너지의 언덕인 Peierls–Nabarro 장벽을 넘어야 한다. 순수 금속에서 이 장벽은 낮고 주기적으로 반복되므로 전위는 비교적 쉽게 이동할 수 있다. 그러나 HEA의 불균일한 에너지 지형에서는 두 가지 방식으로 전위 이동이 억제된다.

전위는 자체적으로도 응력장을 갖는다. 칼날 전위를 예로 들면 전위선 위쪽의 격자는 압축되어 있고 아래쪽은 인장되어 있다. 이 전위 고유의 응력장이 격자 내에 이미 존재하는 국소 응력장과 만나면 탄성 상호작용(elastic interaction)이 일어나며, 이를 표현하는 것이 Peach–Koehler 식이다. 전위에 작용하는 힘 F는 $F = (\sigma \cdot b) \times \xi$로 주어지는데, 여기서 σ는 주변의 응력 텐서, b는 버거스 벡터, ξ는 전위선 방향 단위벡터다. HEA에서는 전위가 이동하는 경로의 매 지점마다 σ가 불규칙하게 변하므로, 전위에 작용하는 힘의 크기와 방향이 끊임없이 바뀐다.

이로 인해 두 가지 억제 현상이 나타난다. 첫 번째는 마찰 응력(friction stress, τ_f)의 전반적 증가다. 전위 이동 경로 전체에 걸쳐 에너지 장벽의 평균 높이가 높아지므로, 전위를 움직이는 데 필요한 최소 응력이 증가한다. 이것이 항복 강도 향상의 가장 직접적인 원인이다. 두 번째는 전위 핀닝(dislocation pinning)이다. 전위가 이동하다가 주변보다 에너지가 특히 낮은, 즉 열역학적으로 안정한 격자 자리에 도달하면 그 자리에 고착되는 핀닝 현상이 발생한다. HEA에서는 이러한 핀닝 자리가 격자 전체에 걸쳐 높은 밀도로 분포하므로, 전위는 아주 짧은 거리를 이동한 뒤 곧바로 핀닝된다. 핀닝된 전위를 다시 떼어내 이동시키려면 추가적인 응력이 필요하므로, 재료의 항복 강도와 가공 경화율이 모두 높아진다.

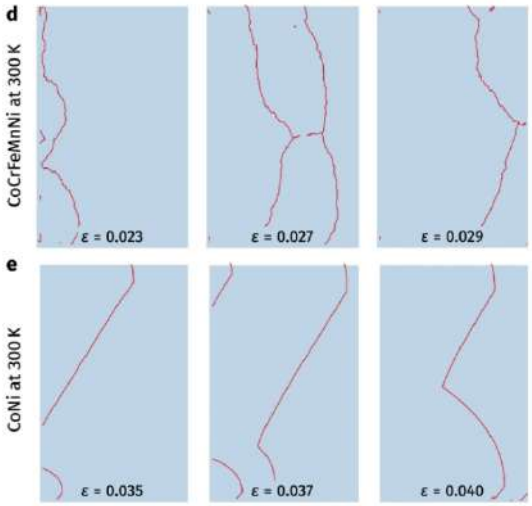


Figure 4. HEA와 기존 합금에서의 전위 이동 경로 비교 (원자 시뮬레이션, 300K)

(위, d) CoCrFeMnNi HEA에서 전위선은 불규칙한 Peierls 장벽에 반복적으로 핀닝되어 파도 모양의 경로를 보인다. (아래, e) CoNi 이원계 합금에서는 전위선이 거의 직선을 유지하며 이동한다. 이 차이가 HEA의 높은 항복 강도의 직접적 원인이다. ε은 인가된 변형률을 나타낸다.

2-4. 격자 왜곡이 전위 이동을 억제하는 원리

기존의 고용 강화 이론은 소량의 용질 원자가 만드는 응력장이 전위 이동을 방해한다는 개념에 기반한다. HEA에서는 이 개념이 극단화되어 모든 원소가 동등하게 용질이자 용매가 된다. Varvenne 등 (Varvenne et al., 2016)은 이를 다성분계 HEA에 맞게 확장하였으며, 고용 강화에 의한 항복 강도 증가분 Δτ는 다음과 같이 근사된다.

$$\Delta\tau \propto \mu \cdot |b| \cdot \sqrt{\sum x_i \cdot \epsilon_i^2}$$

여기서 μ는 전단 탄성계수, $|b|$는 버거스 벡터의 크기, x_i는 원소 i의 몰 분율, ϵ_i는 원소 i가 만드는 국소 탄성 응력장의 세기 파라미터다. 핵심은 $\sum x_i \epsilon_i^2$ 항이다. 기존 합금에서는 소량 첨가된 원소만 이 합산에 기여하므로 값이 작다. 그러나 HEA에서는 5종 이상의 모든 원소가 각자의 ϵ_i 값을 갖고 전체 합산에 동등하게 기여하므로, 이 항이 기존 합금과 비교할 수 없이 커진다. 이것이 바로 HEA가 기존 고용 강화 합금보다 훨씬 높은 강도를 보이는 이유를 정량적으로 설명한다.

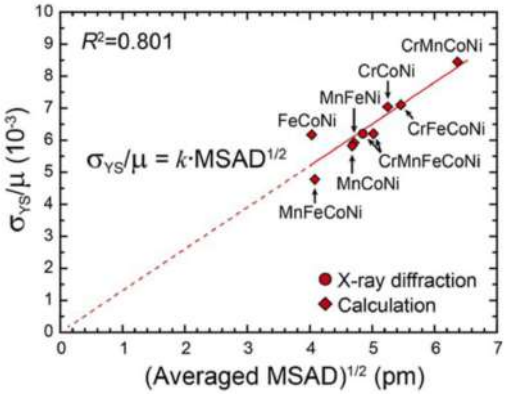


Figure 5. 원소 조합에 따른 고용 강화량 비교

CrMnFeCoNi HEA 및 그 파생 삼원계·사원계 합금의 0K 항복 강도(y축)를 격자 왜곡 정도(MSAD의 제곱근, x축)에 대해 나타낸 그래프. 구성 원소 수가 많아질수록 격자 왜곡이 커지고 항복 강도가 비례하여 증가한다. 이는 $\Delta\tau \propto \sqrt{\sum x_i \epsilon_i^2}$로 표현되는 Varvenne 강화 이론과 일치한다.

3. 방법론: Processing Routes for High-Entropy Alloys

HEA의 합성 공정은 단순히 시편을 제작하는 과정을 넘어, 조성 균일성, 결정립 미세화, 결함 밀도 제어를 통해 강화 기전의 발현 여부를 결정하는 핵심 변수이다. 이는 곧 다른 말로 HEA 내부의 에너지 상태를 결정하는 과정으로 이해할 수 있다. 각 공정은 서로 다른 에너지 입력 방식: 열적 에너지(arc melting), 기계적 에너지(mechanical alloying), 국소적 광열 에너지(additive manufacturing)을 사용하며, 이로 인해 형성되는 미세구조의 자유에너지 상태가 달라진다.

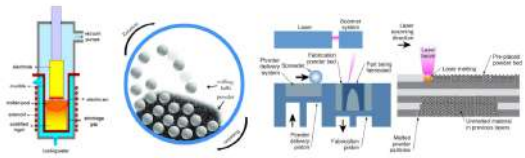


Figure 6. 순서대로 VAM, Mechanical Alloying, Additive Manufacturing의 공정을 시각화 한 것이다.

특히 급속 응고가 일어나는 적층 제조에서는 비평형 상태가 고정되면서, 열역학적으로 완전히 안정하지 않지만 기계적으로 유리한 구조가 유지될 수 있다. 이러한 비평형 구조는 높은 전위 밀도와 미세한 결정립을 포함하며, 이는 전위 이동을 더욱 효과적으로 억제한다. 특히, 이러한 공정 간 차이는 단순한 미세구조 차이를 넘어 전위의 생성, 이동 경로, 상호작용 방식 자체를 변화시키는 요인으로 작용한다. 즉, HEA의 강화 기전은 조성만으로 결정되는 것이 아니라, 공정에 의해 형성된 에너지 지형(energy landscape)에 의해 재구성된다고 볼 수 있다. 또한 각 공정은 시스템의 열역학적 평형 접근 경로(path to equilibrium)를 다르게 설정한다. 용융 기반 공정은 평형 상태로의 접근을 허용하는 반면, 고상 합성이나 적층 제조는 비평형 상태를 고정시키는 방향으로 작용한다. 이러한 차이는 최종적으로 전위 이동에 대한 에너지 장벽 분포를 결정하는 핵심 변수이다.

3-1. Vacuum Arc Melting (VAM)

아크 용해는 HEA 연구의 표준적인 용융 기반 공정으로, 고순도 금속 원소를 고진공 또는 불활성 기체 분위기에서 용융시킨다. 다성분계 합금의 특성상 원소 간 융점 및 확산 속도 차이로 인해 냉각 과정에서 수지상 구조(Dendritic structure)와 화학적 편석(Segregation)이 발생하기 쉽다.

아크 용해 공정에서 형성되는 수지상 구조는 단순한 미세조직이 아니라, 국소적인 조성 차이에 따른 격자 왜곡 분포의 공간적 불균일성을 의미한다. 수지상 중심과 interdendritic 영역 사이의 조성 차이는 원자 반경 불일치 파라미터 δ의 국소적 변화를 유도하며, 이는 전위 이동 시 서로 다른 에너지 장벽을 형성하게 된다. 이러한 구조에서는 전위가 이동할 때 균일한 마찰 응력(friction stress)을 경험하는 것이 아니라, 위치에 따라 다른 저항을 받게 된다. 결과적으로 전위는 특정 영역에서 쉽게 이동하다가, 다른 영역에서는 강하게 핀닝(pinning)되는 비균일 변형 거동을 보이게 된다.

이를 해결하기 위해 최소 5회 이상의 반복 용해(Flip-and-Remelt)와 장시간의 균질화 열처리가 필수적이다. 기초 연구용으로는 유용하나, 에너지 산업용 대형 구조물이나 복잡 형상 부품 제조 시에는 국부적 상분리 제어에 한계가 있다.

또한 장시간 균질화 열처리는 조성 균일성을 향상시키는 동시에, 결정립 성장(grain growth)을 유도한다. 이는 Hall–Petch 관계:

$$\sigma_y = \sigma_0 + kd^{-1/2}$$

에서 결정립 크기 d의 증가에 따라 항복 강도가 감소할 수 있음을 의미한다 (Cantor et al., 2004). 따라서 아크 용해 공정에서는 조성 균일성 확보와 결정립 미세화 사이의 trade-off를 고려한 공정 최적화가 필수적이다. 에너지 관점에서, 이 공정은 높은 열에너지 입력을 필요로 하며, 대형 구조물 제작 시 에너지 효율성이 제한될 수 있다. 따라서 산업적 응용에서는 공정 효율과 미세구조 제어 능력 간의 균형이 중요한 설계 변수로 작용한다.

3-2. Mechanical Alloying (MA) Solid-state Synthesis

기계적 합금화는 용융 과정을 거치지 않고 금속 분말을 고에너지 볼밀링하여 고체 상태에서 원자 단위의 혼합을 유도한다. 이 공정은 융점 차이가 극심한 원소들 간의 합금화에 유리하며, 반복적인 소성 변형을 통해 나노 결정립(Nanocrystalline) 조직을 형성함으로써 항복 강도를 극대화한다.

다만, 밀링 매체로부터의 불순물 혼입과 산화 제어가 핵심 과제이며, 에너지 구조재로의 적용을 위해서는 방전 플라즈마 소결(SPS) 등을 통한 후속 치밀화 과정에서 기공을 완벽히 제거해야 한다. 기계적 합금화 과정에서 반복적인 소성 변형은 단순히 분말 혼합을 넘어, 결정 구조 내부에 고밀도의 전위(dislocation density, ρ)를 축적시키는 과정으로 이해할 수 있다. 이때 전위 밀도는 $\sigma \propto \alpha\mu b\sqrt{\rho}$와 같은 관계로 기계적 강도에 기여한다. 즉, 밀링 과정에서 축적된 전위는 이후 소결된 벌크 재료에서도 잔존할 수 있으며, 이는 초기 항복 강도의 증가로 이어진다 (Yeh et al., 2004).

또한 MA 공정에서 형성되는 나노 결정립 구조는 단순한 크기 감소 효과를 넘어, 결정립 경계(grain boundary)가 전위 이동의 주요 장애물로 작용하는 구조를 형성한다. 특히 HEA에서는 이러한 경계가 조성적으로도 불균일할 수 있어, 전위가 경계를 통과할 때 추가적인 에너지 장벽이 발생한다. 비평형 상태에서 형성된 구조는 열처리 과정에서 부분적으로 안정화되지만, 완전히 평형 상태로 이동하지 않고 kinetic trapping 상태로 유지될 수 있다. 이 상태는 열역학적으로는 불안정하지만, 기계적으로는 높은 강도를 제공하는 특징을 가진다 (Yeh et al., 2004).

에너지 시스템 적용 측면에서 MA 공정은 고융점 원소 기반 HEA, 특히 내화성 HEA(Refractory HEA)의 합성에 매우 유리하다. 다만 분말 기반 공정 특성상 기공(porosity) 제거가 중요하며, 이는 최종 구조재의 피로 수명과 수소 취성 저항성에 직접적인 영향을 미친다.

3-3. Additive Manufacturing (AM)

적층 제조(L-PBF 등)는 2026년 현재 HEA 산업화의 핵심 공정이다. 레이저의 국소적 입열을 통한 $10^3 - 10^6$ K/s 수준의 급속 응고는 원자의 장범위 확산을 억제하여 용질 트래핑(Solute trapping) 효과를 유발하며 이는 거시적 편석이 억제된 준안정 고용체 형성을 가능케 한다. 적층 제조에서의 급속 응고는 단순한 미세화 효과를 넘어, 확산 지배(diffusion-controlled)에서 계면 지배(interface-controlled) 성장 메커니즘으로의 전환을 유도한다. 이로 인해 전통적인 평형 응고 과정에서는 형성되기 어려운 준안정 고용체와 비평형 미세구조가 형성된다.

특히 solute trapping 현상은 고속 응고 과정에서 고체–액체 계면이 이동하는 속도가 원자의 확산 속도를 초과할 때 발생하며, 이로 인해 용질 원소가 평형 농도보다 높은 상태로 고용된다. 이는 결과적으로 격자 왜곡의 공간적 균일성 증가로 이어져 전위 이동 억제 효과를 강화한다 (Zhang et al., 2014). 또한 AM 특유의 미세한 세포 구조(Cellular structure) 경계에서 발생하는 나노 스케일의 화학적 변조(Chemical modulation)는 전위의 이동을 다단계로 차단하는 계층적 구조를 형성하기 때문에 수소 저장 장치나 항공우주용 경량 구조물의 내구성 향상에 특히 기여한다. AM에서 형성되는 세포 구조(cellular structure)는 μm 이하의 크기를 가지며, 그 경계에는 원소 농도 차이에 의한 미세한 화학적 변조가 존재한다. 이 구조는 다음과 같은 다단계 전위 저항 메커니즘을 형성한다. 세포 내부에선 격자 왜곡에 의한 고용 강화가 발생하며, 세포 경계에서는 화학적 조성 변화에 의한 전위 산란, 그리고 미세 결정립에선 결정립 경계에 의한 전위 차단이 발생한다. 이러한 구조는 전위를 단일 메커니즘이 아닌, 계층적(hierarchical)으로 억제하는 구조를 형성한다. 더 나아가, AM 공정에서 형성되는 높은 열 구배는 방향성 응고(directionally solidified microstructure)를 유도할 수 있으며, 이는 특정 방향에서의 기계적 성질을 강화하는 이방성(anisotropy)을 유발한다 (Zhang et al., 2014).

이러한 특성은 항공우주 구조물과 같이 하중 방향이 명확한 시스템에서 오히려 장점으로 활용될 수 있다. 에너지 산업 관점에서 보았을 때, AM 기반 HEA는 매우 중요한 의미를 가진다. 수소 저장 시스템에서는 극저온 환경에서의 균열 억제가 중요하며, AM으로 형성된 미세구조는 균열 전파 경로를 복잡하게 만들어 파괴 저항성을 증가시킬 수 있다. 또한 항공우주 분야에서는 경량화와 고강도가 동시에 요구되는데, AM은 내부 격자 구조(lattice structure)를 설계하여 강도 대비 무게 비율을 극대화할 수 있다.

4. Discussion: Strengthening Mechanisms and Energy-Relevant Design Strategies

“HEA의 강화 기전을 에너지 관점에서 보면, 외부에서 가해진 에너지가 어떻게 분배되는가?” 라는 질문이 주가 된다. 금속이 변형될 때 외부에서 가해진 일(work)은 다음 세 가지로 분배된다. 전위 이동에 사용되는 에너지, 열로 소산되는 에너지, 그리고 미세구조 변화(상변화, 결함 생성)에 사용되는 에너지이다. HEA에서는 격자 왜곡과 국소 응력장이 존재하기 때문에, 전위 이동에 필요한 에너지 비율이 증가한다. 이는 같은 변형을 일으키기 위해 더 많은 외부 에너지가 필요하다는 것을 의미하며, 결과적으로 높은 강도로 나타난다.

4-1. Lattice Distortion and Dislocation Mobility

HEA의 강화 기전은 다성분계 원자 간 반경 및 탄성계수 차이에서 기인하는 격자 왜곡(Lattice Distortion)에 기반한다. 이러한 국소적 응력장은 전위(Dislocation) 이동에 대한 마찰 응력(Friction stress)으로 작용하여 항복 강도를 높인다. 하지만 모든 HEA에서 격자 왜곡이 심각하다고 일반화하는 것은 위험하며, 원자 반경 편차가 큰 원소를 포함한 합금에서 선별적으로 극대화된다. 최근에는 격자 왜곡뿐만 아니라 화학적 단범위 질서(Short-range order, SRO)가 전위 선의 에너지를 변화시켜 추가적인 강화 효과를 제공한다는 점이 규명되어, 이를 정밀 제어하는 설계가 중요해지고 있다 (Miracle and Senkov, 2017), (George et al., 2019). 격자 왜곡에 의해 형성된 국소 응력장은 전위 이동에 대한 단순한 저항을 넘어서, 전위의 이동 경로 자체를 무작위화(randomization)하는 효과를 가진다. 즉, 전위는 이상적인 슬립면을 따라 직선적으

로 이동하지 못하고, 에너지적으로 유리한 경로를 탐색하며 비선형적인 이동 경로를 따른다. 이 과정은 Peierls–Nabarro 장벽의 공간적 분포를 불균일하게 만들며, 전위 이동에 필요한 임계 응력(critical stress)을 증가시킨다. 특히 HEA에서는 다음 두 가지 강화 메커니즘인 평균 마찰 응력 증가 (global strengthening), 국소 핀닝에 의한 이동 억제 (local pinning)이 동시에 작용한다. 또한 최근 연구에서는 화학적 단범위 질서(SRO)가 전위 코어(dislocation core)의 에너지 상태를 변화시켜, 전위의 핵 생성 및 이동을 추가적으로 억제하는 것으로 보고되고 있다. 이는 기존의 단순 고용 강화 모델을 넘어, 화학적 질서–기계적 물성 간 상호작용을 고려한 새로운 설계 방향을 제시한다 (George et al., 2019). 에너지 관점에서 보면, 이러한 구조는 전위 이동에 필요한 활성화 에너지 $\Delta G_{activation}$를 증가시키며, 이는 외부에서 가해진 기계적 에너지가 더 높은 임계값을 넘어야 변형이 시작됨을 의미한다.

4-2. Metastability Engineering: TRIP and TWIP Effects

강도와 연성의 트레이드-오프(Trade-off) 관계를 극복하기 위해 의도적으로 상 안정성을 낮추는 준안정성 공학이 활용된다. 변형 중 면심입방구조(FCC)가 마르텐사이트 구조로 변하는 TRIP 효과와 결정 쌍정을 형성하는 TWIP 효과는 가공 경화율을 지속적으로 유지시켜 균열의 진전을 효과적으로 지연시킨다. 이러한 손상 허용(Damage-tolerant) 특성은 압력 변화가 극심한 극저온 수소 탱크나 고속 진동이 발생하는 우주 항공 부품의 신뢰성을 보장하는 핵심 기전이다. TRIP 및 TWIP 효과는 단순한 변형 메커니즘이 아니라, 기계적 에너지의 내부 변환 경로를 확장하는 메커니즘으로 이해할 수 있다.

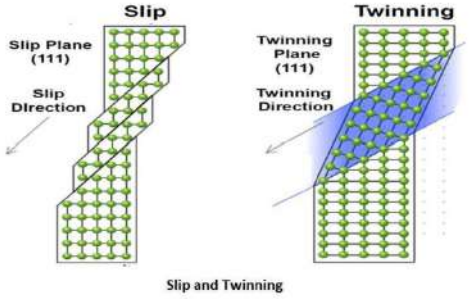


Figure 7. Slip에서는 원자층이 그대로 미끄러지지만, Twinning에서는 일부 영역의 구조 방향이 바뀌어 새로운 경계가 형성됨

외부에서 가해진 일(work)은 다음과 같은 형태로 변환된다:

$$W_{\text{external}} = W_{\text{dislocation}} + W_{\text{heat}} + W_{\text{phase}}$$

여기서 W_{phase}는 상변화에 사용되는 에너지이다. TRIP 효과가 활성화된 HEA에서는 이 항의 비중이 증가하여, 전위 이동에 집중되는 에너지를 분산시킨다.

또한 TWIP 효과는 결정 내 쌍정(twin boundary)을 형성하여, 전위 이동 경로를 반복적으로 차단하는 동적 장애물(dynamic obstacle)을 생성한다. 이로 인해 전위 밀도는 지속적으로 증가하며, 가공 경화율이 높은 수준으로 유지된다. 이러한 메커니즘은 특히 극저온 환경에서 더욱 두드러지는데, 온도가 낮아질수록 전위 이동이 제한되면서 상변화 및 쌍정 형성의 기여도가 상대적으로 증가하기 때문이다(Li et al., 2016).

에너지 산업 측면에서, 이러한 특성은 단순한 강도 향상을 넘어 손상 허용 설계(damage-tolerant design)를 가능하게 한다. 즉, 재료가 파괴되기 전에 더 많은 에너지를 흡수하고 분산할 수 있기 때문에, 수소 저장 탱크나 항공 구조물과 같이 안전성이 중요한 시스템에서 매우 중요한 역할을 한다.

4-3. Sluggish Diffusion and Thermal Stability

초기 가설인 확산 지연(Sluggish diffusion) 효과는 고온 환경에서의 결정립 성장 및 석출상 조대화를 억제하는 열적 안정성의 근거로 제시되었다. 그러나 최신 연구들에 따르면 확산 속도는 합금계와 온도에 따라 일반 합금과 유사하거나 오히려 빠른 경우도 관찰되므로, 이를 보편적 특징으로 단정해서는 안 된다. 따라서 고온 에너지 시스템 적용을 위해서는 해당 온도 영역에서의 개별적 확산 거동과 내산화성에 대한 실험적 검증이 반드시 선행되어야 한다.

확산 거동은 HEA의 고온 안정성을 결정하는 핵심 요소 중 하나이다. 일반적으로 확산 계수는 다음과 같이 표현된다.

$$D = D_0 \exp\left(-\frac{Q}{RT}\right)$$

여기서 Q는 확산 활성화 에너지이다. HEA에서는 다성분계 상호작용으로 인해 Q가 증가할 가능성이 있으며, 이는 확산 속도를 감소시키는 방향으로 작용할 수 있다 (Miracle and Senkov, 2017).

그러나 최근 연구들은 HEA의 확산 거동이 단순히 “느리다”라고 일반화될 수 없음을 보여준다. 일부 합금에서는 기존 합금과 유사하거나, 특정 원소에 대해 더 빠른 확산이 관찰되기도 한다. 이는 HEA 내부의 국소 화학 환경(local chemical environment)이 매우 다양하기 때문이다 (Miracle and Senkov, 2017),(Zhang et al., 2014).

따라서 HEA의 열적 안정성은 단순히 확산 속도 하나로 설명될 수 없으며, 격자 왜곡에 의한 에너지 장벽, 다성분계 결합 에너지 분포, 산화 및 부식 반응성이 복합적으로 작용한다. 특히 고온 에너지 시스템(터빈, 원자로 등)에서는 확산, 산화, 크리프 거동을 통합적으로 고려해야 하며, 이는 HEA 설계의 중요한 연구 방향으로 남아 있다.

4-4. Non-equiatomic Design and Economic Viability

전통적인 등원자비(Equiatomic) 설계에서 벗어나 특정 원소의 비율을 조절하는 비등원자비 설계는 성능 최적화와 비용 절감을 동시에 달성하는 전략이다. Co, Ni 등 고가 원소의 비중을 낮추면서도 엔트로피 안정화 효과를 유지하는 Lean-HEA 설계는 자원 공급망 리스크가 큰 에너지 산업에서 HEA의 상용화 가능성을 높이는 실질적인 방안이다.

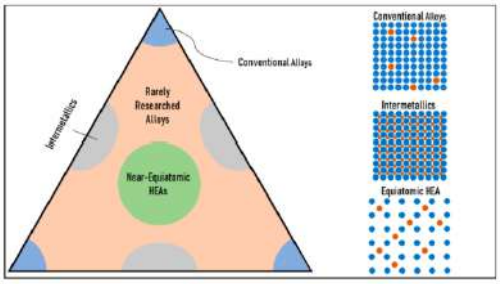


Figure 8. Conventional alloys, intermetallics, and equiatomic HEAs의 원자 배열 비교를 통해, HEA가 다성분 무작위 고용체 구조를 형성함을 보여

비등원자비 설계는 단순한 비용 절감 전략을 넘어, 물성–열역학–공정성 간 균형을 최적화하는 다변수 설계 문제로 이해할 수 있다. 등원자비 HEA는 최대 혼합 엔트로피를 제공하지만, 이는 항상 최적의 기계

적 물성을 보장하지는 않는다. 특정 원소의 비율을 조절하면 격자 왜곡 정도 조절, 상 안정성 변화 (FCC ↔ BCC), 그리고 TRIP 활성화 조건 조절 효과를 유도할 수 있다.
예를 들어, Mn 함량을 증가시키면 FCC 안정성이 감소하여 TRIP 효과가 활성화될 수 있으며, 이는 연성 향상으로 이어진다 (George et al., 2019). 또한 Lean-HEA 설계는 Co, Ni와 같은 전략 금속의 사용량을 줄여 공급망 리스크를 완화할 수 있으며, 이는 에너지 산업에서 매우 중요한 요소이다. 특히 수소 경제에서는 대량 생산이 필수적이기 때문에, 비용과 성능을 동시에 고려한 설계가 요구된다.

4-5. CALPHAD and AI-Driven Discovery

조성 공간의 방대함으로 인한 탐색 비용 문제는 CALPHAD(상태도 계산)와 머신러닝의 결합으로 해결되고 있다. CALPHAD는 열역학적 상 안정성을 예측하며, AI는 방대한 데이터셋을 바탕으로 격자 왜곡도와 기계적 물성 간의 상관관계를 도출한다. 2026년 현재 이 기술들은 실험적 시행착오를 줄이는 보조 도구를 넘어, 인간이 예측하기 어려운 새로운 결정 구조를 제안하는 자율 주행 실험실(SDL)의 핵심 엔진으로 기능하고 있다.

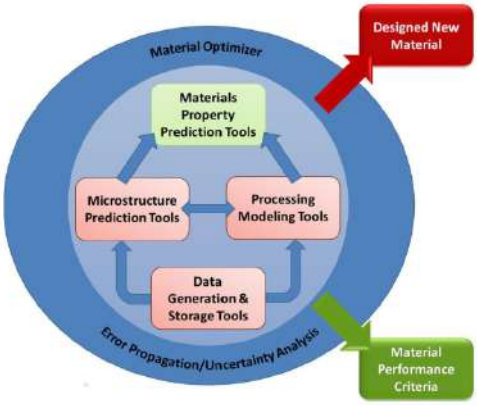


Figure 9. CALPHAD의 흐름도

Figure 9. CALPHAD의 흐름도

HEA의 가장 큰 도전 과제 중 하나는 방대한 조성 공간이다. 5개 원소만 고려하더라도 조합 가능한 조성은 사실상 무한에 가깝다. 이러한 문제를 해결하기 위해 CALPHAD와 AI 기반 설계가 결합된 접근법이 활용되고 있다. CALPHAD는 각 상의 자유에너지를 계산하여 상 안정성을 예측하는 방법으로, 다성분계에서도 상대적으로 신뢰성 있는 예측을 제공한다. 이를 통해 특정 조성이 단일 고용체를 형성할지, 금속간 화합물이 생성될지를 사전에 판단할 수 있다.한편 머신러닝은 격자 왜곡 파라미터 δ 예측, 항복강도 및 연성 예측, 그리고 상 형성 경향 분류와 같은 역할을 주로 수행한다.

최근에는 이러한 기술들이 결합되어, 자율 주행 실험실(Self-Driving Laboratory, SDL) 개념이 등장하고 있다. SDL은 계산–예측–실험–피드백 과정을 자동화하여, 인간 연구자가 직접 탐색하기 어려운 조성 영역까지 탐색할 수 있도록 한다(Zhang et al., 2014).

에너지 산업에서는 이러한 기술이 특히 중요하다. 수소 저장, 고온 구조재, 원자로 재료는 매우 엄격한 조건을 만족해야 하며, 기존의 경험적 설계만으로는 한계가 있다. CALPHAD와 AI는 이러한 복잡한 요구조건을 동시에 만족하는 최적 조성을 탐색하는 데 핵심적인 도구로 활용되고 있다.

5. Conclusion

고엔트로피 합금은 다성분계 전역에서 발생하는 격자 왜곡, 준안정 상변화, 미세조직 제어를 통합적으로 활용하는 차세대 소재 설계 플랫폼이다. 본 탐구 결과, HEA의 우수성은 단순히 무질서 자체에 있는 것이 아니라, 원자 단위의 무질서를 정밀하게 제어하여 전위 이동을 억제하고 에너지 흡수 능력을 극대화하는 데 있음을 확인하였다.

향후 HEA의 성공적인 산업화를 위해서는 4대 핵심 효과에 대한 무조건적인 수용보다 각 합금계의 특성에 맞는 물리적 검증이 지속되어야 한다. 특히 적층 제조 공정 최적화와 AI 기반의 비등원자비 설계 기술을 결합한다면, HEA는 탄소 중립을 위한 수소 인프라와 민간 우주 시대의 극한 환경을 극복할 가장 강력한 재료적 솔루션이 될 것이다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

Q1. **물성 분석** ***:*** **원자 반경 편차 수식을 바탕으로, 특정 합금계에서 격자 왜곡이 항복 강도 향상에 기여하는 정도를 정량화하고, 이때 화학적 단범위 질서*** ***(SRO)*** **가 미치는 영향을 논하시오*** ***.***

Q2. **환경 대응** ***:*** **극저온 환경에서 일부** ***HEA*** **가 일반 합금과 달리 파괴 인성이 상승하는 이유를 변형 유도 쌍정*** ***(Twinning)*** **및 격자 마찰 응력 변화와 결합하여 설명하시오*** ***.***

Q3. **공정 설계** ***:*** **적층 제조*** ***(AM)*** **시 발생하는 급속 응고가** ***HEA*** **의 상 분리를 억제하는 열역학적 원리를 설명하고, 이 과정에서 발생할 수 있는 잔류 응력 문제를 해결하기 위한 열처리 전략을 제안하시오*** ***.***

Q4. **산업 적용** ***: HEA*** **가 가진 여러 장점 중 하나만 선택해 실제 산업에 적용한다면, 어떤 분야에 가장 먼저 활용하고 싶은가*** ***?*** **그리고 그 이유는 무엇인가*** ***?***

References

Cantor, B., I. T. H. Chang, P. Knight, and A. J. B. Vincent. 2004. Microstructural Development in Equiatomic Multicomponent Alloys. Materials Science and Engineering: A **375–377**:213–218.

Gali, A., and E. P. George. 2013. Tensile Properties of High- and Medium-Entropy Alloys. Intermetallics **39**:74–78.

George, E. P., D. Raabe, and R. O. Ritchie. 2019. High-Entropy Alloys. Nature Reviews Materials **4**:515–534.

Gludovatz, B., A. Hohenwarter, D. Catoor, E. H. Chang, E. P. George, and R. O. Ritchie. 2014. A Fracture-Resistant High-Entropy Alloy for Cryogenic Applications. Science **345**:1153–1158.

He, Q.-F., and Y. Yang. 2018. On Lattice Distortion in High Entropy Alloys. Frontiers in Materials **5**:42.

Li, Z., K. G. Pradeep, Y. Deng, D. Raabe, and C. C. Tasan. 2016. Metastable High-Entropy Dual-Phase Alloys Overcome the Strength–Ductility Trade-Off. Nature **534**:227–230.

Maresca, F., D. Dragoni, G. Csányi, N. Marzari, and W. A. Curtin. 2022. The Origin of Jerky Dislocation Motion in High-Entropy Alloys. Nature Communications **13**:4929.

Miracle, D. B., and O. N. Senkov. 2017. A Critical Review of High Entropy Alloys and Related Concepts. Acta Materialia **122**:448–511.

Okamoto, N. L., K. Yuge, K. Tanaka, H. Inui, and E. P. George. 2016. Atomic Displacement in the CrMnFeCoNi High-Entropy Alloy – A Scaling Factor to Predict Solid Solution Strengthening. AIP Advances **6**:125008.

Roylance, D., 2001. The Dislocation Basis of Yield and Creep. Technical report, Massachusetts Institute of Technology, Department of Materials Science and Engineering. https://web.mit.edu/course/3/3.11/www/modules/dn.pdf.

University of Cambridge, 2004. Dislocations in 3D – DoITPoMS Teaching and Learning Package. https://www.doitpoms.ac.uk/tlplib/dislocations/dislocations_in_3D.php.

Varvenne, C., A. Luque, and W. A. Curtin. 2016. Theory of Strengthening in FCC High Entropy Alloys. Acta Materialia **118**:164–176.

Yeh, J.-W., S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang. 2004. Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes. Advanced Engineering Materials **6**:299–303.

Zhang, Y., Y. J. Zhou, J. P. Lin, G. L. Chen, and P. K. Liaw. 2008. Solid-Solution Phase Formation Rules for Multi-Component Alloys. Advanced Engineering Materials **10**:534–538.

Zhang, Y., T. T. Zuo, Z. Tang, M. C. Gao, K. A. Dahmen, P. K. Liaw, and Z. P. Lu. 2014. Microstructures and Properties of High-Entropy Alloys. Progress in Materials Science **61**:1–93.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Sangho Oh

M. shoh@kentech.ac.kr

Dr. Sang Ho Oh is Professor and Director of the Center for Shared Research Facilities at the Korea Institute of Energy Technology (KENTECH) in Naju, Republic of Korea. He received his Ph.D. in Materials Science and Engineering from POSTECH in 2002. Before joining KENTECH in 2022, he held faculty positions at Sungkyunkwan University (2016–2022) and POSTECH (2009–2015), and research appointments at the Max Planck Institute for Metals Research, the Korea Research Institute of Standards and Science (KRISS), the Erich Schmid Institute in Leoben, Austria, Oak Ridge National Laboratory, and the Korea Basic Science Institute (KBSI). He has co-authored more than 200 publications on topics including the atomic and electronic structures of interfaces, deformation behavior of nanostructured materials and natural biomaterials, switching mechanisms in non-volatile devices, and strain distribution in epitaxial heterostructures. His current research focuses on the atomic-scale characterization of energy materials and devices, with emphasis on surface and interface structures, in situ electron microscopy, electron holography, ptychography, atomic-resolution HRTEM/STEM imaging, and 4D-STEM. He also develops machine-learning-based approaches for electron microscopy data analysis.

KENTECH
Korea Institute of Energy Technology

Vol. 05 (2026)

Effects of Lattice Distortion on Dislocation Migration Inhibition and Mechanical Properties in Multicomponent Solid Solutions: Analysis of Strengthening Mechanisms in High-Entropy Alloys (HEA)

Abstract

High-entropy alloys (HEAs) are next-generation structural materials made by mixing five or more elements in nearly equal proportions. Their high mixing entropy, expressed as $\Delta S_{\mathrm{mix}} \geq 1.5R$, helps stabilize a single solid-solution phase and fundamentally redefines the traditional paradigm of alloy design. HEAs can exhibit an exceptional combination of high strength and toughness that conventional alloys often cannot achieve in extreme environments, such as cryogenic liquid hydrogen storage tanks, aircraft turbine blades, and nuclear structural components. As of 2026, the rapid expansion of the hydrogen economy and the technological advancement of the aerospace industry have created an unprecedented demand for structural materials that can withstand extreme conditions. In this context, HEAs are emerging as one of the most promising candidate materials. This study analyzes and explains, at the atomic scale, the physical mechanism by which lattice distortion, one of the main origins of the excellent mechanical properties of HEAs, suppresses the motion of dislocations responsible for plastic deformation.

Introduction

Modern research on metallic materials is no longer limited to finding materials that are simply stronger and lighter. It is increasingly moving toward the design of material structures at the atomic scale. Within this trend, a material that overturns conventional ideas has emerged: the high-entropy alloy (HEA). Traditional alloy design is based on one principal element, such as iron, aluminum, or copper, with small amounts of additional elements. For decades, it was widely believed that mixing many elements in large amounts would form complex intermetallic compounds and make the material brittle. However, in 2004, Professor Yeh and his research team in Taiwan reported that when five or more elements are mixed in nearly equal proportions, they can instead form a simple and stable single-phase solid solution. This discovery had a major impact on the field of materials science (Yeh et al., 2004). The reason why this counterintuitive phenomenon is possible lies in entropy, a key concept in thermodynamics. This is why these materials are called high-entropy alloys.

The main properties of HEAs can be summarized by four core effects. The first is the high-entropy effect. When multiple elements are mixed, the mixing entropy, ΔS_{mix}, increases. According to the Gibbs free energy relationship, $\Delta G = \Delta H - T\Delta S$, a larger entropy term lowers the free energy at high temperatures and thermodynamically stabilizes a simple solid-solution phase. In simple terms, thermodynamics explains the paradox that a more complex chemical mixture can produce a simpler crystal structure.

The second is the severe lattice distortion effect. This effect occurs when elements with different atomic sizes and properties randomly share a single crystal lattice, causing local distortion of the lattice. This is the most direct reason why HEAs show high strength, and it is the main focus of this research paper.

The third is the sluggish diffusion effect. In HEAs, the coexistence of many different elements creates a complex energy landscape that slows atomic movement. As a result, the microstructure remains more stable even

at high temperatures, and the alloy can show excellent creep resistance.
The fourth is the cocktail effect. This effect refers to the synergistic properties that arise from interactions among multiple elements. The unique simultaneous improvement in strength and toughness, which cannot be predicted by a simple rule of mixtures, is related to this effect.

Because of these properties, HEAs are receiving particular attention in two key industries as of 2026. In the hydrogen economy, CrMnFeCoNi-based HEAs are being actively studied as candidate materials for tanks and pipelines that store and transport liquid hydrogen at $-253\,^{\circ}\mathrm{C}$. Conventional metallic materials can undergo brittle fracture at cryogenic temperatures. In contrast, this alloy shows the unusual property of improved toughness even at $-196\,^{\circ}\mathrm{C}$ (Gludovatz et al., 2014). This property is directly connected to the safety of liquid hydrogen storage vessels. In the aerospace industry, aircraft engine turbine blades must be exposed to high-temperature, high-pressure, and oxidative environments above $1,200\,^{\circ}\mathrm{C}$ for thousands of hours. Refractory high-entropy alloys (RHEAs), especially MoNbTaVW-based systems, show high-temperature strength beyond that of conventional superalloys and are attracting attention as next-generation blade materials.

However, HEA research still faces several challenges. The greatest difficulty is the complexity of phase diagram prediction. In multicomponent systems containing five or more elements, the dimensionality of the phase diagram increases dramatically. Therefore, it is very difficult to predict without experiments which compositions will form a single solid solution and which will form intermetallic compounds. To address this problem, CALPHAD, a thermodynamic calculation method, and machine-learning-based composition prediction models are being actively developed. In addition, the process of uniformly mixing multiple high-purity elements and manufacturing them into large components is more expensive than conventional alloy production. It is also difficult to control quality variation during processing. These issues of manufacturing cost and process complexity remain major obstacles to commercialization. Paradoxically, these challenges show why it is necessary to understand the strengthening mechanisms of HEAs more precisely and to systematically investigate the relationship between composition, structure, and properties.

1. Background Knowledge: Crystal Lattices and Dislocations

1-1. crystal structure of metals

The atoms that make up metals are not randomly scattered. Instead, they form a regular three-dimensional arrangement called a crystal lattice. Depending on the arrangement of atoms, metallic crystals can be classified into several structures, such as face-centered cubic (FCC), body-centered cubic (BCC), and hexagonal close-packed (HCP) structures. CrMnFeCoNi-based HEAs mainly have an FCC structure. In this structure, atoms are located at the eight corners of each unit cell and at the centers of the six faces. The FCC structure is important because it has 12 slip systems, which allow plastic deformation to be distributed in several directions. This crystallographic feature explains why CrMnFeCoNi alloys can maintain high ductility even at cryogenic temperatures.

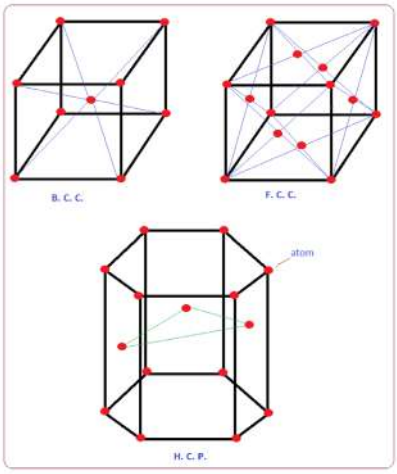


Figure 1. Schematic unit-cell structures of BCC, FCC, and HCP crystals. CrMnFeCoNi HEAs have an FCC structure. FCC crystals have 12 independent slip systems along the {111} planes, which makes plastic deformation easier at room temperature compared with BCC crystals, which have many slip systems but high critical shear stress, and HCP crystals, which have only a limited number of slip systems.

1-2. Main Carrier of Plastic Deformation: Dislocation

When an external force is applied to a metal, the distance between atoms changes slightly at first. If the force is removed, the metal returns to its original shape. This is called elastic deformation. However, when the applied force exceeds a critical value called the yield strength, the material begins to undergo plastic deformation, in which its shape changes permanently. At first glance, plastic deformation may seem as if an entire atomic layer slides at the same time. However, the theoretical shear strength required for this process is about 100 to 1,000 times larger than the experimentally measured value. This large discrepancy is explained by a linear defect called a dislocation.

A dislocation is an imperfect line in the atomic arrangement inside a crystal lattice. Plastic deformation does not occur because the entire lattice moves at once. Instead, it occurs because dislocations move through the lattice step by step, one atomic spacing at a time. This process is similar to moving a long carpet by pushing a wrinkle forward, rather than pulling the entire carpet at once. This is why plastic deformation can occur with much less energy.

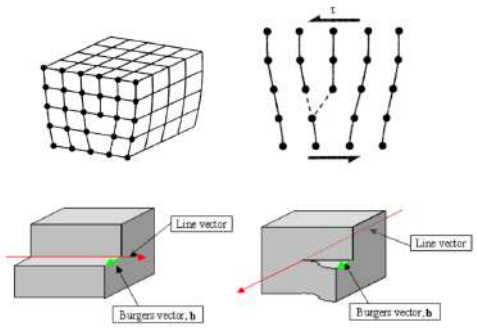


Figure 2. Schematic illustration of the formation and motion of dislocations.

The upper-left part of the figure shows the structure of an edge dislocation. The upper-right part shows that, under shear stress τ, a dislocation moves step by step over atomic distances, resulting in plastic deformation. The lower part compares an edge dislocation and a screw dislocation. In an edge dislocation, the Burgers vector $\mathbf{b}$ is perpendicular to the dislocation line. In a screw dislocation, the Burgers vector $\mathbf{b}$ is parallel to the dislocation line. The magnitude $|\mathbf{b}|$ corresponds to the actual distance that atoms move during dislocation motion.

The characteristics of a dislocation can be quantified by the Burgers vector, $\mathbf{b}$. This vector describes how much the atomic arrangement around a dislocation deviates from a perfect crystal lattice. Its magnitude, $|\mathbf{b}|$, corresponds to the distance that atoms move when a dislocation advances. In an edge dislocation, $\mathbf{b}$ is perpendicular to the dislocation line. In a screw dislocation, $\mathbf{b}$ is parallel to the dislocation line. The critical resolved shear stress (CRSS) required for dislocation motion can be approximated as $\tau_{\mathrm{CRSS}} \propto \mu|\mathbf{b}|/L$. Here, μ is the shear modulus, and L is the average spacing between dislocations or obstacles. As the number of obstacles along the dislocation path increases, L decreases and the CRSS increases. As a result, the material becomes stronger. The main argument of this research paper is that lattice distortion in HEAs acts as this type of obstacle to dislocation motion.

2. Theory and Mecahnism

2-1. Origin of Lattice Distortion in HEAs

In pure metals and conventional binary alloys, only one or two types of atoms occupy the lattice, so the lattice spacing is nearly uniform. In HEAs, however, five or more elements with different atomic radii randomly occupy the lattice sites of the crystal structure. Taking the representative CrMnFeCoNi alloy as an example, the differences in atomic radius among Cr (128 pm), Mn (127 pm), Fe (126 pm), Co (125 pm), and Ni (124 pm) cause local size mismatch throughout the lattice. In addition, differences in electronegativity, $\Delta\chi$, further distort the lattice by changing the directionality and strength of interatomic bonding. For example, the electronegativities of Cr and Ni are 1.66 and 1.91, respectively, giving a difference of 0.25. This difference creates an imbalance in electron density within the lattice and changes the interatomic bond distances.

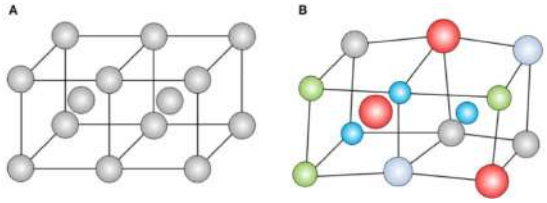


Figure 3. Schematic comparison of the crystal lattices of a pure metal and an HEA.

(A) In a pure metal, atoms of the same size occupy lattice sites uniformly, so the lattice remains regular. (B) In an HEA, five or more elements with different atomic radii (distinguished by color) randomly occupy the lattice sites, producing overall lattice distortion. This distortion forms internal stress fields and serves as the starting point of the strengthening mechanism that suppresses dislocation motion.

The degree of lattice distortion is quantified by the atomic size mismatch parameter, δ. This value is defined in a manner similar to the standard deviation of atomic radii relative to their weighted average, and it is calculated as follows:

$$\delta = \sqrt{\sum x_i \left(1 - \frac{r_i}{\bar{r}}\right)^2} \times 100\,(\%)$$

Here, x_i is the mole fraction of element i, r_i is its atomic radius, and $\bar{r} = \sum x_i r_i$ is the average atomic radius of the alloy. A larger δ means more severe lattice distortion. In general, when $\delta \geq 1\%$, the solid-solution strengthening effect becomes significant. Zhang et al. (2008) showed that when $\delta < 6.6\%$, a simple solid solution tends to form stably, whereas intermetallic compounds are more likely to form above this value. In other words, there is an optimization condition in HEAs: lattice distortion must be large enough to strengthen the alloy, but it must also remain below a certain limit in order to preserve the solid-solution structure.

2-2. Internal Stress Fields Generated by Lattice Distortion

When elements of different sizes randomly occupy lattice sites, the effect is not limited to a simple increase or decrease in lattice spacing. From the viewpoint of elastic mechanics, if an atom with a larger radius than its neighbors occupies a lattice site, it pushes the surrounding lattice outward and creates a compressive stress field around itself. In contrast, a smaller atom creates a tensile stress field because the surrounding lattice is pulled inward. In pure metals, where all atoms have the same size, such local stress fields do not exist, and the potential energy landscape of the crystal has uniform periodicity. In HEAs, however, five or more elements are arranged randomly, so these local stress fields are distributed irregularly and non-uniformly throughout the entire crystal. As a result, the potential energy landscape of the crystal no longer resembles the smooth surface of a calm lake. Instead, it resembles a rugged mountain range, with hills and valleys of different sizes spread disorderly across the structure.

2-3. How Lattice Distortion Suppresses Dislocation Motion

For a dislocation to move, it must overcome the Peierls–Nabarro barrier, which is the periodic potential energy barrier created by the crystal lattice. In pure metals, this barrier is relatively low and repeats periodically, so dislocations can move rather easily. In the non-uniform energy landscape of an HEA, however, dislocation motion is hindered in two main ways.

A dislocation itself also has its own stress field. In the case of an edge dislocation, for example, the lattice above the dislocation line is under compression, whereas the lattice below it is under tension. When this intrinsic stress field of the dislocation interacts with the local stress fields already present in the lattice, elastic interaction occurs. This interaction is described by the Peach–Koehler equation. The force F acting on a dislocation is given by $F = (\sigma \cdot b) \times \xi$, where σ is the local stress tensor, b is the Burgers vector, and ξ is the unit vector along the dislocation line. In HEAs, σ changes irregularly at each point along the dislocation path, so both the magnitude and the direction of the force acting on the dislocation continuously change.

This leads to two suppressing effects. The first is the overall increase in friction stress, τ_f. Because the average height of the energy barriers along the entire dislocation path increases, the minimum stress required to move the dislocation also increases. This is the most direct reason for the improvement in yield strength. The second is dislocation pinning. As a dislocation moves, it may reach a lattice site that is especially low in energy and therefore thermodynamically stable compared with its surroundings. When this happens, the dislocation becomes pinned at that site. In HEAs, such pinning sites are distributed at high density throughout the lattice, so a dislocation is pinned again after moving only a very short distance. Additional stress is then required to detach and move the pinned dislocation. As a result, both the yield strength and the work-hardening rate of the material increase.

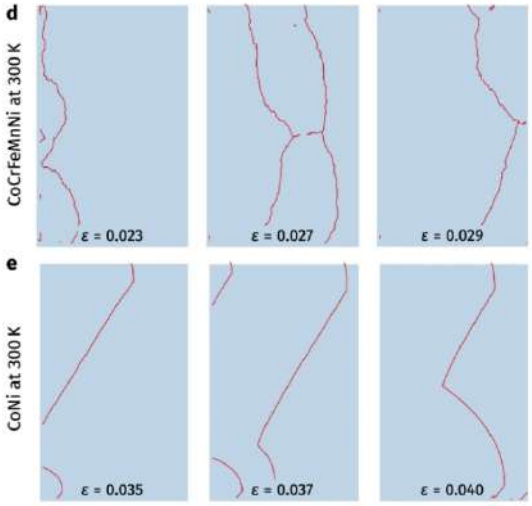


Figure 4. Comparison of dislocation motion paths in an HEA and a conventional alloy (atomic simulation at 300 K).

(Top, d) In the CoCrFeMnNi HEA, the dislocation line is repeatedly pinned by irregular Peierls barriers and follows a wavy path. (Bottom, e) In the CoNi binary alloy, the dislocation line remains nearly straight during motion. This difference is the direct origin of the high yield strength of the HEA. ε denotes the applied strain.

2-4. Quantitative Description of Strengthening by Lattice Distortion

Conventional solid-solution strengthening theory is based on the concept that the stress fields created by a small number of solute atoms hinder dislocation motion. In HEAs, this concept is taken to the extreme because every element acts simultaneously as both solute and solvent. Varvenne et al. (Varvenne et al., 2016) extended this idea to multicomponent HEAs and showed that the increase in yield strength due to solid-solution strengthening, $\Delta\tau$, can be approximated as follows:

$$\Delta\tau \propto \mu \cdot |b| \cdot \sqrt{\sum x_i \cdot \epsilon_i^2}$$

Here, μ is the shear modulus, $|b|$ is the magnitude of the Burgers vector, x_i is the mole fraction of element i, and ϵ_i is the parameter representing the strength of the local elastic stress field created by element i. The key term is $\sum x_i \epsilon_i^2$. In conventional alloys, only a small amount of added solute contributes to this summation, so the value remains relatively small. In HEAs, however, all five or more constituent elements have their own ϵ_i values and contribute equally to the total summation. As a result, this term becomes much larger than in conventional alloys. This quantitatively explains why HEAs show much higher strength than ordinary solid-solution-strengthened alloys.

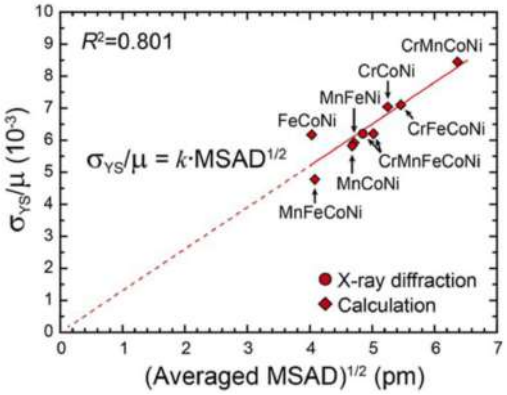


Figure 5. Comparison of solid-solution strengthening according to elemental combination.

This graph shows the 0 K yield strength (y-axis) of the CrMnFeCoNi HEA and its derived ternary and quaternary alloys as a function of the degree of lattice distortion (x-axis), expressed as the square root of the MSAD. As the number of constituent elements increases, lattice distortion becomes larger, and the yield strength increases proportionally. This trend is consistent with Varvenne's strengthening theory, expressed as $\Delta\tau \propto \sqrt{\sum x_i \epsilon_i^2}$.

3. Methodology: Processing Routes for High-Entropy Alloys

The synthesis process of HEAs is not merely a procedure for fabricating specimens. It is a key variable that determines whether strengthening mechanisms can be activated by controlling compositional uniformity, grain refinement, and defect density. In other words, the processing route can be understood as a process that determines the internal energy state of the HEA. Each processing method uses a different type of energy input: thermal energy in arc melting, mechanical energy in mechanical alloying, and localized photothermal energy in additive manufacturing. As a result, the free-energy state of the resulting microstructure differs depending on the processing route.

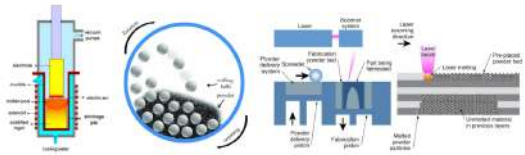


Figure 6. Schematic visualization of the processing routes of VAM, mechanical alloying, and additive manufacturing, respectively.

In particular, in additive manufacturing, rapid solidification can freeze a non-equilibrium state. This allows a structure that is not fully thermodynamically stable but is mechanically favorable to be retained. Such non-equilibrium structures contain high dislocation density and fine grains, which suppress dislocation motion more effectively.

The differences among processing routes are not limited to simple microstructural differences. They also change the formation, motion paths, and interaction modes of dislocations. In this sense, the strengthening mechanism of HEAs is not determined only by composition. Rather, it is reconstructed by the energy landscape formed during processing.

Each processing route also sets a different path to thermodynamic equilibrium. Melting-based processes allow the alloy to approach an equilibrium state, whereas solid-state synthesis and additive manufacturing tend to freeze non-equilibrium states. This difference is a key factor that determines the distribution of energy barriers against dislocation motion.

3-1. Vacuum Arc Melting (VAM)

Vacuum arc melting is a standard melting-based process used in HEA research. In this process, high-purity metallic elements are melted under a high-vacuum or inert gas atmosphere. Because HEAs are multicomponent alloys, differences in melting point and diffusion rate among elements can easily lead to dendritic structures and chemical segregation during cooling.

The dendritic structure formed during arc melting is not merely a microstructural feature. It also represents spatial non-uniformity in the distribution of lattice distortion caused by local compositional differences. The compositional difference between dendrite cores and interdendritic regions induces local changes in the atomic size mismatch parameter, δ. This creates different energy barriers during dislocation motion. In this type of structure, a moving dislocation does not experience uniform friction stress. Instead, it experiences different levels of resistance depending on its position. As a result, the dislocation may move easily in some regions but become strongly pinned in others, leading to non-uniform deformation behavior.

To solve this problem, at least five cycles of flip-and-remelt processing and long homogenization heat treatment are generally required. Although this method is useful for fundamental research, it has limitations in controlling local phase separation when producing large structural components or complex-shaped parts for energy-related applications.

Long homogenization heat treatment improves compositional uniformity, but it can also induce grain growth. According to the Hall–Petch relationship,

$$\sigma_y = \sigma_0 + k d^{-1/2},$$

an increase in grain size, d, can reduce the yield strength (Cantor et al., 2004). Therefore, in arc melting, process optimization is essential to balance compositional uniformity and grain refinement.

From an energy perspective, this process requires a large thermal energy input. Its energy efficiency may therefore be limited when fabricating large structural components. For industrial applications, the balance between processing efficiency and microstructural control becomes an important design variable.

3-2. Mechanical Alloying (MA) and Solid-State Synthesis

Mechanical alloying induces atomic-scale mixing in the solid state by high-energy ball milling of metallic powders, without using a melting process. This route is advantageous for alloying elements with extremely different melting points. It also maximizes yield strength by forming nanocrystalline structures through repeated plastic deformation.

However, contamination from the milling media and oxidation control are major challenges. For application as energy structural materials, pores must also be completely removed during subsequent densification processes such as spark plasma sintering (SPS). The repeated plastic deformation during mechanical alloying should be understood not only as powder mixing but also as a process that accumulates high dislocation density, ρ, inside the crystal structure. This dislocation density contributes to mechanical strength through a

relationship such as

$$\sigma \propto \alpha \mu b \sqrt{\rho}.$$

Here, the dislocations accumulated during milling may remain in the sintered bulk material. This can lead to an increase in the initial yield strength (Yeh et al., 2004).

The nanocrystalline structure formed by mechanical alloying is not simply a result of size reduction. It also creates a structure in which grain boundaries act as major obstacles to dislocation motion. In HEAs, these boundaries can also be compositionally non-uniform. Therefore, additional energy barriers can form when dislocations pass through grain boundaries. Structures formed under non-equilibrium conditions may become partially stabilized during heat treatment. However, they may not fully move toward equilibrium and can remain in a kinetically trapped state. This state is thermodynamically unstable, but it can provide high mechanical strength (Yeh et al., 2004).

From the viewpoint of energy-system applications, mechanical alloying is highly useful for synthesizing HEAs based on high-melting-point elements, especially refractory high-entropy alloys. However, because this is a powder-based process, porosity control is crucial. Porosity directly affects the fatigue life and hydrogen embrittlement resistance of the final structural material.

3-3. Additive Manufacturing (AM)

As of 2026, additive manufacturing, including laser powder bed fusion (L-PBF), is a key processing route for the industrialization of HEAs. Rapid solidification at rates of approximately 10^3–$10^6\ \mathrm{K/s}$ occurs through localized heat input from a laser. This suppresses long-range atomic diffusion and causes solute trapping, which enables the formation of metastable solid solutions with reduced macroscopic segregation.

Rapid solidification in additive manufacturing does more than refine the microstructure. It also induces a transition from diffusion-controlled growth to interface-controlled growth. As a result, metastable solid solutions and non-equilibrium microstructures, which are difficult to form under conventional equilibrium solidification, can be produced.

Solute trapping occurs when the moving speed of the solid–liquid interface exceeds the diffusion speed of solute atoms during rapid solidification. As a result, solute elements remain dissolved at concentrations higher than their equilibrium values. This increases the spatial uniformity of lattice distortion and strengthens the suppression of dislocation motion (Zhang et al., 2014).

In addition, nanoscale chemical modulation at the boundaries of the fine cellular structures formed during additive manufacturing creates a hierarchical structure that blocks dislocation motion in multiple stages. This feature can be especially useful for improving the durability of hydrogen storage devices and lightweight aerospace structures. The cellular structures formed during additive manufacturing have sizes below the micrometer scale. Their boundaries contain fine chemical modulation caused by local differences in elemental concentration.

This structure creates a multistage dislocation-resistance mechanism. Inside the cells, solid-solution strengthening occurs due to lattice distortion. At the cell boundaries, dislocation scattering occurs due to changes in chemical composition. In the fine grains, grain boundaries block dislocation motion. Therefore, this structure does not suppress dislocations through a single mechanism. Instead, it suppresses them hierarchically.

Furthermore, the high thermal gradient formed during additive manufacturing can induce a directionally solidi-

fied microstructure. This can cause anisotropy, which strengthens mechanical properties in specific directions (Zhang et al., 2014). This characteristic can be advantageous in systems such as aerospace structures, where the loading direction is clearly defined.

From the viewpoint of the energy industry, AM-based HEAs have important significance. In hydrogen storage systems, crack suppression under cryogenic conditions is essential. The microstructure formed by additive manufacturing can make crack propagation paths more complex and improve fracture resistance. In the aerospace field, both lightweight design and high strength are required. Additive manufacturing can maximize the strength-to-weight ratio by allowing the design of internal lattice structures.

4. Discussion: Strengthening Mechanisms and Energy-Relevant Design Strategies

From an energy perspective, the strengthening mechanism of HEAs can be understood through the following question: how is externally applied energy distributed inside the material? When a metal is deformed, the external work is mainly distributed into three forms: energy used for dislocation motion, energy dissipated as heat, and energy used for microstructural changes such as phase transformation and defect formation. In HEAs, lattice distortion and local stress fields increase the fraction of energy required for dislocation motion. This means that more external energy is needed to produce the same amount of deformation. As a result, the material exhibits higher strength.

4-1. Lattice Distortion and Dislocation Mobility

The strengthening mechanism of HEAs is based on lattice distortion caused by differences in atomic radius and elastic modulus among multiple elements. These local stress fields act as friction stress against dislocation motion, thereby increasing the yield strength. However, it is risky to generalize that all HEAs have severe lattice distortion. This effect is selectively maximized in alloys containing elements with large atomic radius differences.

Recent studies have shown that not only lattice distortion but also short-range order (SRO) can provide additional strengthening by changing the energy state of dislocation lines. Therefore, precise control of SRO has become an important design strategy (Miracle and Senkov, 2017; George et al., 2019).

The local stress fields formed by lattice distortion do more than simply resist dislocation motion. They also randomize the motion path of dislocations. In other words, dislocations cannot move linearly along ideal slip planes. Instead, they search for energetically favorable paths and follow non-linear trajectories. This process makes the spatial distribution of the Peierls–Nabarro barrier non-uniform and increases the critical stress required for dislocation motion. In HEAs, two strengthening mechanisms operate at the same time: an increase in average friction stress as a global strengthening effect and suppression of motion by local pinning.

Recent research has also reported that SRO changes the energy state of the dislocation core and further suppresses dislocation nucleation and motion. This suggests a new design direction beyond the conventional solid-solution strengthening model, in which the interaction between chemical order and mechanical properties must be considered (George et al., 2019). From an energy perspective, this structure increases the activation energy for dislocation motion, $\Delta G_{\mathrm{activation}}$. This means that externally applied mechanical energy must exceed a higher critical threshold before plastic deformation can begin.

4-2. Metastability Engineering: TRIP and TWIP Effects

Metastability engineering intentionally lowers phase stability to overcome the trade-off between strength and ductility. The transformation-induced plasticity (TRIP) effect, in which the face-centered cubic (FCC) structure transforms into a martensitic structure during deformation, and the twinning-induced plasticity (TWIP) effect, in which deformation twins are formed, continuously maintain the work-hardening rate. As a result, crack propagation can be effectively delayed.

These damage-tolerant properties are key mechanisms for ensuring the reliability of cryogenic hydrogen tanks exposed to large pressure changes and aerospace components subjected to high-speed vibration. The TRIP and TWIP effects are not merely deformation mechanisms. They can also be understood as mechanisms that expand the internal conversion pathways of mechanical energy.

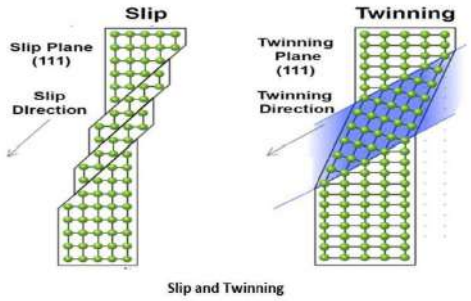


Figure 7. In slip, atomic layers slide directly over one another, whereas in twinning, the structural orientation of a local region changes and a new boundary is formed.

The externally applied work can be expressed as follows:

$$W_{\text{external}} = W_{\text{dislocation}} + W_{\text{heat}} + W_{\text{phase}}$$

Here, W_{phase} represents the energy used for phase transformation. In HEAs where the TRIP effect is activated, the contribution of this term increases. As a result, the energy that would otherwise be concentrated on dislocation motion is distributed into phase transformation.

The TWIP effect also forms twin boundaries inside the crystal. These twin boundaries act as dynamic obstacles that repeatedly block dislocation motion. Therefore, dislocation density continues to increase, and the work-hardening rate remains high. This mechanism becomes especially significant under cryogenic conditions because dislocation motion is restricted at lower temperatures. As a result, the relative contribution of phase transformation and twin formation increases (Li et al., 2016).

From the perspective of the energy industry, these properties are important not only for increasing strength but also for enabling damage-tolerant design. Since the material can absorb and dissipate more energy before fracture, it plays a critical role in safety-sensitive systems such as hydrogen storage tanks and aerospace structures.

4-3. Sluggish Diffusion and Thermal Stability

The initial hypothesis of sluggish diffusion was proposed as a basis for thermal stability because it could suppress grain growth and precipitate coarsening at high temperatures. However, recent studies have shown that diffusion rates in HEAs can be similar to, or even faster than, those in conventional alloys depending on the alloy system and temperature. Therefore, sluggish diffusion should not be assumed to be a universal feature of all HEAs.

For high-temperature energy-system applications, experimental verification of diffusion behavior and oxidation resistance in the relevant temperature range must be conducted first. Diffusion behavior is one of the key factors that determine the high-temperature stability of HEAs. In general, the diffusion coefficient is expressed as follows:

$$D = D_0 \exp\left(-\frac{Q}{RT}\right)$$

Here, Q is the activation energy for diffusion. In HEAs, multicomponent interactions may increase Q, which can reduce the diffusion rate (Miracle and Senkov, 2017).

However, recent studies show that the diffusion behavior of HEAs cannot simply be generalized as “slow.” In some alloys, diffusion is similar to that in conventional alloys, and faster diffusion has also been observed for certain elements. This occurs because the local chemical environments inside HEAs are highly diverse (Miracle and Senkov, 2017; Zhang et al., 2014).

Therefore, the thermal stability of HEAs cannot be explained only by diffusion rate. It is determined by the combined effects of energy barriers caused by lattice distortion, the distribution of multicomponent bonding energies, and oxidation and corrosion reactions. In high-temperature energy systems such as turbines and nuclear reactors, diffusion, oxidation, and creep behavior must be considered together. This remains an important research direction in HEA design.

4-4. Non-Equiatomic Design and Economic Viability

Non-equiatomic design, which adjusts the ratio of specific elements instead of using traditional equiatomic compositions, is a strategy that can optimize performance and reduce cost at the same time. Lean-HEA design maintains the entropy-stabilization effect while reducing the amount of expensive elements such as Co and Ni. This is a practical approach for improving the commercial viability of HEAs in energy industries where resource supply-chain risks are critical.

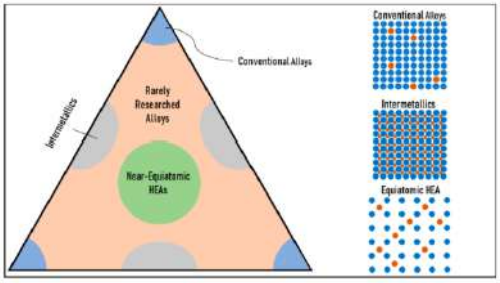


Figure 8. Comparison of atomic arrangements in conventional alloys, intermetallic compounds, and equiatomic HEAs. This schematic shows that HEAs form multicomponent random solid-solution structures.

Non-equiatomic design is more than a cost-reduction strategy. It is a multivariable design problem that optimizes the balance among mechanical properties, thermodynamics, and processability. Equiatomic HEAs provide maximum mixing entropy, but this does not always guarantee the best mechanical properties. By adjusting the ratio of specific elements, it is possible to control the degree of lattice distortion, change phase stability between FCC and BCC structures, and tune the conditions for TRIP activation.

For example, increasing the Mn content can reduce FCC stability and activate the TRIP effect, which can improve ductility (George et al., 2019). Lean-HEA design can also reduce the use of strategic metals such as Co and Ni, thereby lowering supply-chain risks. This factor is especially important in the energy industry. In the hydrogen economy, large-scale production is essential, so alloy design must consider both cost and performance.

4-5. CALPHAD and AI-Driven Discovery

The problem of high exploration cost caused by the vast compositional space of HEAs is being addressed through the combination of CALPHAD and machine learning. CALPHAD predicts thermodynamic phase stability, whereas AI derives correlations between lattice distortion and mechanical properties from large datasets. As of 2026, these technologies are no longer only auxiliary tools for reducing experimental trial and error. They are becoming core engines of self-driving laboratories (SDLs), which can propose new crystal structures that are difficult for humans to predict.

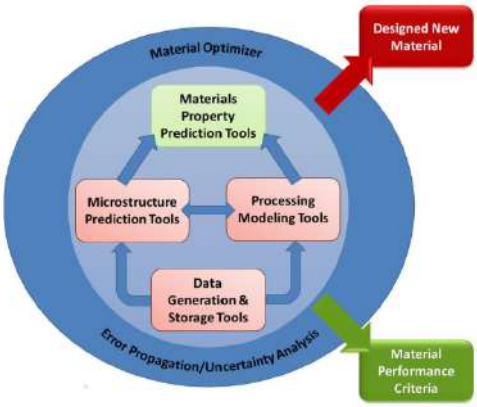


Figure 9. CALPHAD의 흐름도

Figure 9. Flowchart of the CALPHAD-based alloy design process.

One of the greatest challenges in HEA research is the enormous compositional space. Even when only five elements are considered, the number of possible compositions is almost infinite. To address this problem, CALPHAD and AI-based design approaches are being combined. CALPHAD calculates the free energy of each phase and predicts phase stability. It can provide relatively reliable predictions even in multicomponent systems. This makes it possible to estimate in advance whether a certain composition will form a single solid solution or intermetallic compounds.

Meanwhile, machine learning is mainly used to predict the lattice distortion parameter, δ, estimate yield strength and ductility, and classify phase-formation tendencies. Recently, these technologies have been integrated into the concept of the self-driving laboratory. An SDL automates the cycle of computation, prediction, experiment, and feedback, allowing researchers to explore compositional regions that would be difficult to investigate manually (Zhang et al., 2014).

These technologies are especially important in the energy industry. Materials for hydrogen storage, high-temperature structural components, and nuclear reactors must satisfy very strict requirements. Conventional empirical design has clear limitations in meeting these complex demands. CALPHAD and AI are therefore becoming essential tools for discovering optimal compositions that can satisfy multiple requirements at the same time.

5. Conclusion

High-entropy alloys are next-generation materials design platforms that integrate lattice distortion, metastable phase transformation, and microstructural control across multicomponent systems. This research shows that the superiority of HEAs does not come simply from disorder itself. Rather, it comes from the precise control of atomic-scale disorder, which suppresses dislocation motion and maximizes energy absorption capacity.

For the successful industrialization of HEAs, the four core effects should not be accepted unconditionally. Instead, physical verification must continue for each alloy system. In particular, if additive manufacturing optimization is combined with AI-based non-equiatomic design, HEAs can become one of the most powerful materials solutions for overcoming extreme environments in hydrogen infrastructure for carbon neutrality and in the emerging era of commercial space technology.

Example Questions

Research may be conducted on any free topic other than the following example questions.

Q1. Property Analysis: Based on the equation for atomic radius mismatch, quantify the extent to which lattice distortion contributes to the improvement of yield strength in a specific alloy system, and discuss the effect of short-range order (SRO) on this relationship.

Q2. Environmental Response: Explain why the fracture toughness of some HEAs increases in cryogenic environments, unlike conventional alloys, by connecting this behavior to deformation-induced twinning and changes in lattice friction stress.

Q3. Process Design: Explain the thermodynamic principle by which rapid solidification during additive manufacturing (AM) suppresses phase separation in HEAs, and propose a heat-treatment strategy to address the residual stress that may occur during this process.

Q4. Industrial Application: If you could select only one advantage of HEAs and apply it to a real industrial field, which field would you choose first, and why?

References

Cantor, B., I. T. H. Chang, P. Knight, and A. J. B. Vincent. 2004. Microstructural Development in Equiatomic Multicomponent Alloys. Materials Science and Engineering: A **375–377**:213–218.

Gali, A., and E. P. George. 2013. Tensile Properties of High- and Medium-Entropy Alloys. Intermetallics **39**:74–78.

George, E. P., D. Raabe, and R. O. Ritchie. 2019. High-Entropy Alloys. Nature Reviews Materials **4**:515–534.

Gludovatz, B., A. Hohenwarter, D. Catoor, E. H. Chang, E. P. George, and R. O. Ritchie. 2014. A Fracture-Resistant High-Entropy Alloy for Cryogenic Applications. Science **345**:1153–1158.

He, Q.-F., and Y. Yang. 2018. On Lattice Distortion in High Entropy Alloys. Frontiers in Materials **5**:42.

Li, Z., K. G. Pradeep, Y. Deng, D. Raabe, and C. C. Tasan. 2016. Metastable High-Entropy Dual-Phase Alloys Overcome the Strength–Ductility Trade-Off. Nature **534**:227–230.

Maresca, F., D. Dragoni, G. Csányi, N. Marzari, and W. A. Curtin. 2022. The Origin of Jerky Dislocation Motion in High-Entropy Alloys. Nature Communications **13**:4929.

Miracle, D. B., and O. N. Senkov. 2017. A Critical Review of High Entropy Alloys and Related Concepts. Acta Materialia **122**:448–511.

Okamoto, N. L., K. Yuge, K. Tanaka, H. Inui, and E. P. George. 2016. Atomic Displacement in the CrMnFeCoNi High-Entropy Alloy – A Scaling Factor to Predict Solid Solution Strengthening. AIP Advances **6**:125008.

Roylance, D., 2001. The Dislocation Basis of Yield and Creep. Technical report, Massachusetts Institute of Technology, Department of Materials Science and Engineering. https://web.mit.edu/course/3/3.11/www/modules/dn.pdf.

University of Cambridge, 2004. Dislocations in 3D – DoITPoMS Teaching and Learning Package. https://www.doitpoms.ac.uk/tlplib/dislocations/dislocations_in_3D.php.

Varvenne, C., A. Luque, and W. A. Curtin. 2016. Theory of Strengthening in FCC High Entropy Alloys. Acta Materialia **118**:164–176.

Yeh, J.-W., S.-K. Chen, S.-J. Lin, J.-Y. Gan, T.-S. Chin, T.-T. Shun, C.-H. Tsau, and S.-Y. Chang. 2004. Nanostructured High-Entropy Alloys with Multiple Principal Elements: Novel Alloy Design Concepts and Outcomes. Advanced Engineering Materials **6**:299–303.

Zhang, Y., Y. J. Zhou, J. P. Lin, G. L. Chen, and P. K. Liaw. 2008. Solid-Solution Phase Formation Rules for Multi-Component Alloys. Advanced Engineering Materials **10**:534–538.

Zhang, Y., T. T. Zuo, Z. Tang, M. C. Gao, K. A. Dahmen, P. K. Liaw, and Z. P. Lu. 2014. Microstructures and Properties of High-Entropy Alloys. Progress in Materials Science **61**:1–93.

Contributors



Minseong Eom

M. eomminseong@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi

M. mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Sangho Oh

M. shoh@kentech.ac.kr

Dr. Sang Ho Oh is Professor and Director of the Center for Shared Research Facilities at the Korea Institute of Energy Technology (KENTECH) in Naju, Republic of Korea. He received his Ph.D. in Materials Science and Engineering from POSTECH in 2002. Before joining KENTECH in 2022, he held faculty positions at Sungkyunkwan University (2016–2022) and POSTECH (2009–2015), and research appointments at the Max Planck Institute for Metals Research, the Korea Research Institute of Standards and Science (KRISS), the Erich Schmid Institute in Leoben, Austria, Oak Ridge National Laboratory, and the Korea Basic Science Institute (KBSI). He has co-authored more than 200 publications on topics including the atomic and electronic structures of interfaces, deformation behavior of nanostructured materials and natural biomaterials, switching mechanisms in non-volatile devices, and strain distribution in epitaxial heterostructures. His current research focuses on the atomic-scale characterization of energy materials and devices, with emphasis on surface and interface structures, in situ electron microscopy, electron holography, ptychography, atomic-resolution HRTEM/STEM imaging, and 4D-STEM. He also develops machine-learning-based approaches for electron microscopy data analysis.

KENTECH
Korea Institute of Energy Technology

IX

분산 에너지 자원의 미래, 마이크로 그리드

p. 202 - 220

Vol. 6 (2026)

분산 에너지 자원의 미래, 마이크로그리드

Abstract 탄소 중립을 향한 전 세계적 흐름 속에서 기존 중앙집중형 전력 시스템의 한계가 드러나면서, 분산 에너지 자원을 기반으로 한 마이크로그리드가 차세대 에너지 인프라로 주목받고 있다. 본 탐구 자료는 마이크로그리드의 정의와 핵심 구성요소를 정리하고, KENTECH 캠퍼스에서 전기자전거 배터리를 이동형 ESS로 활용한 실제 사례를 통해 마이크로그리드의 실용적 가능성을 검토한다. 나아가 마이크로그리드의 본격적인 상용화를 위해 기술, 경제, 제도 측면에서 해결해야 할 과제를 제시하며, 마이크로그리드가 분산 에너지 시대의 핵심 인프라로 자리잡기 위한 방향성을 논의한다.

Introduction

우리가 일상에서 사용하는 전기는 발전소에서 생산되어 송전선과 변전소를 거쳐 가정, 학교, 공장 등으로 공급된다. 이러한 전력 공급 체계인 전력망(Grid) 은 오랫동안 안정적인 전기 공급을 가능하게 하며 현대 사회를 뒷받침해 왔다. 그러나 기후변화로 인한 자연재해의 증가, 노후화된 송전 인프라, 그리고 태양광·풍력과 같은 재생에너지의 확대는 기존 중앙집중형 전력망의 한계를 점차 드러내고 있다. 기존 전력망은 대규모 발전소에서 생산한 전기를 먼 거리까지 전달하는 구조이기 때문에 송전 손실이 발생하고, 주요 설비에 장애가 생기면 넓은 지역이 동시에 정전될 수 있다. 또한 발전량이 날씨에 따라 변하는 재생에너지가 증가하면서 기존 방식만으로는 전력의 생산과 소비를 안정적으로 균형 맞추는 데 어려움이 커지고 있다. 이러한 문제를 해결하기 위한 대안으로 마이크로그리드(Microgrid) 가 주목받고 있다. 마이크로그리드는 일정 지역 안에서 전기를 생산·저장·소비하는 소규모 전력망으로, 평상시에는 기존 전력망과 연계되어 운영되지만 필요할 경우 독립적으로 전력을 공급할 수 있다. 본 탐구 자료에서는 기존 전력망의 구조와 한계를 살펴본 뒤, 마이크로그리드의 개념과 핵심 구성요소를 소개하고, KENTECH 캠퍼스의 실제 사례를 통해 마이크로그리드의 활용 가능성과 미래 발전 방향을 살펴보고자 한다.

1. 기존 그리드 시스템

현재 대부분의 국가에서 운영되는 전력 시스템은 발전, 송전, 배전 소비의 단계적 구조를 따른다. 발전 단계 에서는 화력, 원자력, 수력 등 대규모 중앙 발전소에서 수백 MW GW 규모의 전력을 생산한다. 생산된 전력은 송전 효율을 높이기 위해 154kV 765kV의 초고압으로 승압된다. 송전 단계에서는 승압된 전력이 송전탑과 고압 송전선을 통해 수백 킬로미터를 이동하여 변전소까지 전달된다. 이 과정에서 저항 손실(줄 손실)로 인해 전체 발전량의 약 5 8%의 전력이 열로 소실된다. 배전 단계에서는 변전소에서 22.9kV 수준으로 강압된 뒤, 주택·상업시설·공장 등 최종 소비자에게 220V 또는 380V로 공급된다. 이 구조에서 전력의 흐름은 발전소에서 소비자의 단방향으로만 이루어지며, 소비자는 전력을 수동적으로 공급받는 주체에 머문다.

1-1. 중앙집중형 전력망의 한계

이러한 단방향·집중형 구조는 안정적인 환경에서는 규모의 경제를 실현하지만, 현대의 에너지 환경에서는 다음과 같은 구조적 한계를 드러낸다.

- **문제가 연속적으로 퍼질 수 있는 구조**
  발전소, 주요 변전소, 핵심 송전선 중 어느 한 곳에 장애가 발생하면 해당 계통 전체에 정전이 연쇄적으로 파급된다. 2003년 미국·캐나다 대정전 당시 하나의 송전선 장애가 연쇄 반응을 일으켜 5,500만 명이 영향을 받은 것이 대표적인 사례이다. 기후변화로 인해 태풍, 폭설, 폭염 등 극한 기상 이벤트가 잦아지면서 이 취약성은 더욱 부각되고 있다.

- **재생에너지 통합의 어려움**
  태양광·풍력은 날씨에 따라 발전량이 급변하는 간헐성을 가진다. 기존 중앙 계통은 수요와 공급을 실시간으로 정밀하게 일치시키도록 설계되어 있어, 예측 불가능한 분산 전원이 대량 연계될 경우 전압·주파수 불안정, 역조류 등의 문제가 발생한다.

- **원거리 송전으로 인한 손실 및 비효율**
  발전소와 소비지 간의 물리적 거리가 멀수록 송전 손실이 커지며, 소비지 인근에 발전원이 없는 도서·산간·개발도상국 지역은 전력망 구축 자체에 막대한 비용이 소요된다. 그 예로 대한민국의 울릉도가 있다. 울릉도는 육지와 200km 정도 떨어진 섬으로 자체적으로 발전하여 6.6kV 배전 선로를 통해 전력을 공급한다.

2. 마이크로그리드

2-1. 마이크로그리드의 정의

마이크로그리드는 중앙집중식 전력망(Grid)에 의존해 전력을 일방적으로 공급받고 소비하는 것이 아닌, 기존의 광역적인 전력 시스템으로부터 독립된 분산전원을 중심으로 한 국소적인 전력 공급 시스템, 즉 '자급자족' 할 수 있는 소규모의 전력 체계를 의미한다. 추가적으로 미국 DoE는 마이크로 그리드를 명확한 전기적 경계 내에서 상호 연결된 부하(전력 소비)와 분산 에너지 자원(전력 생산 및 저장)의 그룹으로 정의한다.



Figure 1. 마이크로그리드의 구성 및 전력 흐름도

기존 전력 시스템은 한전 계통으로부터 일방적으로 전력을 공급받는 구조로 정전이나 계통 장애 발생 시 전력 공급이 즉시 중단된다. 그러나 위 그림과 같이 마이크로그리드를 적용하면 태양광·풍력·에너지저장장치 등 내부 분산 에너지 자원을 통해 상위계통이 차단되더라도 자립적으로 전력을 공급할 수 있으며, 잉여 전력은 인근 마이크로그리드와 서로 공유하는 것도 가능하다. (Asmus et al., 2018)

2-2. 마이크로그리드 도입의 필요성

마이크로그리드는 기존 전력망과 관련된 여러 문제에 대한 해결방안을 제공하며 다음과 같은 장점이 있다. (Wood, 2022)

- **복원력 향상**
  마이크로그리드는 주요 전력망에 대한 압력을 줄이고 극한의 날씨 또는 정전 시 백업 전력을 제공하여 중요한 부하에 대한 안정적인 전력 공급을 보장할 수 있다. 이러한 점은 주 전력망에 대한 접근이 제한된 지역이나 정전 또는 자연 재해가 발생하기 쉬운 지역에서 특히 유용하다.
- **효율성 향상**
  마이크로그리드는 지역 에너지 자원과 부하를 관리함으로써 공급과 수요의 균형을 효율적으로 맞추고, 송전 손실을 줄이며, 에너지 사용량을 조정할 수 있다. 피츠버그 국제공항의 경우 태양광 및 천연가스 마이크로그리드로 전환하여 첫해에 100만 달러의 비용을 절감하였다. 캘리포니아의 한 와이너리는 태양광 발전을 중심으로 마이크로그리드를 구축하여 월 에너지 요금을 15,000달러에서 1,000달러로 줄였다.
- **지속가능성**
  마이크로그리드는 재생 에너지원을 전력 시스템에 통합할 수 있도록 하여 전반적인 온실가스 배출을 줄이는데 기여할 수 있다. 태양광 발전, 풍력 발전 같은 재생 에너지의 특성상 변동성이 크다. 또한, 최근에는 가정집 창문, 학교 옥상등에 태양광 발전기를 부착함에 따라 용량 대비 관리해야 하는 갯수가 늘어나고 있다.. 이를 마이크로 그리드에서 관리하게 된다면 전체 그리드의 부담이 줄어들게 되어 재생 에너지를 전력 시스템에 통합하기 쉬워진다.

2-3. 마이크로그리드의 핵심 구성요소

마이크로그리드를 운영하기 위해서는 에너지 공급, 저장, 제어와 같이 분야에 따라 필요한 요소들이 있다. (McGrath and Gomstyn, 2024)

- **전원 공급원**
  마이크로그리드에 필요한 전기를 생산하는 발전기이다. 태양 전지판, 풍력 터빈 및 수력 발전 시스템과 같은 재생 가능한 자원뿐만 아니라 디젤 또는 천연 가스 발전기와 같은 자원이 포함될 수 있다.
- **에너지 저장 시스템**
  배터리 또는 플라이 휠과 같은 에너지 저장 장치로 마이크로그리드에서 생성된 초과 전력을 저장한다. 이렇게 저장된 에너지는 수요가 생산량을 초과할 때 또는 간헐적인 발전 기간(예: 태양광 발전의 경우 야간) 동안 사용할 수 있다. 전원 공급원과 에너지 저장 시스템을 통틀어 분산 에너지 자원(Distributed Energy Resources, DER)이라 한다.
- **분배 인프라**
  전원에서 부하로 전력을 분배하는 데 필요한 물리적 인프라(예: 전선, 변압기, 스위치)이다.
- **제어시스템**
  마이크로그리드의 '두뇌' 역할로 전력 공급의 균형을 맞추고, 재생 에너지원을 통합하며 에너지 저장을 관리한다. 또한 필요에 따라 마이크로그리드를 주요 전력망에서 분리했다가 다시 연결할 수 있다. 제어 시스템에는 전력 수요가 증가하거나 감소함에 따라 공급을 조정하는 부하 관리 도구와 전력 생산 및 소비를 측정하는 계량 장치가 포함되어 데이터를 제공한다.

- **부하 (Load)**
  주택, 상업시설, 공장 등 전기를 실제로 사용하는 수요처이다.

2-4. 중앙집중형 전력망의 한계를 보완하는 마이크로그리드

- **계통 분리(Island Mode) 운전**
  단일 장애점 문제는 계통 분리 운전을 통해 해결한다. 마이크로그리드는 상위 계통에 장애가 발생하는 순간 PCC(공통결합지점)를 통해 즉각 계통과 분리되어 독립 운전으로 전환한다. 내부의 분산에너지 자원을 활용해 자립적으로 전력을 공급하기 때문에, 광역 정전 상황에서도 해당 구역의 전력 공급은 중단 없이 유지된다.

- **ESS와 재생에너지의 협조 운전**
  재생에너지의 간헐성 문제는 ESS와 제어시스템의 협조 운전으로 대응한다. 태양광 발전이 왕성한 낮 시간대에 잉여 전력을 배터리에 저장해두었다가, 발전이 불가능한 야간이나 흐린 날에 방전하는 방식으로 수급 불균형을 능동적으로 완충한다. 이를 통해 기존 계통에서 재생에너지 연계 시 발생하던 전압·주파수 불안정 문제를 내부적으로 흡수할 수 있다.

- **소비지 인근 직접 발전·소비**
  원거리 송전으로 인한 손실 문제는 소비지 인근에서의 직접 발전·소비 구조로 해소한다. 발전원이 수요처와 지리적으로 가깝기 때문에 기존 중앙집중형 계통에서 불가피하게 발생하던 송전 손실이 대폭 줄어들며, 전력망 구축 비용이 과도하게 드는 도서·산간 지역에도 경제적인 전력 공급이 가능해진다.

- **EMS 기반 실시간 동적 부하 조절**
  수요 대응의 경직성은 EMS(에너지 관리 시스템)의 실시간 동적 부하 조절로 극복한다. EMS는 수요와 발전량 데이터를 실시간으로 수집·분석하여 ESS 충방전, 발전원 출력, 부하 우선순위를 유연하게 조정한다. 이로써 피크 수요에 맞춰 과잉 설비를 구축할 필요 없이, 주어진 자원 안에서 효율적인 에너지 운용이 가능해진다.

특히 PCC(공통결합지점)를 통해 평상시에는 기존 계통과 연계하여 안정성을 확보하고, 계통 이상 시에는 즉각 분리하여 독립 운전으로 전환하는 유연한 이중 운전 모드가 마이크로그리드의 핵심 강점이다. 이는 기존 전력망을 대체하는 것이 아니라, 분산된 지역 단위에서 에너지 회복력을 층층이 쌓아 올리는 방식으로 전체 계통의 신뢰성을 높이는 접근이다. (아이온커뮤니케이션즈, 2017)

3. 실제 사례 분석 - 켄텍 캠퍼스 마이크로그리드

3-1. 프로젝트 개요

KENTECH은 캠퍼스 내 태양광 발전 설비를 적극적으로 운영하고 있으나, 캠퍼스 내 전력 소비 패턴은 강의, 연구 스케줄과 기숙사 생활 패턴에 따라 태양광 발전 패턴과 상이한 양상을 보인다. 이를 해결하기 위한 대안으로 대형 고정형 ESS가 존재하지만, 높은 초기 투자 비용과 공간적 제약, 평상시 자산 가동률 저하라는 명확한 한계가 있다. V2G 기술 또한 대안으로 주목받고 있으나, 인프라 부족과 사용자의 심리적 장벽, 접속 불확실성으로 인해 캠퍼스 환경에 즉각 도입하기 어렵다는 현실적 제약이 있다. 이에 KENTECH 학생들이 일상적으로 사용하는 전기자전거의 배터리를 이동형 ESS로 활용하는 새로운 캠퍼스 마이크로그리드 운영 모델을 제안하였다. K-BEMS(Building Energy Management System) 데이터를 기반으로 전력 생산·소비 패턴을 분석하고, 전기자전거 배터리의 충·방전 최적 시기를 설계하여, 전력 부족 상황에서 이를 보조 전력원으로 활용하는 방안을 검증하는 것을 목적으로 프로젝트를 진행하였다. (오승지, 2026)

3-2. 연구 방법

본 연구의 전체 시스템은 K-BEMS가 15분 간격으로 수집하는 태양광 발전 및 부하 데이터를 에너지 관리 시스템(EMS)이 수집, 전처리한 뒤, EMS가 전기자전거 도킹스테이션(양방향 충·방전기), 파워뱅크(고정형 ESS), 태양광 인버터, 모사 부하에 제어 명령을 배분하는 구조로 설계되었다. 주요 설비 사양은 전기자전거 10대(배터리 288 Wh, 최대 충전 150 W, 최대 방전 450 W), 400 W 태양광 패널 6장(총 2.4 kW), 파워뱅크 4.0 kWh이며, 부하는 RC 교육생활관의 전력 수요를 축소·모사하였다. 데이터는 K-BEMS로부터 RC 생활관 부하와 캠퍼스 태양광 발전량을 15분 간격으로 수집하였으며, 결측 보간 및 이상치 제거 등의 전처리를 거쳐 실험 설비의 정격에 맞게 축소·스케일링하여 시뮬레이션에 활용하였다.



Figure 2. 켄텍 캠퍼스 마이크로그리드 구성도

3-3. 연구 결과

시뮬레이션 결과, 전기자전거 배터리와 파워뱅크의 협조 운전을 통해 낮 시간대의 잉여 태양광 전력을 효과적으로 흡수하고, 이른 아침과 저녁 피크 시간대의 전력 수요를 평탄화하는 데 기여하였다. 특히 본 결과는 전기자전거 배터리를 단순한 보조 배터리로 활용한 것이 아니라, 태양광 발전, 이동형 ESS, 고정형 ESS, 부하를 하나의 시스템으로 연계하고 EMS를 통해 충·방전을 제어한 캠퍼스 마이크로그리드 실증 연구라는 점에서 의미가 있다. 또한 대여 상황에서도 SoC 하한이 안정적으로 유지되어 사용자 편의성을 보장하면서 구매 전력 의존도를 낮출 수 있음이 확인되었으며, 이동형 ESS가 실제 마이크로그리드 운영 자원으로 기능할 수 있음을 시범적으로 입증하였다. 이를 통해 KENTECH 캠퍼스 사례는 기존 자원을 활용한 저비용·고활용도 기반의 마이크로그리드 운영 모델로서, 잉여 전력 흡수, 피크 감쇄, 구매 전력 저감을 동시에 달성할 수 있는 가능성을 보여주었다.

4. 마이크로그리드의 미래

마이크로그리드의 본격적인 상용화를 위해서는 기술, 경제, 제도 세 측면에서의 노력이 병행되어야 하며, KENTECH 사례는 이 세 축이 현장에서 어떻게 맞물리는지를 압축적으로 보여준다.

기술적으로는 이기종 분산 에너지 자원 간 표준화된 통신 프로토콜 확립과 계통 연계·분리 시 전압·주파수 안정화 기술의 고도화, 그리고 AI 기반 예측·제어 알고리즘의 실계통 적용 신뢰성 개선이 요구된다. KENTECH 사례에서 전기자전거 도킹스테이션, 파워뱅크, 태양광 인버터, 부하를 하나의 EMS로 묶어 제어한 구조는, 결국 서로 다른 기기들이 안정적으로 신호를 주고받을 수 있어야 비로소 성립한다. 실제로 이 팀은 인프라 구축 과정에서 기기 간 데이터 통신 보안의 중요성을 확인하고 후속 연구로 TLS 보안 프로토콜 최적화에까지 나아갔는데, 이는 분산자원을 엮는 일이 단순한 배터리 연결이 아니라 통신·제어·보안이 결합된 시스템 통합 문제임을 보여주는 단적인 예다.

경제적으로는 높은 초기 구축 비용이 여전히 소규모 커뮤니티의 진입 장벽으로 작용한다. KENTECH

사례의 의의는 바로 이 지점에서 드러난다. 학생들이 이미 매일 사용하는 전기자전거 배터리를 이동형 ESS로 재해석함으로써, 대규모 신규 투자 없이도 잉여 태양광 흡수와 피크 완화를 달성한 것이다. 이처럼 기존 자원의 가동률을 끌어올리는 저비용 접근법의 발굴은, 자본 여력이 부족한 캠퍼스나 중소 규모 커뮤니티가 마이크로그리드에 진입할 수 있는 현실적인 경로가 된다.
제도적으로는 마이크로그리드 간 전력 거래와 P2P 에너지 거래를 허용하는 법·제도 정비, 그리고 사용자의 자발적 참여를 이끌어낼 인센티브 체계 구축이 뒷받침되어야 한다. KENTECH 팀이 배터리 충전 상태(SoC) 하한선을 두어 이동 편의를 보장한 것은 기술적 안전장치인 동시에 참여 유도의 출발점이지만, 방전에 기여한 사용자에게 명확한 보상이 연결될 때 자발적 참여는 비로소 지속 가능해진다. 또한 전기자전거 배터리처럼 작고 이동하는 자원을 전력 시스템 안에서 정당한 자원으로 인정하고 보상하는 표준과 제도가 마련될 때, 개별 캠퍼스의 성공은 다른 공간으로 복제될 수 있다. 이러한 기술·경제·제도적 노력이 함께 이루어질 때, 마이크로그리드는 분산 에너지 시대의 핵심 인프라로 자리잡을 수 있을 것이다.

Conclusions

본 탐구를 통해 마이크로그리드는 기존 중앙집중형 전력망의 한계를 보완할 수 있는 현실적인 분산 에너지 시스템임을 확인하였다. 마이크로그리드는 단순한 백업 전원 시스템을 넘어 재생에너지의 간헐성을 보완하고 전력 수급을 유연하게 조절할 수 있는 능동적인 에너지 인프라로 활용될 수 있음을 보여주었다. 특히 KENTECH 캠퍼스 사례는 대규모 고정형 ESS를 추가로 구축하지 않고도 기존 전기자전거 배터리를 이동형 ESS로 활용하여 저비용 마이크로그리드 운영이 가능함을 보여주었다는 점에서 의미가 있다.

앞으로 마이크로그리드가 분산 에너지 시대의 핵심 인프라로 자리 잡기 위해서는 기술 표준화와 경제성 확보, 제도적 지원이 함께 이루어져야 한다. 이러한 기반이 마련된다면 마이크로그리드는 기존 중앙집중형 전력망을 보완하며, 더욱 안전하고 지속 가능한 분산형 전력 시스템 구축과 탄소중립 실현에 중요한 역할을 할 것으로 기대된다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

1. **학교 주변에 마이크로그리드를 도입한다면, 어떤 분산 에너지 자원을 조합하는 것이 가장 현실적까*?***

2. **마이크로그리드가 전력을 안정적으로 공급하려면 '제어 시스템'이 필수인데, *AI*가 이 역할을 맡는다면 어떤 데이터를 학습하고, 어떤 판단을 내려야 할까*?***

3. **마이크로그리드가 기술적으로 가능해도 사람들이 참여하지 않으면 작동하지 않는다. 사용자들의 자발적 참여를 끌어내기 위해 어떤 인센티브나 제도 설계가 필요할까*?***

4. **마이크로그리드끼리 서로 전력을 사고팔 수 있는 에너지 거래가 허용된다면, 에너지 가격은 어떻게 결정되어야 공정할까*?***

References

Asmus, P., A. Forni, and L. Vogel, 2018. Microgrid Analysis and Case Studies Report: California, North America, and Global Case Studies. Final Project Report CEC-500-2018-022, California Energy Commission. https://microgridresources.org/wp-content/uploads/2021/12/CEC-500-2018-022.pdf.

McGrath, A., and A. Gomstyn, 2024. 마이크로그리드란 무엇인가요? https://www.ibm.com/kr-ko/think/topics/microgrid.

Wood, E., 2022. How a Microgrid Saved Pittsburgh International Airport $1 Million. https://www.microgridknowledge.com/editors-choice/article/11427168/how-a-microgrid-saved-pittsburgh-international-airport-1-million.

아이온커뮤니케이션즈, 2017. [기술칼럼] 분산 에너지 자원 전력시장의 새로운 트렌드, Micro Grid(마이크로 그리드). https://www.i-on.net/news/newsletter/column/1214925_7965.html.

오승지. 2026. 켄텍 학부 1기, 전기자전거로 캠퍼스 마이크로그리드 실증. 전기신문 https://www.electimes.com/news/articleView.html?idxno=364582.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Jinoh Lee

M. jolee@kentech.ac.kr

Jin-oh Lee received the B.S. and Ph.D. degrees in Electrical and Computer Engineering from Seoul National University (SNU) in 2013 and 2019, respectively. He is currently an Assistant Professor in the Department of Energy Engineering at Korea Institute of Energy Technology (KENTECH). Prior to his current position, he worked as a Senior Researcher at the Korea Electrotechnology Research Institute (KERI) from 2020 to 2024, and served as an Assistant Professor at the KERI School, Korea National University of Science & Technology (UST) in 2024. Previously, he was a Postdoctoral Researcher in Electrical and Computer Engineering at Seoul National University (SNU) from 2019 to 2020. His research interests include power system design, operation, and control, inverter-based distributed energy resource (DER) control, energy management, grid restoration, smart grids, and microgrids. His research focuses on analyzing the static and dynamic states of power grids to derive optimal solutions for the stable transition to sustainable future power systems, ultimately contributing to tackling the climate crisis and achieving carbon neutrality.

KENTECH
Korea Institute of Energy Technology

Vol. 6 (2026)

The Future of Distributed Energy Resources, Microgrids

Abstract As the global trend toward carbon neutrality reveals the limitations of the existing centralized power system, microgrids based on distributed energy resources are drawing attention as the next generation energy infrastructure. This inquiry report organizes the definition and core components of microgrids and examines their practical potential through an actual case at the KENTECH (Korea Institute of Energy Technology) campus, where electric bicycle batteries are used as a mobile ESS (Energy Storage System). Furthermore, it presents the technical, economic, and institutional challenges that must be addressed for the full-scale commercialization of microgrids, and discusses the direction microgrids must take to establish themselves as core infrastructure in the era of distributed energy.

Introduction

The electricity we use in daily life is generated at power plants and supplied to homes, schools, and factories through transmission lines and substations. This electric power supply system, known as the grid, has long enabled stable electricity supply and has underpinned modern society. However, the increase in natural disasters caused by climate change, aging transmission infrastructure, and the expansion of renewable energy sources such as solar and wind power are gradually exposing the limitations of the existing centralized grid. Because the existing grid is structured to deliver electricity generated at large power plants over long distances, transmission losses occur, and a fault in a major facility can cause simultaneous blackouts across wide areas. In addition, as the proportion of renewable energy—whose output varies with weather conditions—increases, it becomes increasingly difficult to reliably balance electricity production and consumption using conventional methods alone. As an alternative solution to these problems, microgrids are drawing attention. A microgrid is a small-scale power network that generates, stores, and consumes electricity within a specific area; it normally operates in connection with the existing grid but can supply power independently when necessary. This inquiry report first examines the structure and limitations of the existing grid, then introduces the concept and core components of microgrids, and explores the potential applications and future development direction of microgrids through an actual case at the KENTECH campus.

1. The Existing Grid System

The power system currently operated in most countries follows a staged structure of generation, transmission, distribution, and consumption. In the generation stage, large centralized power plants—thermal, nuclear, hydroelectric, etc.—produce electricity on the scale of hundreds of MW to GW. The generated power is stepped up to ultra-high voltages of 154kV–765kV to improve transmission efficiency. In the transmission stage, the stepped-up power travels hundreds of kilometers to substations via transmission towers and high-voltage transmission lines. During this process, approximately 5–8% of total generated power is lost as heat due to resistive (Joule) losses. In the distribution stage, power is stepped down to about 22.9kV at substations and then supplied to end consumers—households, commercial facilities, factories—at 220V or 380V. In this structure, the flow of electricity is one-directional, from power plant to consumer, and consumers remain passive recipients of power.

1-1. Limitations of the Centralized Grid

While this one-directional, centralized structure achieves economies of scale in a stable environment, it reveals the following structural limitations in today's energy environment.

- **A structure in which problems can spread in a chain reaction**
  If a fault occurs at any one of the power plants, major substations, or key transmission lines, blackouts can cascade throughout the entire connected system. A representative example is the 2003 Northeast blackout in the United States and Canada, where a single transmission line fault triggered a chain reaction affecting 55 million people. As extreme weather events such as typhoons, heavy snow, and heat waves become more frequent due to climate change, this vulnerability is becoming even more pronounced.
- **Difficulty integrating renewable energy**
  Solar and wind power have intermittent output that fluctuates with weather conditions. Because the existing centralized grid is designed to precisely match supply and demand in real time, large-scale integration of unpredictable distributed generation can cause problems such as voltage and frequency instability and reverse power flow.
- **Losses and inefficiency due to long-distance transmission**
  The greater the physical distance between the power plant and the consumption area, the greater the transmission losses, and regions such as islands, mountainous areas, or developing countries that lack nearby generation sources face enormous costs simply to build grid infrastructure. Ulleungdo, an island in South Korea, is one such example. Located about 200km from the mainland, Ulleungdo generates its own power and supplies it through a 6.6kV distribution line.

2. Microgrids

2-1. Definition of a Microgrid

A microgrid refers to a localized power supply system centered on distributed generation that is independent from the existing wide-area power system, rather than relying on a centralized grid to unilaterally receive and consume electricity—in other words, a small-scale, "self-sufficient" power system. Additionally, the U.S. Department of Energy (DoE) defines a microgrid as a group of interconnected loads (power consumption) and distributed energy resources (power generation and storage) within clearly defined electrical boundaries.

The existing power system is structured to unilaterally receive power from the utility grid, so power supply is immediately interrupted in the event of an outage or grid fault. However, as shown in the figure above, when a microgrid is applied, internal distributed energy resources such as solar power, wind power, and energy storage systems can supply power independently even if the upper-level grid is disconnected, and surplus power can also be shared with neighboring microgrids. (Asmus et al., 2018)

2-2. The Need for Microgrid Adoption

Microgrids provide solutions to several problems associated with the existing grid and offer the following advantages. (Wood, 2022)

- **Improved resilience**
  Microgrids can reduce pressure on the main grid and provide backup power during extreme weather or



Figure 1. Composition and power flow diagram of a microgrid

outages, ensuring stable power supply to critical loads. This is particularly useful in regions with limited access to the main grid or areas prone to outages or natural disasters.

- **Improved efficiency**
  By managing local energy resources and loads, microgrids can efficiently balance supply and demand, reduce transmission losses, and adjust energy usage. Pittsburgh International Airport, for example, switched to a solar and natural gas microgrid and saved $1 million in costs in the first year. A winery in California built a microgrid centered on solar power generation, reducing its monthly energy bill from $15,000 to $1,000.

- **Sustainability**
  Microgrids enable the integration of renewable energy sources into the power system, helping to reduce overall greenhouse gas emissions. Renewable energy sources such as solar and wind power are highly variable by nature. In addition, as solar panels are increasingly attached to household windows and school rooftops, the number of units that must be managed relative to capacity is growing. If microgrids manage this, the burden on the overall grid is reduced, making it easier to integrate renewable energy into the power system.

2-3. Core Components of a Microgrid

Operating a microgrid requires various elements depending on the field, such as energy supply, storage, and control. (McGrath and Gomstyn, 2024)

- **Power sources**
  Generators that produce the electricity needed for the microgrid. These can include renewable resources such as solar panels, wind turbines, and hydroelectric systems, as well as resources such as diesel or natural gas generators.

- **Energy storage systems**
  Energy storage devices such as batteries or flywheels that store excess power generated by the microgrid. This stored energy can be used when demand exceeds generation or during periods of intermittent

generation (e.g., nighttime for solar power). Power sources and energy storage systems are collectively referred to as Distributed Energy Resources (DER).

- **Distribution infrastructure**
  The physical infrastructure (e.g., wires, transformers, switches) needed to distribute power from sources to loads.
- **Control system**
  Acting as the ”brain” of the microgrid, it balances power supply, integrates renewable energy sources, and manages energy storage. It can also disconnect the microgrid from the main grid and reconnect it as needed. The control system includes load management tools that adjust supply as power demand rises or falls, and metering devices that measure power production and consumption to provide data.
- **Load**
  The demand side—households, commercial facilities, factories, etc.—that actually consumes electricity.

2-4. Microgrids Complementing the Limitations of the Centralized Grid

- **Island mode operation**
  The single-point-of-failure problem is resolved through island mode operation. The moment a fault occurs in the upper-level grid, the microgrid immediately disconnects from the grid through the PCC (Point of Common Coupling) and switches to independent operation. Because it uses internal distributed energy resources to supply power on its own, power supply in that area continues without interruption even during a wide-area blackout.
- **Coordinated operation of ESS and renewable energy**
  The intermittency problem of renewable energy is addressed through coordinated operation of the ESS and the control system. Surplus power generated during peak solar hours in the daytime is stored in batteries and then discharged at night or on cloudy days when generation is not possible, actively buffering supply-demand imbalances. This allows the voltage and frequency instability problems that occur when renewable energy is integrated into the existing grid to be absorbed internally.
- **Direct generation and consumption near the point of use**
  The problem of losses from long-distance transmission is resolved through a structure of direct generation and consumption near the point of use. Because the generation source is geographically close to the demand area, the transmission losses that inevitably occur in the existing centralized grid are greatly reduced, and economical power supply becomes possible even in island or mountainous regions where grid construction costs would otherwise be prohibitive.
- **Real-time dynamic load control based on EMS**
  The rigidity of demand response is overcome through real-time dynamic load control by the EMS (Energy Management System). The EMS collects and analyzes demand and generation data in real time to flexibly adjust ESS charging/discharging, generation source output, and load priority. This enables efficient energy operation within available resources without the need to build excess facilities to meet peak demand.

In particular, the flexible dual operating mode—maintaining stability by connecting to the existing grid under normal conditions via the PCC (Point of Common Coupling), and immediately disconnecting to switch to independent operation in the event of a grid abnormality—is the core strength of microgrids. This is not an approach

that replaces the existing grid, but rather one that increases the reliability of the overall system by building up layers of energy resilience at the level of distributed local units. (아이온커뮤니케이션즈, 2017)

3. Case Study Analysis: The KENTECH Campus Microgrid

3-1. Project Overview

KENTECH actively operates solar power generation facilities on campus, but the campus's power consumption pattern differs from the solar generation pattern due to lecture and research schedules and dormitory living patterns. Large fixed ESS installations exist as an alternative solution to this problem, but they have clear limitations, including high initial investment costs, spatial constraints, and low asset utilization during normal operation. V2G (Vehicle-to-Grid) technology has also drawn attention as an alternative, but practical constraints such as insufficient infrastructure, psychological barriers among users, and connection uncertainty make it difficult to implement immediately in a campus environment. To address this, a new campus microgrid operating model was proposed that uses the batteries of electric bicycles, which KENTECH students use on a daily basis, as a mobile ESS. The project analyzed power generation and consumption patterns based on K-BEMS (Building Energy Management System) data, designed optimal charging and discharging timing for the electric bicycle batteries, and aimed to verify a method of using them as an auxiliary power source during periods of power shortage. (오승지, 2026)

3-2. Research Method

The overall system in this study was designed so that solar generation and load data collected by K-BEMS at 15-minute intervals are gathered and preprocessed by the Energy Management System (EMS), which then distributes control commands to the electric bicycle docking stations (bidirectional chargers/dischargers), power bank (fixed ESS), solar inverter, and simulated load. The main equipment specifications were: 10 electric bicycles (288 Wh battery, maximum charging 150 W, maximum discharging 450 W), six 400 W solar panels (total 2.4 kW), and a 4.0 kWh power bank; the load was a scaled-down simulation of the power demand of the RC (Residential College) education and living hall. Data on the RC dormitory load and campus solar generation were collected from K-BEMS at 15-minute intervals, and after preprocessing—including interpolation of missing values and removal of outliers—the data were scaled down to match the rated specifications of the experimental equipment for use in the simulation.



Figure 2. Configuration diagram of the KENTECH campus microgrid

3-3. Research Results

The simulation results showed that coordinated operation between the electric bicycle batteries and the power bank effectively absorbed surplus solar power during the daytime and contributed to flattening power demand during the early morning and evening peak periods. In particular, these results are significant in that the electric bicycle batteries were not used as mere auxiliary batteries, but rather as part of a campus microgrid demonstration in which solar generation, mobile ESS, fixed ESS, and load were linked into a single system with charging and discharging controlled by the EMS. It was also confirmed that even during battery rental, the lower SoC (State of Charge) limit was stably maintained, ensuring user convenience while reducing dependence on purchased power, demonstrating that a mobile ESS can function as an actual microgrid operating resource. Through this, the KENTECH campus case demonstrated the potential of a low-cost, high-utilization microgrid operating model using existing resources, capable of simultaneously achieving surplus power absorption, peak reduction, and reduced purchased power.

4. The Future of Microgrids

For the full-scale commercialization of microgrids, efforts must be made in parallel across three areas—technology, economics, and institutions—and the KENTECH case compactly illustrates how these three axes interact in the field.

On the technical side, what is needed is the establishment of standardized communication protocols among heterogeneous distributed energy resources, the advancement of voltage and frequency stabilization technologies for grid connection and disconnection, and improved reliability of AI-based prediction and control algorithms in actual grid applications. In the KENTECH case, the structure that integrated the electric bicycle docking stations, power bank, solar inverter, and load under a single EMS ultimately depends on the different devices being able to exchange signals reliably. In fact, the team recognized the importance of data communication security during infrastructure construction and extended their follow-up research to TLS security protocol optimization—a clear example showing that linking distributed resources is not simply a matter of connecting batteries, but a system integration challenge combining communication, control, and security.

On the economic side, high initial construction costs remain a barrier to entry for small communities. This is precisely where the significance of the KENTECH case lies. By reinterpreting electric bicycle batteries that students already use every day as a mobile ESS, the project achieved surplus solar absorption and peak mitigation without large-scale new investment. This kind of low-cost approach that raises the utilization rate of existing resources offers a realistic pathway for campuses or small communities with limited capital to adopt microgrids.

On the institutional side, what is needed is the establishment of laws and institutions that allow power trading between microgrids and P2P energy trading, along with incentive systems that encourage voluntary user participation. The KENTECH team's decision to set a lower SoC limit to ensure convenient mobility was both a technical safeguard and a starting point for encouraging participation, but voluntary participation can only become sustainable once clear compensation is linked to users who contribute through discharging. Furthermore, when standards and institutions are established that recognize and compensate small, mobile resources like electric bicycle batteries as legitimate resources within the power system, the success of an individual campus can be replicated elsewhere. When these technical, economic, and institutional efforts come together, microgrids will be able to establish themselves as core infrastructure in the era of distributed energy.

Conclusions

Through this inquiry, it was confirmed that microgrids are a realistic distributed energy system capable of complementing the limitations of the existing centralized grid. It was shown that microgrids can be used not merely as backup power systems but as active energy infrastructure capable of compensating for the intermittency of renewable energy and flexibly adjusting power supply and demand. In particular, the KENTECH campus case is significant in that it demonstrated that low-cost microgrid operation is possible by utilizing existing electric bicycle batteries as a mobile ESS, without the need to build additional large-scale fixed ESS.

Going forward, in order for microgrids to establish themselves as core infrastructure in the era of distributed energy, technical standardization, economic viability, and institutional support must be achieved together. Once this foundation is in place, microgrids are expected to complement the existing centralized grid and play an important role in building a safer, more sustainable distributed power system and achieving carbon neutrality.

Example Questions

Research may also be conducted on a free topic other than these questions.

1. If a microgrid were introduced near your school, which combination of distributed energy resources would be most realistic?

2. A "control system" is essential for a microgrid to supply power stably. If AI were to take on this role, what data should it learn from, and what decisions should it make?

3. Even if a microgrid is technically feasible, it will not function if people do not participate. What incentives or institutional designs are needed to encourage voluntary user participation?

4. If energy trading were allowed that lets microgrids buy and sell power to one another, how should energy prices be determined to ensure fairness?

References

Asmus, P., A. Forni, and L. Vogel, 2018. Microgrid Analysis and Case Studies Report: California, North America, and Global Case Studies. Final Project Report CEC-500-2018-022, California Energy Commission. https://microgridresources.org/wp-content/uploads/2021/12/CEC-500-2018-022.pdf.

McGrath, A., and A. Gomstyn, 2024. 마이크로그리드란 무엇인가요? https://www.ibm.com/kr-ko/think/topics/microgrid.

Wood, E., 2022. How a Microgrid Saved Pittsburgh International Airport $1 Million. https://www.microgridknowledge.com/editors-choice/article/11427168/how-a-microgrid-saved-pittsburgh-international-airport-1-million.

아이온커뮤니케이션즈, 2017. [기술칼럼] 분산 에너지 자원 전력시장의 새로운 트렌드, Micro Grid(마이크로 그리드). https://www.i-on.net/news/newsletter/column/1214925_7965.html.

오승지. 2026. 켄텍 학부 1기, 전기자전거로 캠퍼스 마이크로그리드 실증. 전기신문 https://www.electimes.com/news/articleView.html?idxno=364582.

Contributors



Mincheol Kim

M. kimmc050930@kentech.ac.kr

I am currently studying at KENTECH especially power semiconductor. My goal is to become a professor dedicated to advancing research in the field of power semiconductors.



Sumin Lee

M. jetson@kentech.ac.kr

Undergraduate student at KENTECH with interests in artificial intelligence, robotics, and energy policy. Belonging Prof SJ Lee's VIEW lab as undergraduated researcher. I am researching a robot that can explore for long periods of time by efficiently tracking solar energy.



Suhyeon Oh

M. suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Jinoh Lee

M. jolee@kentech.ac.kr

Jin-oh Lee received the B.S. and Ph.D. degrees in Electrical and Computer Engineering from Seoul National University (SNU) in 2013 and 2019, respectively. He is currently an Assistant Professor in the Department of Energy Engineering at Korea Institute of Energy Technology (KENTECH). Prior to his current position, he worked as a Senior Researcher at the Korea Electrotechnology Research Institute (KERI) from 2020 to 2024, and served as an Assistant Professor at the KERI School, Korea National University of Science & Technology (UST) in 2024. Previously, he was a Postdoctoral Researcher in Electrical and Computer Engineering at Seoul National University (SNU) from 2019 to 2020. His research interests include power system design, operation, and control, inverter-based distributed energy resource (DER) control, energy management, grid restoration, smart grids, and microgrids. His research focuses on analyzing the static and dynamic states of power grids to derive optimal solutions for the stable transition to sustainable future power systems, ultimately contributing to tackling the climate crisis and achieving carbon neutrality.

KENTECH
Korea Institute of Energy Technology

X

우주용 전력원으로서의 페로브스카이트 태양전지 :

소재 특성과 방사선 내구성

p. 222 - 275

Vol. 07 (2026)

우주용 전력원으로서의 페로브스카이트 태양전지: 소재 특성과 방사선 내구성

Abstract

인공지능과 클라우드 서비스의 확산으로 데이터 처리 수요가 급증하면서, 지상 데이터센터의 전력난과 냉각 문제를 극복하기 위해 연산 설비 일부를 우주 궤도로 옮기려는 시도가 나타나고 있다. 궤도 상의 설비는 지상 전력망을 쓸 수 없어 현지에서 직접 전력을 생산해야 하며, 이에 따라 태양전지가 핵심 전원으로 주목받고 있다. 그러나 현재 우주에서 쓰이는 III-V족 다중접합(Multijunction) 전지는 비싸고 무거운 단결정 기판 위에서만 제조할 수 있어, 수천 기 규모의 위성군이나 대면적 어레이를 요구하는 새로운 임무에는 맞지 않는다.

본 탐구 자료는 그 대안으로 거론되는 페로브스카이트(Perovskite)를 소개한다. 먼저 우주 태양전지가 만족해야 할 조건이 지상과 어떻게 다른지 정리하고, 현재 소재의 한계가 어디에서 비롯되는지를 분석한다. 이어 페로브스카이트의 구조적 유래와 조성 설계 원리, 뛰어난 광전 특성과 결함 수용성(Defect tolerance)을 차례로 살펴본 뒤, 이 결함 수용성이 어떻게 방사선 내구성과 '자가 회복(Self-healing)'으로 이어지는지를 설명한다. 나아가 성층권 기구부터 국제우주정거장 장기 노출까지 이어진 궤도 실증 사례를 검토하고, 실용화를 위해 소재·검증·시스템 세 측면에서 해결해야 할 과제를 제시한다.

서론

생성형 AI 모델의 학습과 추론, 클라우드 서비스의 확산으로 지상의 데이터 처리 수요와 전력 소비가 폭증하고 있다. 대형 데이터센터의 전력과 냉각 수요는 수백 MW에 달하지만, 지상에서는 계통 접속 지연, 냉각수 부족, 부지 규제로 증설에 한계를 맞이했다. 이러한 문제를 극복하기 위한 대안으로 풍부한 태양광과 진공 환경을 활용하는 '우주 궤도 연산 설비(궤도 데이터센터)' 구상이 등장했다.

궤도 연산 설비가 최근 다시 논의되는 이유는 지상의 제약이 커진 동시에 궤도의 이점이 수치로 계산 가능해졌기 때문이며, 이는 네 가지로 정리된다.

첫째, 태양광의 가용성이다. 지상과 달리 궤도에는 밤, 구름, 대기 감쇄가 없다. 특히 새벽-황혼 태양동기 궤도(Sun-synchronous orbit, 지구의 밤낮 경계선을 따라 돌며 지구 그림자에 들어가지 않는 궤도)를 쓰면 위성이 끊이지 않고 빛을 받아 지상보다 최대 8배 많은 전력을 얻을 수 있다 (**?**).

둘째, 냉각 방식이 다르다. 진공에서는 대류와 전도가 없어 열을 복사로만 버려야 하므로 방열판 면적이

커지지만, 냉각수가 전혀 필요 없다. 즉, 궤도의 열 문제는 '물의 문제'에서 태양전지와 동일한 '면적과 질량의 문제'로 전환된다.

셋째, 부지와 규제에서 자유롭다. 토지 매입이나 주민 동의, 계통 접속 대기 등 지상의 증설 걸림돌이 존재하지 않는다.

넷째, 발사 비용이 내려갔다. 재사용 발사체 보급으로 kg당 수송 단가가 낮아져 대량 수송이 현실화되었다.

실제 움직임도 나타나고 있다. 이미 저궤도(Low Earth Orbit, LEO) 통신 위성군이 수천 기 규모로 운용 중이며 (?), 2025년 11월 구글은 태양동기궤도에 위성 군집을 배치해 연산용 반도체를 올리는 연구 계획을 공개했다 (?). 다만 열 관리, 방사선 내성 등 미검증 항목이 많아 현 단계는 개념 검증에 가깝다.

궤도 인프라는 전력을 현지에서 직접 확보해야 한다. RTG(Radioisotope Thermoelectric Generator)나 소형 원자로는 출력 한계(수백 W)나 안전 검증 문제로 한계가 있어, MW GW급의 대규모 연산 설비 전력을 지속 공급하려면 태양전지가 유일한 대안이다 (?). 저궤도 위성군 역시 수천 기를 발사해야 하므로 발사체 페어링(Fairing) 부피와 질량 제약이 함께 걸린다 (?). 그러나 기존 우주용 태양전지는 단일 위성 장기 운용에 맞춰져 있어 이러한 대규모 임무에는 적합하지 않으므로, 핵심 소재 선택부터 다시 검토해야 한다.

본 탐구 자료에서는 먼저 우주 태양전지가 만족해야 하는 조건을 정리한 뒤(1장), 현재 쓰이는 III-V족 다중접합 전지의 한계가 어디에서 비롯되는지 분석한다(2장). 이어 대안으로 거론되는 페로브스카이트가 어떤 물질이며 왜 방사선에 강한지를 살펴보고(3장), 실제 궤도에서 수행된 실증 시험 사례를 검토한 뒤(4장), 실용화를 위해 남은 과제를 소재·검증·시스템 세 축으로 정리한다(5장).

1. 우주에서 전기를 만든다는 것

1.1 태양광 조건: AM0

태양전지의 성능을 평가하고 작동 환경을 이해하려면 가장 먼저 광원(빛) 조건을 파악해야 한다. 지상이든 우주든 태양광이 도달하는 환경의 차이를 나타내는 공통 기준이 바로 '대기질량(Air Mass, AM)'이다.

AM은 태양빛이 지구 대기층을 얼마나 길게 뚫고 지나왔는지를 나타내는 비율이다. 태양이 머리 바로 위에 똑바로 떠 있을 때 대기를 통과하는 길이를 기준인 1(AM1)로 두고, 태양이 기울어질수록 길어지는 경로를 상대적인 수치로 나타낸다. 태양이 기울어진 각도(천정각(Zenith angle) θ)를 이용하면 다음과 같이 간단히 계산할 수 있다.

$$\mathrm{AM} = \frac{1}{\cos\theta}$$

우리가 지상용 태양전지 스펙에서 흔히 접하는 "효율 몇 퍼센트"라는 수치는 모두 지상 표준 조건인 AM1.5G에서 측정한 결과다. 이는 중위도 지역의 연평균 태양 위치(약 48.2° 기울어진 각도)를 기준으로,

햇빛의 세기(조사도(Irradiance))를 1000 W/m^2로 맞춰놓고 평가하는 방식이다. 반면, 우주는 대기를 전혀 지나지 않으므로 경로 길이가 0이라는 뜻에서 AM0를 표준 조건으로 사용한다. 대기막에 가려지지 않은 우주 공간에서는 햇빛의 세기가 약 1361–1366 W/m^2 수준으로 지상보다 훨씬 강하다 (**?**). 그림 1처럼 태양이 바로 위에 있으면 AM1, 태양이 비스듬히 기울어질수록 대기를 지나야 하는 기하학적 경로가 길어져 AM1.5, AM2로 숫자가 커지며, 대기가 아예 없는 우주 궤도 환경이 곧 AM0가 된다.

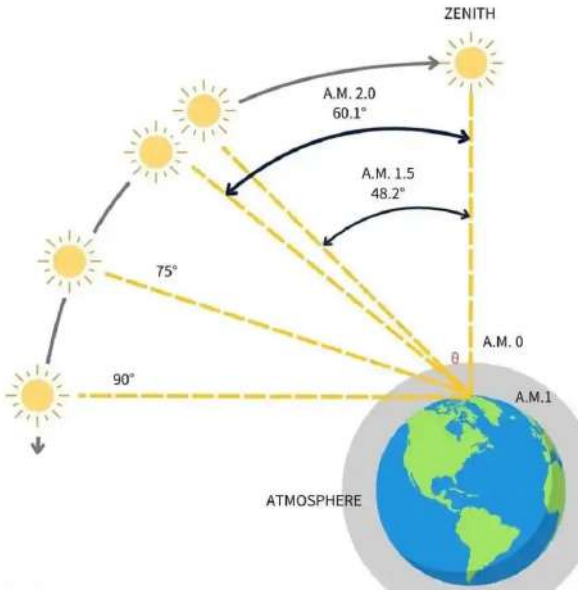


Figure 1. 태양 천정각에 따른 대기질량의 정의

두 조건의 차이는 햇빛의 세기뿐만 아니라 빛의 성분(스펙트럼(Spectrum)의 모양)에서도 크게 나타난다. 지상에서는 대기 중의 오존이 자외선(Ultraviolet, UV)을 막아주고 수증기와 이산화탄소가 근적외선(Near-infrared, NIR)을 흡수하지만, 대기가 없는 AM0 환경에는 이러한 차단막이 없다. 두 스펙트럼을 겹쳐 놓은 그림 2에서 이 차이를 확인할 수 있으며, 이는 우주용 태양전지 설계에 다음 세 가지 결정적인 영향을 미친다.

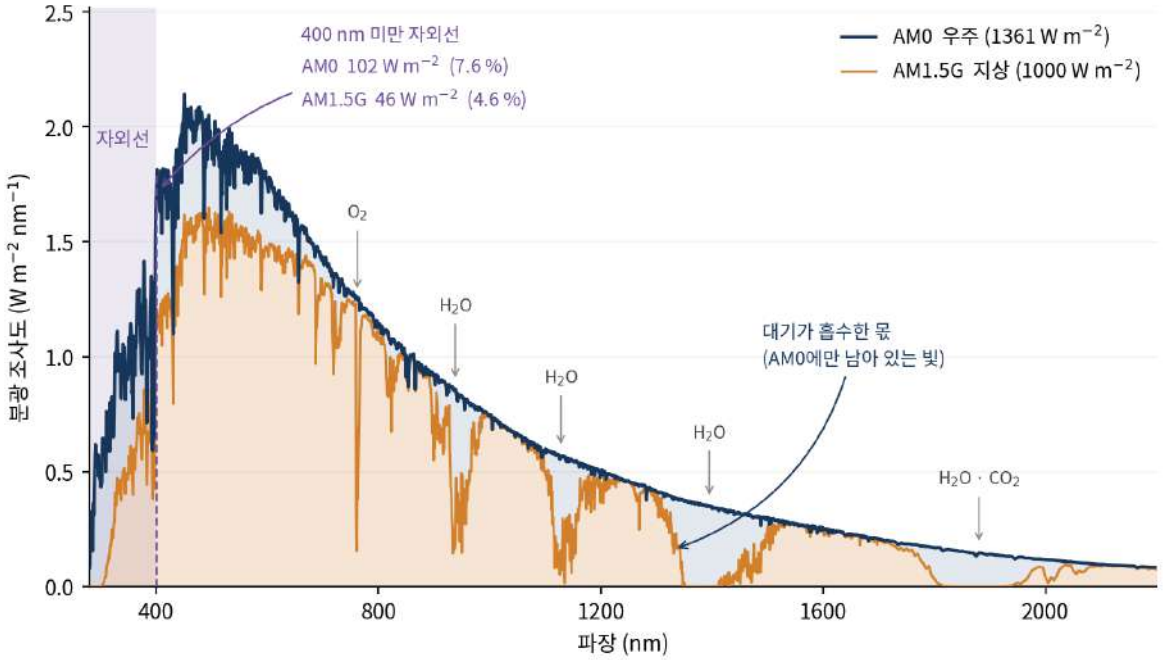


Figure 2. AM0와 AM1.5G 태양 스펙트럼의 비교 (**?**)

첫째, 자외선 비중이 높아져 소재가 쉽게 손상된다. 우주에서는 에너지가 매우 강한 자외선(파장 400 nm 미만)의 비율이 눈에 띄게 높다. 자외선 광자(Photon)는 에너지가 커서 태양전지 내부의 유기물 화학 결합을 끊어버리므로, 지상에서보다 유기 전하수송층과 보호막의 열화가 훨씬 빠르게 진행된다.

둘째, 지상 대기에서 깎여 나갔던 근적외선 영역까지 모두 활용할 수 있다. 지상(AM1.5G) 스펙트럼에서는 대기 중 수증기 때문에 파여 있던 근적외선 대역이 AM0에서는 채워져 있다. 따라서 좁은 밴드갭을 가진 흡수체 소자가 흡수할 수 있는 빛(광자)의 절대적인 양이 늘어난다.

셋째, 지상에서 최고 성능을 내는 소재가 우주에서도 최고라는 보장이 없다 (**?**). 빛의 성분이 달라지므로 태양전지가 빛을 가장 잘 흡수할 수 있는 **'최적 밴드갭(빛 흡수 에너지 대역)'**의 위치 자체가 이동한다. 실제로 같은 태양전지를 지상과 우주 조건에서 측정하면 우주(AM0)에서의 효율이 대체로 낮게 측정된다. 따라서 지상용 효율 수치를 그대로 우주 성능으로 간주해서는 안 되며, 우주용 소자는 반드시 AM0 기준에 맞춰 빛 흡수 대역을 새롭게 설계하고 평가해야 한다.

1.2 우주 태양전지의 네 가지 요구 조건

광원 조건 외에도 우주용 태양전지가 만족해야 할 핵심 항목들이 있는데, 이는 지상용 태양전지에서는 거의 고려하지 않는 조건들이다.

방사선 내구성. 궤도 환경에는 지구 자기장에 포획된 고에너지 양성자(Proton)와 전자(Electron)가 존재하는 '밴앨런대(Van Allen belt)' 영역이 있으며, 태양 입자 폭발(Solar Particle Event, SPE)과 우주선(Cosmic

ray)이 지속적으로 쏟아진다. 이들이 반도체 격자에 입사하면 원자를 제자리에서 밀어내 결함(Defect)을 만든다. 이 손상은 시간에 따라 축적되며 대부분의 반도체에서 되돌릴 수 없으므로 (**?**), 우주태양전지의 실제 수명은 단순 초기 효율이 아니라 '열화 속도(Degradation rate)'로 결정되는 경우가 많다.

출력 대 질량비. 단위 질량당 얼마의 전력을 내는지를 나타내는 지표로, 발사체에 실을 수 있는 질량이 엄격히 제한되는 우주 임무에서 매우 중요하게 다루어진다. 주의할 점은 분모에 전지 내부의 흡수층만 들어가는 것이 아니라 기판, 봉지재, 전극 배선, 전개 구조체의 무게가 모두 포함된다는 사실이다. 아무리 빛 흡수층을 얇게 만들어도 지지하는 기판이 무거우면 의미가 없다.

수납 부피와 전개성. 위성을 다수 배치하는 임무에서는 질량보다 부피가 실제 제약 요소가 되는 경우가 많다. 즉, 발사체 페어링(탑재체 보호 덮개) 안의 제한된 공간에 얼마나 작고 밀도 높게 접어 넣을 수 있는지가 핵심이다. 작게 접거나 말아서 궤도에 올린 뒤 펼칠 수 있는 유연한 구조가 유리한 이유다.

열 사이클 내성. 저궤도 위성은 약 90분마다 햇빛을 받는 일조 구간과 지구 그림자에 가려지는 식(Eclipse) 구간을 반복한다. 이 과정에서 소자 온도가 대략 −150 °C에서 +120 °C 사이를 극심하게 오가는데, 하루에 16번, 1년이면 5,000번이 넘는 충격이 가해진다. 박막을 구성하는 각 층마다 열팽창계수(Coefficient of Thermal Expansion, CTE)가 다르므로 계면에 물리적 응력이 반복적으로 걸리고, 이것이 쌓이면 층이 벗겨져 소자가 파괴된다.

이 네 항목은 서로 독립적이지 않으며 팽팽한 상충 관계를 이룬다. 방사선을 막으려고 두꺼운 보호층을 두면 출력 대 질량비가 떨어지고, 유연성과 수납성을 얻으려고 얇은 고분자 기판을 쓰면 원자상 산소(Atomic oxygen)나 자외선 공격에 취약해진다. 우주태양전지 설계는 결국 이러한 요소들 간의 최적 균형점을 찾는 작업이며, 이는 5장에서 다시 다룬다.

2. 현재의 우주태양전지와 한계

2.1 III-V족 다중접합 태양전지

현재 궤도에서 쓰이는 태양전지는 대부분 III-V족 화합물 다중접합 전지로, 주기율표 13족과 15족 원소로 만든 화합물 반도체를 여러 층 쌓은 구조다. 대표적인 조합은 InGaP, InGaAs, Ge 세 층이다.

여러 층을 쌓는 이유는 열 손실을 줄이기 위해서다. 단일 흡수층으로 태양광을 받으면 밴드갭보다 에너지가 작은 광자는 그냥 통과하고, 밴드갭보다 큰 광자는 흡수되더라도 밴드갭을 초과하는 남은 에너지가 열로 버려진다. 밴드갭이 다른 층을 겹쳐 쌓으면 위층이 파장이 짧은 고에너지 빛을, 아래층이 파장이 긴 저에너지 빛을 나눠 맡아 효율적으로 흡수하게 되며, 이 방식으로 AM0 조건에서 30%가 넘는 높은 효율을 얻는다.

신뢰성 자료도 충분하다. 궤도 운용 실적이 수십 년에 이르고 방사선에 따른 열화를 예측하는 모형도 정립되어 있어 (**?**), 임무 수명 동안 출력이 얼마나 줄어들지 설계 단계에서 정량적으로 계산할 수 있다.

우주 임무는 한번 발사하면 고치러 갈 수 없으므로 이 점은 결코 사소하지 않으며, 새로운 소재가 이 자리를 대체하려면 성능뿐만 아니라 이러한 예측 가능성도 함께 확보해야 한다.

2.2 단결정 기판 하나가 세 가지를 묶는다

III-V족 전지의 제약은 흔히 고단가, 고중량, 유연성 부족(강성) 세 가지로 요약된다. 그런데 이 셋은 별개의 결점이 아니라, 이 전지가 단결정 기판 위에서만 만들어져야 한다는 하나의 원인에서 파생된 결과다.

먼저 왜 단결정 기판이 필요한지 살펴보자. 다중접합 전지는 성질이 다른 반도체를 층층이 쌓는데, 이때 각 층의 원자 간격(격자 상수(Lattice constant))이 아래층과 맞지 않으면 계면에 전위(Dislocation)가 생긴다. 전위는 결정 배열이 어긋난 선 모양의 결함으로, 생성된 전하를 붙잡아 소멸시키는 통로가 된다. 결함 없이 격자를 맞추며 결정을 올려 쌓으려면 원자 배열이 완벽히 정렬된 Ge 또는 GaAs 단결정 웨이퍼가 필수적이며, 성장 방식 역시 유기금속 화학기상증착(Metal-Organic Chemical Vapor Deposition, MOCVD)이나 분자선 에피택시(Molecular Beam Epitaxy, MBE)처럼 고온·고진공 환경을 쓰는 정밀 장비를 동원해야 한다.

고단가. 단결정 웨이퍼 자체가 비싼 데다 Ge는 매장량이 적고 정제가 어렵다. 여기에 성장 속도가 느리고 원료 이용률이 낮으며 장비가 고가인 에피택셜 성장(Epitaxial growth) 공정이 더해지므로, 단위 면적당 제조비가 지상용 실리콘 전지보다 두세 자릿수 높아진다 (**?**). 위성 몇 기를 만들 때는 감당할 수 있지만, 수천 기 규모의 인프라를 구축할 때는 이야기가 달라진다.

고중량. 웨이퍼는 고온의 공정 도중 휘거나 깨지지 않아야 하므로 두께를 150 μm 안팎으로 유지해야 하는데, 밀도가 5.32 $\mathrm{g/cm^3}$인 Ge 기판의 면적 질량은 약 800 $\mathrm{g/m^2}$에 달한다. 그런데 빛을 실제로 흡수해 전기를 만드는 층의 두께는 다 합쳐도 수 μm에 불과하므로, 질량의 대부분이 발전에 기여하지 않는 단순 지지체에 불과한 셈이다. 그림 3은 이 구조를 페로브스카이트 소자와 나란히 비교한 것이다. 공정 후 기판을 떼어내고 재사용하는 기술(Epitaxial lift-off)도 연구되고 있으나, 추가 공정으로 인해 단가가 더 올라가는 문제가 발생한다.

유연성 부족. Ge와 GaAs는 힘을 받으면 변형되지 않고 깨져버리는 취성(Brittle) 재료인 데다, 두께가 두꺼울수록 구부릴 때 표면에 걸리는 변형률이 커진다. 결과적으로 잘 구부러지지 않아 접거나 말 수 없다. 실제 위성에서는 평평한 판을 경첩으로 연결해 병풍처럼 접는 방식을 쓰는데, 이는 발사체 페어링 내부 공간을 많이 차지할 뿐만 아니라 전개 기구 자체가 고장 위험이 있는 복잡한 가동부가 된다.

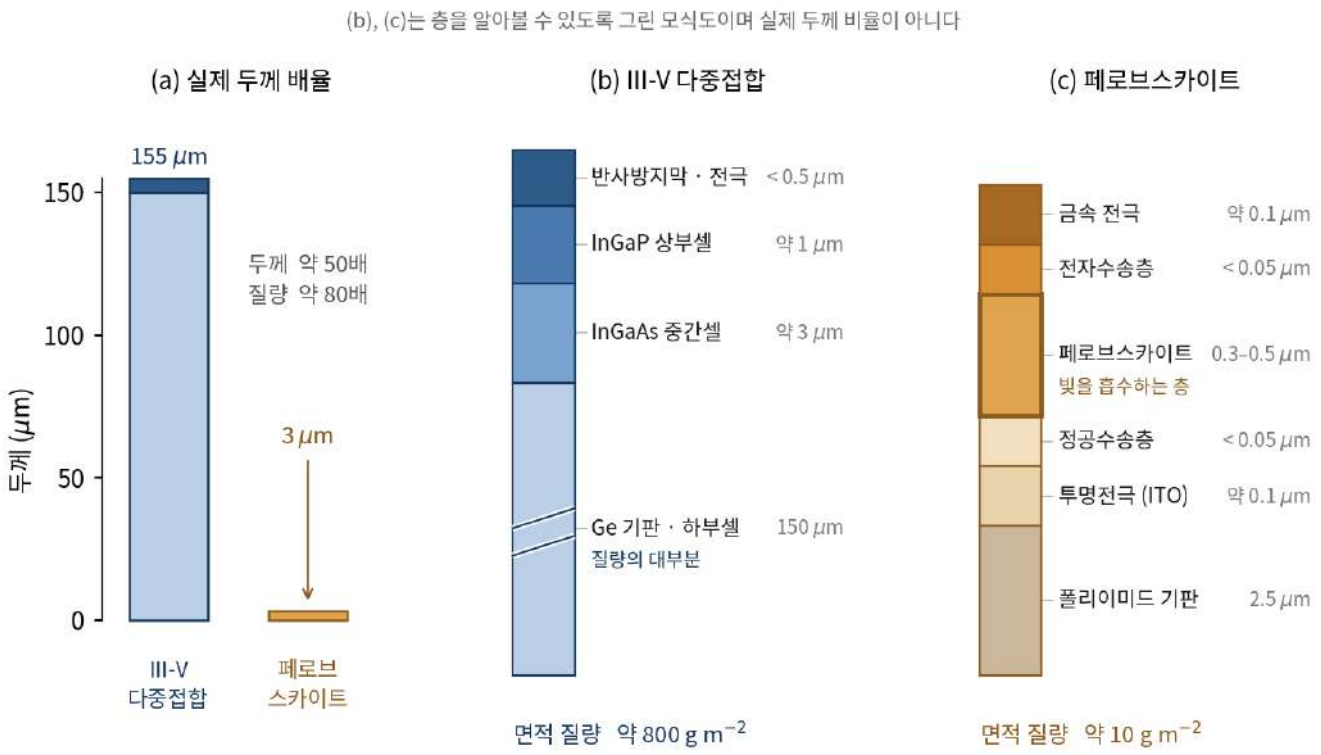


Figure 3. III-V족 다중접합 전지와 페로브스카이트 전지의 단면 구조 비교. (a) 두 소자의 실제 두께 배율, (b) III-V족 층 구성, (c) 페로브스카이트 층 구성

정리하면, 단결정 기판을 써야 한다는 조건 하나가 단가와 질량, 형상(유연성) 세 가지를 동시에 묶고 있다. 따라서 이 한계를 극복하려면 단순한 공정 개선이 아니라 아예 다른 방식의 소재를 도입해야 한다. 주요 후보 소재들의 특성을 표 1에 정리했다.

Table 1. 우주용 태양전지 후보의 특성 비교 (**?**)

항목	III-V 다중접합	결정질 Si	페로브스카이트
AM0 효율	30% 이상	15–18%	개발 단계
흡수층 두께	수 μm	150 μm 이상	0.3–0.5 μm
기판	Ge/GaAs 단결정	Si 웨이퍼	폴리이미드 등
출력 대 질량비	약 5.5 W/g	0.5–2 W/g	23–30 W/g
유연성	없음	제한적	있음
제조 온도	600 °C 이상	1400 °C 이상	150 °C 이하
우주 검증	충분	일부	초기 단계

2.3 여러 대안 중 페로브스카이트가 주목받는 이유

단결정 기판 없이 유연하게 만들 수 있는 박막(Thin film) 태양전지 계열에는 페로브스카이트 외에도 여러 대안 후보가 존재한다. 그렇다면 왜 유독 페로브스카이트가 차세대 우주용 소재로 강력하게 거론되는가. 기존 박막 후보들을 차례로 살펴보면 그 이유가 분명해진다.

CIGS. 구리, 인듐, 갈륨, 셀레늄으로 만든 화합물 반도체로, 유연 기판 제작이 가능하고 방사선 내구성이 우수해 실제로 우주 적용이 검토된 바 있다. 하지만 효율 상승이 오래전부터 정체되어 있고, 고진공 증착 장비가 필요해 제조 단가 절감 폭이 크지 않다. 또한 희유금속인 인듐의 매장량이 제한적이라는 원료 공급상의 한계도 안고 있다.

CdTe. 지상 상용화에 성공한 대표적 박막 전지이지만, 대부분 두꺼운 유리 기판 위에서 제작되므로 유연화가 쉽지 않다. 여기에 독성 물질인 카드뮴 사용에 따른 환경 규제도 우주용 대량 생산 시 부담으로 작용한다.

유기 태양전지(Organic solar cell). 매우 가볍고 유연하여 질량 측면에서는 가장 유리하다. 그러나 절대적인 광전 변환 효율(Power Conversion Efficiency, PCE)이 낮고 자외선에 치명적으로 약하다. 앞서 보았듯 AM0 조건에서는 자외선 비중이 매우 크기 때문에 우주 환경에 치명적인 약점을 가진다.

페로브스카이트. 위 후보들이 한두 가지씩만 만족시키던 요구 조건을 다음과 같이 동시에 충족한다. 첫째, 효율이 이미 실용 수준에 도달했다. 단일접합 인증 효율이 27%대, 실리콘과 결합한 탠덤(Tandem) 구조는 34%대에 이르며, 이는 기존의 박막 전지들이 도달하지 못한 성과다. 둘째, 흡수계수(Absorption coefficient)가 매우 높아 300 nm 남짓의 얇은 두께로도 태양광을 충분히 흡수한다. 이 정도 두께에서는 단결정 기판 없이 수 μm 두께의 얇은 고분자 필름만으로도 구조를 충분히 지탱할 수 있다. 출력 대 질량비가 기존 III-V족 전지보다 한 자릿수 이상 높게 보고되는 근거가 바로 여기에 있다 (**?**). 셋째, 150°C 이하의 저온 용액 공정(Solution process)으로 제조가 가능하다. 열에 약한 유연 고분자 기판을 그대로 사용할 수 있으며, 고가의 고진공 장비가 필수가 아니므로 제조 단가를 획기적으로 낮출 수 있다. 넷째, 방사선 손상에 대한 구조적 고유 내성과 자가 회복 특성을 가진다. 이는 CIGS 등의 단순 방사선 내성과는 달리, 페로브스카이트 고유의 특유한 전자 구조에서 비롯되는 성질이며 3장에서 집중적으로 다룬다.

한 가지 덧붙이면, 페로브스카이트에 ”차세대”라는 기대가 모이는 진짜 이유는 어느 한 지표만 뛰어난 것이 아니라 효율, 질량, 단가, 방사선 내성이라는 상충하는 조건들을 동시에 만족시킬 잠재력을 보여주기 때문이다.

3. 페로브스카이트란 무엇인가

3.1 물질의 이름이 아니라 구조의 이름

먼저 분명히 해 둘 점이 있다. 페로브스카이트는 특정 화합물의 이름이 아니라 결정 구조(Crystal structure)의 이름이며, 어떤 원소로 만들었든 원자가 특정 방식으로 배열되어 있으면 페로브스카이트라 부른다. 이 개념을 명확히 하지 않으면 이후의 논의가 헷갈릴 수 있다.

이름의 유래는 광물이다. 1839년 독일 광물학자 구스타프 로제(Gustav Rose)가 러시아 우랄산맥에서 화학식 $CaTiO_3$의 티탄산칼슘 광물을 발견하고, 러시아 광물학자 레프 페롭스키(Lev Perovski)의 이름을 따서 이를 '페로브스카이트'라 명명했다. 이후 같은 배열을 갖는 화합물이 계속 발견되면서 용어의 쓰임이 넓어졌다.

지금은 $CaTiO_3$와 같은 ABX_3 형태의 결정 배열을 갖는 물질군 전체를 페로브스카이트라 부르며, 원래의 광물은 그중 하나일 뿐이다. 이 구조는 원래 X 자리에 산소가 들어간 산화물 분야에서 오랫동안 연구되어 왔다. $BaTiO_3$는 대표적인 강유전체(Ferroelectric)로 쓰이고 일부 구리 산화물은 고온 초전도체 특성을 보이므로, 페로브스카이트 구조는 태양전지를 위해 새로 발굴된 것이 아니라 재료과학계에 이미 매우 익숙한 구조였다.

태양전지에 쓰이는 것은 X 자리에 산소 대신 할로겐 이온(I, Br, Cl)이 들어간 '할라이드 페로브스카이트(Halide perovskite)'다. 이 계열의 유기-무기 혼합 화합물은 1978년 베버(Dieter Weber)가 처음 합성해 보고했는데 (**?**), 당시에는 태양전지와 무관한 순수 결정학 연구였다. 이 재료가 태양전지로 처음 쓰인 것은 2009년이다. 미야사카(Tsutomu Miyasaka) 연구진이 염료감응형 태양전지(Dye-Sensitized Solar Cell, DSSC)의 감광체로 적용해 3.8%의 효율을 보고한 것이 시작이었다 (**?**). 이후 소자 구조가 비약적으로 개선되면서, 2012년 고체 정공수송층(Hole Transport Layer, HTL) 도입으로 10%를 돌파했고, 2026년 기준 단일접합 인증 효율은 27%대, 실리콘과 결합한 탠덤 구조는 34%대에 도달했다 (**?**). 기존 결정질 실리콘이 같은 효율을 달성하는 데 걸린 기간의 3분의 1도 되지 않는 성장 속도이며, 주요 소재별 연도별 인증 효율 추이를 그림 4에 나타냈다.

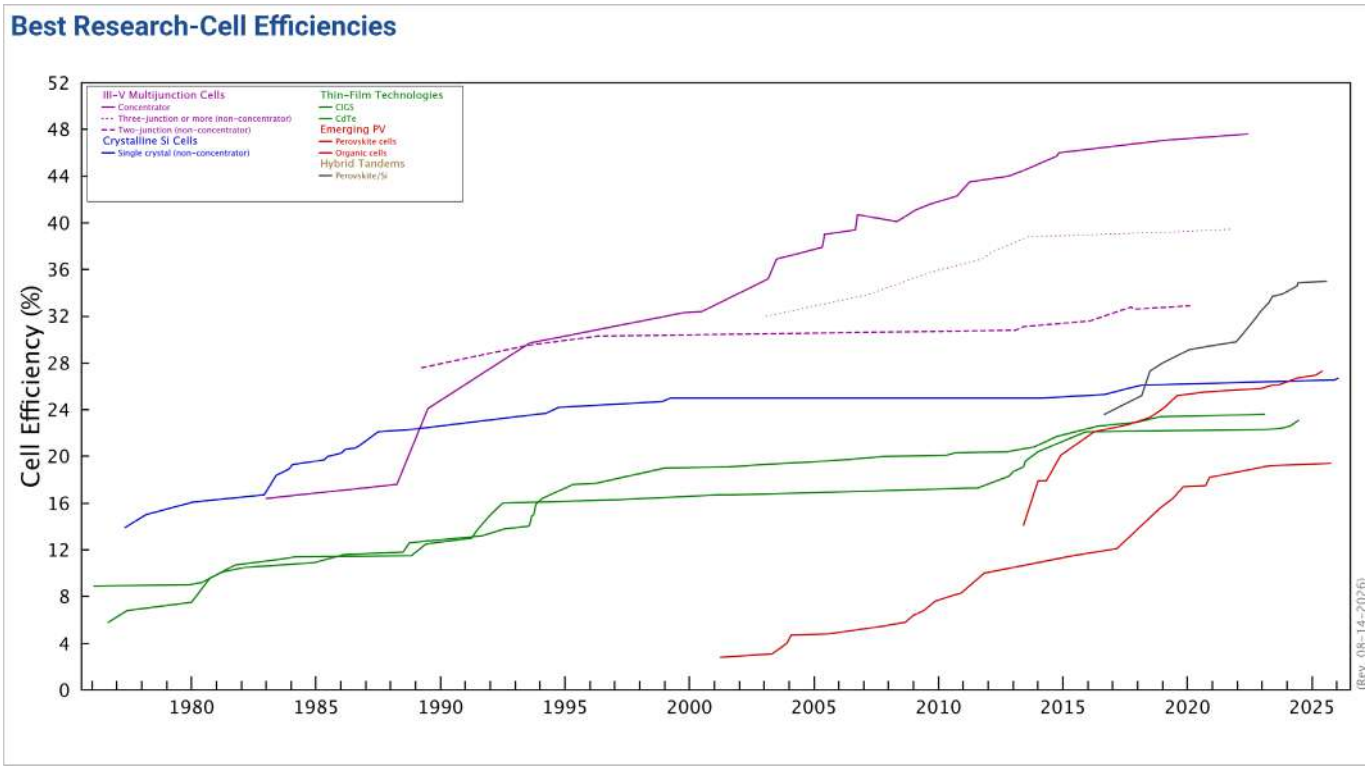


Figure 4. 소재별 인증 효율의 연도별 추이

여기서 눈여겨볼 점은 단지 효율 숫자가 아니라 제조 방식이다. 이 재료는 액체 용액을 기판에 바르고 100℃ 안팎의 저온에서 말리기만 하면 손쉽게 만들어지는데, 그 결과물은 결정립(Grain) 크기가 수백 nm에 불과한 다결정(Polycrystalline) 박막이다. 즉, 무수히 많은 결정립계(Grain boundary)가 촘촘하게 박혀 있다는 뜻이다.

이는 기존 반도체공학의 상식으로는 잘 설명되지 않는 현상이다. 실리콘이나 III-V족 반도체에서는 결정의 완성도가 성능을 완전히 좌우하며, 2.2절에서 보았듯 미세한 선 결함인 전위 하나조차 전하를 소멸시켜 성능을 급격히 떨어뜨린다. 하물며 결정립계는 전위보다 훨씬 거대한 결정 구조의 어긋남인데, 페로브스카이트는 이 결함 덩어리 속에서도 고효율로 정상 작동한다. 비결은 결함이 전혀 없어서가 아니라, "결함이 존재해도 소자 성능에 거의 해를 끼치지 않기 때문"이다. 그 근본적인 이유는 이 물질 특유의 전자 구조에 있으며, 이는 3.3절에서 자세히 다룬다.

3.2 ABX_3 구조와 조성 설계

할라이드 페로브스카이트의 화학식은 ABX_3다. 알파벳 세 자리에 각각 다른 종류의 이온이 들어간다.

A 자리에는 덩치가 큰 양이온이 들어간다. 유기 양이온인 메틸암모늄(Methylammonium, MA^+)과 포름아미디늄(Formamidinium, FA^+), 무기 양이온인 세슘(Cesium, Cs^+)이 주로 쓰인다. B 자리에는 2가 금속 양이온이 들어가는데 대부분 납(Pb^{2+})이며, 납을 피하려는 연구에서는 주석(Sn^{2+})을 쓴다. X 자리에는

할로겐 음이온인 아이오딘(I^-), 브로민(Br^-), 염소(Cl^-)가 들어간다.

이 이온들이 만드는 구조는 다음과 같다. B 이온 하나를 X 이온 여섯 개가 둘러싸 팔면체 하나를 이룬다. 이 팔면체들이 서로 꼭짓점을 맞대고 이어지면서 3차원 골격을 짜고, A 이온은 골격 사이에 생긴 빈방으로 들어간다(그림 5). 즉 A 이온은 골격을 만드는 부재가 아니라 빈방을 채워 구조가 무너지지 않게 받쳐 주는 역할이다.

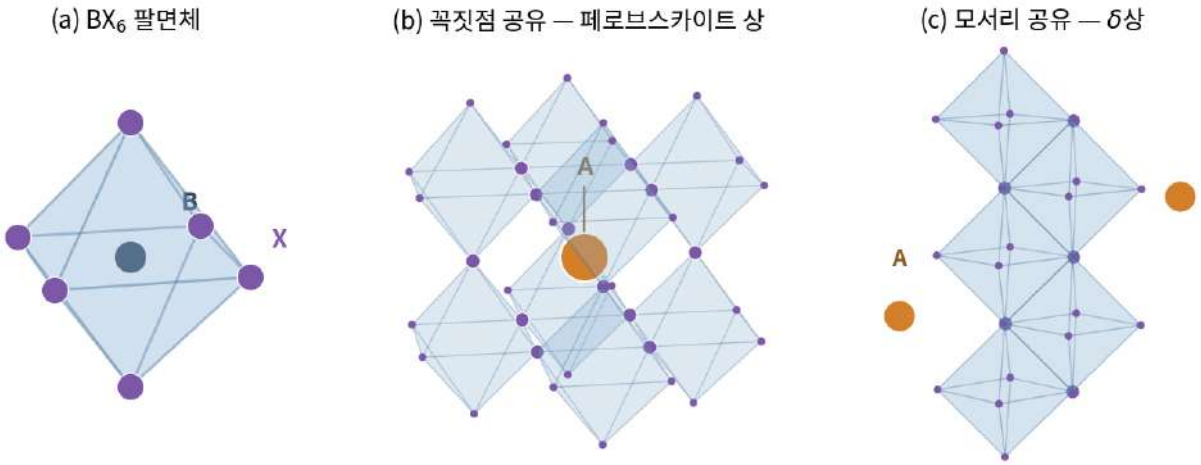


Figure 5. ABX_3 페로브스카이트의 결정 구조. (a) BX_6 팔면체 하나, (b) 팔면체가 꼭짓점을 맞대어 이룬 3차원 골격과 그 빈방을 채우는 A 이온, (c) 팔면체가 모서리를 맞댄 1차원 사슬로 재배열된 δ상

여기서 문제가 생긴다. A 이온이 빈방에 비해 너무 작으면 방을 제대로 채우지 못해 골격이 안쪽으로 기울어지고, 팔면체가 기울어지면 결국 배열이 바뀌면서 빛을 흡수하는 성질까지 달라진다. 반대로 A 이온이 너무 크면 골격이 벌어져 역시 불안정해진다. 결국 세 이온의 크기가 서로 잘 맞아야 구조가 유지된다.

이 크기 궁합을 숫자 하나로 판단하는 지표가 골드슈미트 허용인자(Goldschmidt tolerance factor)다. 기호로 t라 쓰며 다음과 같이 정의한다.

$$t = \frac{r_A + r_X}{\sqrt{2}\,(r_B + r_X)}$$

r_A, r_B, r_X는 각 자리에 들어간 이온의 반지름이다. 분자는 A 이온과 X 이온이 이루는 거리, 분모는 B 이온과 X 이온이 이루는 거리를 기하학적으로 환산한 값이므로, 두 거리가 잘 맞아떨어지면 t가 1에 가까워진다. 다시 말해 t는 "이온들이 서로 얼마나 잘 맞는가"를 나타내는 점수다.

한 가지 주의할 점이 있다. MA^+나 FA^+는 원자 하나가 아니라 여러 원자가 뭉친 분자 이온(Molecular ion)이라 완전한 구가 아니고, 보는 방향에 따라 크기가 다르다. 다만 상온에서는 이들이 빠르게 회전하므로 평균적으로 둥근 공처럼 행동한다. 그래서 회전할 때 휩쓰는 범위를 기준으로 반지름을 잡는다 (**?**).

Table 2. 대표 조성의 허용인자와 상 안정성

조성	허용인자 t	상온 상 안정성	특징
$MAPbI_3$	약 0.91	안정	열과 진공에 약함
$FAPbI_3$	약 0.99	준안정	밴드갭이 좁아 유리
$CsPbI_3$	약 0.84	불안정	열에는 강하나 상전이
$Cs_{0.2}FA_{0.8}PbI_3$	약 0.96	안정	혼합으로 t 조정

표 2에서 세 가지를 읽을 수 있다.

첫째, $CsPbI_3$가 왜 불안정한지가 숫자로 드러난다. Cs^+는 반지름이 작아 A 자리 빈방을 채우지 못하고, 그 결과 t가 0.84까지 내려간다. 상온에서 배열이 바뀌어 꼭짓점을 맞대고 있던 3차원 골격이 무너지고, 팔면체가 모서리를 맞댄 1차원 사슬 구조로 재배열된다. 이 상태를 δ상(δ phase)이라 부르는데, 밴드갭이 넓어지고 광흡수 성능이 크게 떨어져 태양전지로 쓸 수 없다 (**?**). 열에는 강하지만 구조 자체가 유지되지 않는다는 점이 유기물을 뺀 전무기(All-inorganic) 조성의 근본 난제이며, 이는 5.1절에서 다시 다룬다.

둘째, 여러 양이온을 섞으면 t를 원하는 값으로 옮길 수 있다. 작은 Cs와 큰 FA를 함께 넣으면 평균 크기가 조절되어 t가 0.96 부근으로 올라온다. 조성 설계를 감이 아니라 정량적 목표로 다룰 수 있다는 뜻이며, 실제로 최근 고효율 소자는 대부분 두세 종류의 A 자리 양이온을 섞어 쓴다.

셋째, X 자리도 조절 대상이다. I를 Br로 바꾸면 이온 반지름이 작아져 t가 커지는 동시에 밴드갭이 넓어진다. 즉 X 자리 조성 하나로 구조 안정성과 광학 특성을 함께 조절할 수 있다.

다만 t를 만능 기준으로 믿어서는 안 된다. 흔히 인용되는 $0.8 < t < 1.0$ 범위는 산화물 페로브스카이트에서 나온 값이라 할라이드계에 그대로 적용되지 않으며, 할라이드에서 안정한 입방정(Cubic phase)이 실제로 관측되는 구간은 t가 0.9에서 1.0 사이로 훨씬 좁다 (**?**). 따라서 t는 수많은 조성 후보를 1차로 걸러내는 도구로 쓰고, 최종 판단은 계산이나 실험으로 확인하는 것이 적절하다.

3.3 광전 특성과 결함 수용성

이제 이 구조가 태양전지로서 어떤 성질을 갖는지 살펴보고, 3.1절에서 남겨둔 질문에 답한다.

빛을 아주 잘 흡수한다. 페로브스카이트는 직접천이 반도체(Direct bandgap semiconductor)다. 빛을 흡수할 때 전자가 곧바로 위로 올라갈 수 있어 흡수가 매우 효율적이라는 뜻이다. 이를 나타내는 흡수계수가 실리콘보다 훨씬 큰데, 그 결과 두께 300 nm 남짓이면 가시광선의 95%를 흡수한다. 머리카락 굵기의 300분의 1 정도다. 반면 실리콘은 간접천이 반도체(Indirect bandgap semiconductor)라 전자가 올라가는 데 격자 진동(Phonon)의 도움이 필요하고, 그만큼 흡수가 더뎌 같은 일을 하려면 수백 μm가 필요하다. 천 배에 가까운 차이이며, 이것이 2.3절에서 말한 가벼움의 근거다.

전하가 멀리까지 간다. 빛이 만든 전하는 전극까지 이동해야 비로소 전류가 된다. 가는 도중에 사라지지 않고 갈 수 있는 평균 거리를 확산거리(Diffusion length)라 한다. 용액을 발라 만든 다결정 박막에서도 이 거리가 1 μm를 넘고 (**?**), 결정을 크게 키우면 175 μm를 넘는다는 보고도 있다 (**?**).

두 성질을 나란히 놓으면 의미가 분명해진다. 흡수층은 300 nm면 충분한데 전하는 그보다 세 배 이상 멀리까지 간다. 즉 소자 어디서 전하가 생기든 전극까지 무사히 도달한다는 뜻이다. 보통은 이 둘이 함께 좋기 어렵다. 빛을 잘 흡수하는 재료는 대개 결함이 많아 전하가 금방 사라지고, 전하가 오래 사는 재료는 대개 흡수가 약하기 때문이다. 두 조건을 동시에 만족하는 흡수체는 드물다.

전하를 따로 떼어낼 필요가 없다. 빛을 받으면 전자와 정공이 짝을 지어 생긴다. 이 짝을 엑시톤(Exciton)이라 하는데, 서로 붙들려 있으면 전류가 되지 못하므로 떼어내야 하고 떼는 데 드는 에너지(Exciton binding energy)가 재료마다 다르다. 페로브스카이트에서 이 값은 상온에서 재료가 이미 갖고 있는 열에너지보다 작다 (**?**). 다시 말해 가만히 두어도 열 때문에 저절로 떨어져 자유롭게 돌아다닌다. 유기 태양전지가 이 짝을 떼기 위해 두 재료를 잘게 섞은 복잡한 구조를 써야 하는 것과 대비되는 지점이다.

밴드갭을 원하는 값으로 맞출 수 있다. 3.2절에서 본 대로 X 자리 조성을 바꾸면 밴드갭이 달라진다. I를 Br로 치환할수록 넓어져 1.2 eV부터 2.3 eV까지 조절되는데, 이는 실리콘이나 다른 페로브스카이트와 쌓아 올리는 탠덤 구조의 상부셀(Top cell)에 필요한 1.7–1.8 eV를 포함하는 범위다.

결함이 있어도 해롭지 않다. 이제 3.1절의 질문으로 돌아온다. 결함이 그렇게 많은데 왜 작동하는가.

먼저 실리콘과 III-V족의 경우를 보자. 원자와 원자가 결합할 때 전자가 들어갈 자리는 두 종류로 갈라진다. 원자를 서로 붙들어 주는 결합성(Bonding) 자리와, 반대로 밀어내는 반결합성(Antibonding) 자리다. 결합성 자리들이 모여 가전자대(Valence band)를, 반결합성 자리들이 모여 전도대(Conduction band)를 이룬다. 그런데 원자 하나가 빠져 결합이 끊어지면 갈 곳을 잃은 전자가 그 둘 사이, 정확히 밴드갭 한가운데에 자리를 만든다. 이를 심부 트랩(Deep trap)이라 한다. 심부 트랩은 전자와 정공을 함께 붙잡고, 붙잡힌 둘은 만나서 사라진다. 결함 하나가 전하 여러 쌍을 없애는 셈이다.

페로브스카이트는 이 배치가 뒤집혀 있다. 가전자대 꼭대기가 결합성이 아니라 납의 6s 궤도와 아이오딘의 5p 궤도가 만드는 반결합성으로 채워져 있어 통상보다 높은 에너지에 놓인다 (**?**). 전도대 바닥은 납의 6p 궤도가 지배한다.

이 뒤집힌 배치가 결정적이다. 실제로 가장 많이 생기는 결함들, 즉 아이오딘이 빠진 자리와 A 자리 빈자리, 격자 사이에 낀 이온이 밴드갭 한가운데가 아니라 가장자리에 자리를 만든다. 이를 얕은 준위(Shallow level)라 하는데, 상온의 열에너지만으로도 붙잡은 전하를 곧바로 놓아준다 (**?**). 붙잡아 두지 못하니 전하가 사라지지 않는다. 두 경우의 차이를 그림 6에 나타냈다.

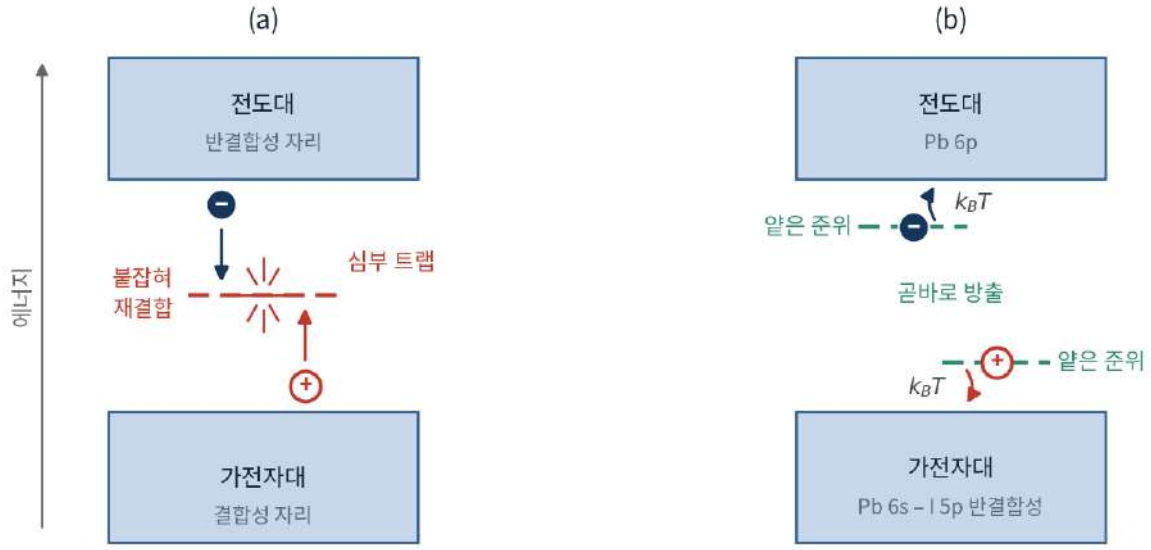


Figure 6. 결함이 만드는 에너지 준위의 비교. (a) 실리콘 · III-V족의 심부 트랩과 전하 재결합, (b) 페로브스카이트의 얕은 준위와 열에너지(k_BT)에 의한 전하 방출

여기에 두 가지가 더해진다. 하나는 이 재료의 유전상수(Dielectric constant)가 크다는 점이다. 전하를 띤 결함이 주변에 전기장을 만들어 지나가는 전하를 끌어당기는데, 유전상수가 크면 그 전기장이 강하게 가려져 힘을 못 쓴다. 다른 하나는 전하가 가볍다는 점이다. 페로브스카이트에서 전하의 유효질량이 작아 결함 주변을 빠르게 지나쳐 버린다 (**?**).

이 세 가지가 합쳐진 결과를 결함 수용성이라 부른다. 결함이 없다는 뜻이 아니라, 있어도 성능을 크게 깎지 않는다는 뜻이다. 용액을 발라 만든 조잡한 다결정 박막이 27%의 효율을 내는 이유가 여기에 있다.

3.4 방사선에 강한 이유

그리고 이 성질이 우주 적용의 출발점이 된다. 방사선이 만드는 손상 역시 결국 같은 종류의 점결함(Point defect)이기 때문이다.

궤도에는 고에너지 하전 입자가 날아다닌다. 지구 자기장에 붙잡힌 양성자와 전자가 주된 성분이고, 여기에 태양이 뿜어내는 입자와 우주선이 더해진다. 이들이 반도체에 부딪히면 갖고 있던 에너지를 두 가지 방식으로 잃는데, 이 구분이 이후 논의의 기초가 된다.

전리 에너지 손실(Ionizing Energy Loss, IEL). 입자가 지나가면서 전자를 들뜨게 한다. 그 과정에서 전자-정공 쌍이 대량으로 만들어지지만, 원자가 있던 자리는 그대로다. 격자 구조는 건드리지 않는 셈이다.

비전리 에너지 손실(Non-Ionizing Energy Loss, NIEL). 입자가 격자의 원자와 직접 부딪힌다. 충돌이 세면 원자가 제자리에서 튕겨 나가고, 원래 있던 자리에는 빈자리가, 튕겨 나간 원자는 격자 사이 틈에 자리

잡는다. 이 한 쌍을 프렌켈 쌍(Frenkel pair)이라 한다. 구조가 실제로 바뀌는 쪽이다.

기존 우주태양전지에서 영구 손상을 남기는 것은 NIEL이다 (?). 프렌켈 쌍이 그대로 심부 트랩으로 굳어지기 때문이다. IEL은 전하를 잠시 늘릴 뿐 구조를 바꾸지 않으므로 대체로 무시한다. 이 전제 위에서 우주태양전지의 수명은 변위손상선량(Displacement Damage Dose, DDD)이라는 값으로 예측한다. 입자마다 종류와 에너지가 제각각이어도 각 입자가 격자를 얼마나 밀어낼 수 있는지를 계산해 모두 합산하면 하나의 숫자로 환산되고, 그 숫자를 기준으로 지상 실험 결과에서 궤도 수명을 추정할 수 있다.

이제 3.3절의 결함 수용성을 여기에 적용해 보자. NIEL이 만드는 빈자리와 격자간 원자는 둘 다 점결함이다. 실리콘과 III-V족에서는 이것이 심부 트랩이 되므로 조사량이 쌓일수록 효율이 계속 떨어진다. 그런데 페로브스카이트에서는 같은 점결함이 대부분 얕은 준위로 귀결된다 (??). 결함이 생기는 것 자체는 막지 못하지만, 생긴 결함이 전하를 붙잡지 않는다. 방사선 내구성은 별개의 특수한 성질이 아니라 결함 수용성이 그대로 이어진 결과인 셈이다.

실험이 이를 뒷받침한다. Lang 등은 68 MeV 양성자를 결정질 실리콘의 손상 임계보다 약 천 배 높은 양까지 조사했다 (?). 그런데 단락전류(Short-circuit current)만 줄었을 뿐 개방전압(Open-circuit voltage)과 충전율(Fill factor)은 거의 그대로였다. 이 둘은 전하가 중간에 사라지는 정도에 민감한 항목이므로, 값이 유지되었다는 것은 새로 생긴 결함이 전하를 없애는 통로로 작동하지 않았다는 뜻으로 읽힌다. 전자를 쬔 실험에서도 비슷한 결과가 나와, 1 MeV 전자를 상당한 양까지 조사한 조건에서 초기 효율의 92% 이상이 유지되었으며 (?), 전 페로브스카이트 다중접합 구조에서도 양성자 내성이 보고되었다 (?).

손상이 스스로 사라진다. 더 주목할 점은 그다음이다. 조사를 멈추면 성능이 일부 되돌아온다. 손상이 덜 생기는 것이 아니라 이미 생긴 손상이 사라진다는 뜻이며, 기존 우주태양전지에서는 볼 수 없던 거동이다.

Lang 등의 실험에서 조사가 끝난 뒤 단락전류가 회복되었고 (?), 이후 여러 연구가 유사한 결과를 보고했다. Afshari 등은 삼중 할라이드(Triple halide) 조성 소자에 양성자를 쬐어 직후 성능이 떨어지는 것을 확인했으나, 두 달이 지난 뒤 조사 전 수준으로 돌아온 것을 관찰했다 (?). Durant 등은 회복을 우연한 관측이 아니라 실험 변수로 다루어, 양성자를 소자의 특정 깊이에 집중시켜 손상을 만든 뒤 전자 이온화를 일부러 유도해 회복이 일어나는지를 확인했다 (?). 가장 직접적인 증거는 Kirmani 등의 이중 조사(Dual irradiation) 실험이다 (?). 먼저 0.06 MeV의 느린 양성자를 쬐었다. 느린 양성자는 얕은 곳에서 멈추므로 원자를 밀어내는 NIEL 비중이 크고, 그 결과 효율이 초기의 74%로 떨어졌다. 이어서 원자를 밀어내는 양은 그대로 두고 전자를 들뜨게 하는 IEL만 수십 배로 키운 1.0 MeV 양성자를 추가로 쬐었다. 그러자 효율이 다시 올라가 NIP 구조에서 83%, PIN 구조에서 85%를 회복했다. 방사선의 한 성분이 다른 성분이 만든 손상을 되돌린 것이다. 지금까지의 주요 보고를 표 3에 정리했다.

Table 3. 양성자 조사 후 회복 거동에 관한 주요 보고

보고	조사 조건	손상	회복
(**?**)	68 MeV 양성자	단락전류 감소, 개방전압 · 충전율 유지	조사 중단 후 단락전류 회복
(**?**)	양성자 조사, 삼중 할라이드	광전 특성 감소	2개월 후 초기 수준 복귀
(**?**)	깊이를 지정한 양성자 조사	국소 변위 손상	전자 이온화로 회복 유도
(**?**)	0.06 MeV 후 1.0 MeV 이중 조사	효율 74%로 감소	83–85%로 회복

회복이 왜 일어나는지 살펴보면 두 가지 기전이 함께 작용한다(그림 7).

첫째, 이온이 원래 자리로 돌아갈 수 있는 유연한 격자 구조다. 페로브스카이트는 이온 결합으로 이루어진 부드러운(Soft) 결정 격자다. 결합이 약한 만큼 이온이 위치를 이동하는 데 필요한 활성화 에너지(Activation energy)도 매우 작아, 아이오딘 이온이 격자 내 빈자리를 옮겨 다니는 데 드는 에너지가 기존 실리콘 반도체보다 몇 배 이상 낮다 (**?**). 이 정도 수준이면 별도로 고온 열처리를 하지 않아도 상온에서 이온이 스스로 움직일 수 있다. 이제 방사선에 의해 형성된 프렌켈 쌍을 떠올려 보자. 튕겨 나간 아이오딘 원자와 그 자리에 남겨진 공공(Vacancy)이 바로 옆에 근접해 있다. 아이오딘이 아주 조금만 움직여도 이 빈자리를 곧바로 메우게 되며, 그 순간 결함 두 개가 한꺼번에 소멸한다. 계산에 따르면 이 둘이 흩어져 멀어지는 것보다 다시 합쳐지는 쪽이 에너지적으로 훨씬 안정하다 (**?**).

둘째, 입자의 전리 손상이 역설적으로 회복을 돕는다. Kirmani 등은 비전리 손실이 만드는 손상 외에, 전리 손실이 격자를 국소적으로 미세하게 진동시키며 그 진동 에너지가 튕겨 나간 원자를 제자리로 밀어 넣는 동력이 된다고 설명한다 (**?**). 즉, 방사선 자체가 손상 부위만 아주 좁게 열처리(Annealing)하는 셈이다. 이 기전은 격자가 부드럽고 전자와 격자 진동 사이의 상호작용(Electron-phonon coupling)이 강한 물질에서 두드러진다. 페로브스카이트는 열전도도까지 낮아 국소적으로 발생한 열이 잘 빠져나가지 않고 미세 열처리에 효과적으로 쓰인다 (**?**). 딱딱한 공유결합(Covalent bond) 형태의 반도체에서는 이러한 현상이 나타나지 않는다.

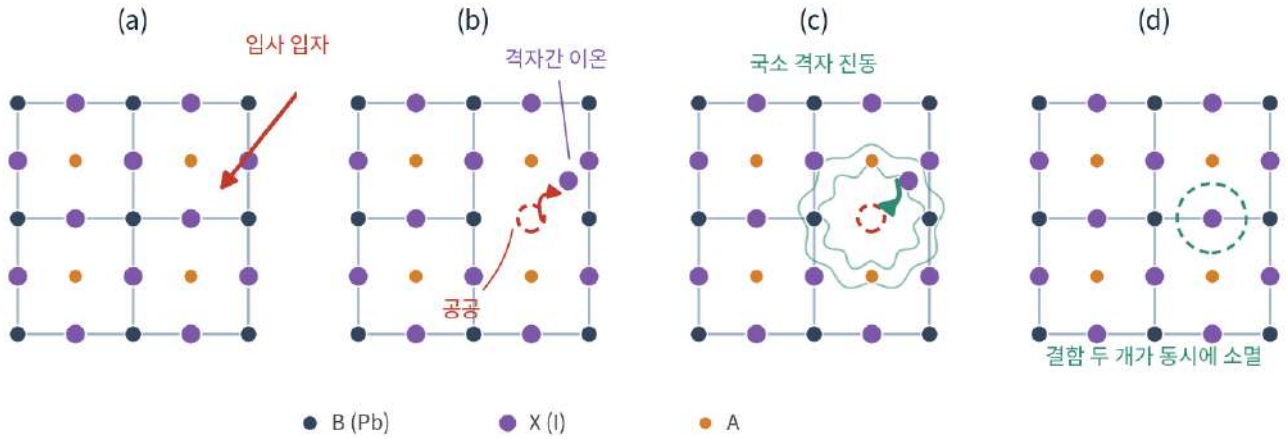


Figure 7. 프렌켈 쌍의 생성과 자가 회복 과정. (a) 고에너지 입자의 입사, (b) 공공과 격자간 이온의 생성, (c) 국소 격자 진동이 돕는 이온의 근거리 이동, (d) 결함 두 개의 동시 소멸

다만 회복에는 조건이 있다. 자가 회복에는 이온이 움직일 수 있는 환경과 이온을 움직이게 할 구동력이 모두 필요하며, 온도와 햇빛, 방사선의 전리 성분이 그 구동력을 제공한다. 뒤집어 말하면 자가 회복은 재료가 항상 지니고 있는 고정된 성질이 아니라 운용 조건에 따라 달라지는 현상이며, 그중에서도 온도에 대단히 민감하다. 이온이 빈자리를 메우려면 활성화 에너지라는 문턱을 넘어야 하는데, 온도가 내려가면 이 문턱을 넘을 수 있는 이온의 비율이 지수함수적으로 급감하기 때문이다. 실제로 햇빛을 받는 일조 구간(약 390 K)과 지구 그림자에 들어가는 식 구간(약 120 K)을 비교하면 회복 속도가 열 자릿수 이상 차이 나며, 극저온인 식 구간 동안에는 회복이 사실상 정지한다고 보아야 한다.

반복되는 손상에서도 마찬가지다. 빛과 어둠을 번갈아 가한 실험에서, 격자 팽창과 밴드갭 이동은 되돌아왔지만 회절 피크(Diffraction peak)의 폭은 사이클을 거듭할수록 계속 넓어졌다 (**?**). 결정이 조금씩 무너지고 무질서한 영역이 남는다는 뜻이다. 조사량이 많아 손상이 한곳에 집중된 경우에도 열과 빛을 함께 가해 효율이 돌아오지 않았다는 보고가 있다 (**?**). 즉 사이클마다 되돌아오는 성분과 되돌아오지 않는 성분이 함께 생기는데, 앞의 것만 사라지고 뒤의 것은 조금씩 쌓인다. 표 4는 이 구분을 정리한 것이다.

Table 4. 반복 손상에서 가역 성분과 비가역 성분의 구분

현상	시간 척도	반복 시 거동
프렌켈 쌍 재결합	매우 빠름(밀리초–초)	반복해도 유지
이온 재분포 · 재혼합	느림(수십 시간–수개월)	온도가 낮으면 정지
결정성 손실 · 비정질화	회복되지 않음	사이클마다 축적
유기 양이온 분해	회복되지 않음	단조 증가
수송층 · 전극 열화	회복되지 않음	단조 증가

문헌에서 "회복"에 걸리는 시간이 밀리초에서 두 달까지 열 자릿수 가까이 흩어져 있는 것도 이 때문이다 (**???**). 측정이 틀렸다기보다 서로 다른 현상을 하나의 이름으로 묶어 부르고 있다고 보는 편이 자연스럽다. 튕겨 나간 이온이 바로 옆 빈자리로 돌아가는 것은 순식간에 끝나지만, 흩어진 이온이 다시 고르게 섞이는 데는 훨씬 오래 걸리고, 한번 무너진 결정이 되살아나는 일은 상온에서 사실상 일어나지 않는다.

따라서 자가 회복은 소자를 언제나 지켜 주는 안전장치가 아니라 특정 조건에서만 작동하는 기전으로 이해해야 한다. 설계의 목표도 회복 능력에 기대는 것이 아니라, 회복이 작동하는 조건 안에 소자를 붙들어 두는 것이 된다. 이는 온도 관리가 방사선 대책의 일부가 된다는 뜻이기도 하다 (**?**).

4. 실제 사례 분석: 궤도 실증

4.1 지상 시험만으로는 확인되지 않는 것

3장에서 살펴본 조사 실험들은 모두 지상 가속기에서 양성자나 전자를 쬔 결과다. 그런데 실제 궤도에서는 진공과 극심한 온도 변화, 원자상 산소, 자외선, 방사선이 동시에 걸린다. 특히 3.4절에서 본 대로 자가 회복이 온도에 지수적으로 의존한다면, 90분마다 일조와 식이 번갈아 오는 저궤도에서는 회복이 진행되다 멈추기를 반복하게 된다. 이런 결합 효과는 응력을 하나씩 따로 가하는 지상 시험으로는 재현되지 않으며, 이 때문에 실제 궤도에 소자를 올려 보는 시험이 이어지고 있다.

4.2 성층권에서 저궤도까지

시작은 성층권이었다. 2018년 $MAPbI_3$ 소자를 기구에 실어 고도 32 km까지 올린 시험에서, 3시간 이상 노출한 뒤 초기 효율의 64%가 유지되었다 (**?**). 흥미로운 것은 손실의 원인이다. 방사선이 아니라 저온에서의 상전이(Phase transition)와 계면 재결합(Interface recombination)으로 분석되었는데, 이는 3.2절에서 본 상 안정성 문제가 실제 환경에서 먼저 드러날 수 있음을 보여준다.

2020년에는 관측 로켓 비행 시험이 이루어져 페로브스카이트와 유기 태양전지를 함께 실어 고도 약 240 km까지 올렸다 (**?**). 비행 중 전기적 특성을 측정한 결과 360초의 측정 구간에서 방사선에 의한 손상은

관측되지 않았고 출력 밀도가 약간 감소한 정도였다. 짧은 시간이지만 우주 환경에서 실제로 작동한 첫 사례다.

2020년부터 2021년에 걸쳐서는 국제우주정거장(International Space Station, ISS)에서 10개월간의 노출 시험이 수행되었다 (**?**). 이 시험은 다음 절에서 따로 다룬다.

2024년에는 큐브샛(CubeSat) 시험이 있었다. Big Red Sat-1은 페로브스카이트 태양전지를 저궤도에서 평가한 1U 규모 큐브샛으로, 2024년 3월 국제우주정거장 보급 임무를 통해 발사되었다. 2025년 10월에는 일본 우주항공연구개발기구(JAXA)의 무인 보급선 HTV-X1에 페로브스카이트 태양전지가 탑재되어, 약 두 달간 우주 환경에 노출하며 전류-전압 특성을 측정하는 계획이 진행 중이다.

4.3 국제우주정거장 10개월 노출 시험

지금까지의 궤도 시험 중 3장의 질문에 가장 직접적으로 답하는 것은 국제우주정거장 시험이다 (**?**).

조건. 봉지(Encapsulation)한 $MAPbI_3$ 시료를 고도 408 km에서 10개월간 노출했다. 이 궤도에서는 90분마다 일조와 식이 번갈아 오므로 시료는 궤도마다 200 °C가 넘는 온도 변화를 겪었고, 그동안 방사선에도 계속 노출되었다. 3.4절에서 자가 회복이 온도에 민감하다고 했던 조건이 실제로 수천 번 반복된 셈이다.

결과. 회수한 시료는 유리가 누렇게 변하고 PbI_2가 생기는 등 부분적인 열화를 보였다. 그런데 지구로 가져와 다시 빛을 쬐자 구조가 상당 부분 복원되면서 결함이 줄고 전하 수명이 늘었다.

의미. 10개월간 손상과 회복이 반복된 뒤에도 회복 능력 자체가 남아 있었다는 뜻이다. 표 4의 표현을 빌리면, 비가역 성분(PbI_2 생성, 유리 변색)이 분명히 쌓였지만 가역 성분을 되돌리는 기전이 완전히 죽지는 않았다. 3.4절에서 우려했던 "반복되면 회복이 무너지는가"라는 질문에 대해, 적어도 저궤도의 10개월 규모에서는 부분적으로 긍정적인 답을 준 자료다.

한계. 다만 이 결과를 과대 해석해서는 안 된다. 첫째, 회복은 지상으로 회수해 상온에서 빛을 쬔 뒤에 확인된 것이지 궤도에서 실시간으로 측정된 것이 아니다. 둘째, 시료는 봉지된 상태였으므로 원자상 산소나 자외선의 직접적인 영향은 걸러졌다. 셋째, 시료 면적이 연구용 수준이라 실제 어레이의 거동을 대표한다고 보기 어렵다.

이런 한계는 다른 시험에도 공통된다. 지금까지의 궤도 실증은 규모가 작고 기간이 짧으며 소자 면적도 연구용 수준이므로, 현 단계는 실용화 이전의 초기 검증으로 보는 것이 적절하다. 다만 지상 시험과 궤도 실측을 대조할 자료가 쌓이기 시작했다는 점에서 의미가 있으며, 이런 시험이 더 필요하다.

5. 페로브스카이트의 미래

페로브스카이트가 우주용 전력원으로 자리 잡기 위해서는 소재, 검증, 시스템 세 측면에서의 노력이 병행되어야 하며, 4장의 국제우주정거장 사례는 이 세 축이 실제로 어떻게 맞물리는지를 압축적으로

보여준다.

5.1 소재: 회복 능력과 상충하는 것들

소재 수준에서 가장 까다로운 점은 회복을 가능하게 하는 성질이 동시에 열화의 원인이라는 것이다.

첫째는 이온 이동이다. 3.4절에서 회복의 근원으로 이온 이동을 들었는데, 같은 성질이 문제도 만든다. 전압이 걸리면 이온이 한쪽으로 쏠려 전류-전압 곡선이 측정 방향에 따라 달라지고(Hysteresis), 혼합 할라이드 조성에서는 빛을 받는 동안 I와 Br이 서로 나뉘기도 한다(Photo-induced phase segregation) (**?**). 밴드갭이 국소적으로 달라진다는 뜻이며 탠덤 상부셀용 조성에서 특히 문제가 된다. 그런데 격자를 딱딱하게 만들어 이온 이동을 억제하면 결함 수용성과 자가 회복을 함께 잃는다. 회복에 필요한 것은 프렌켈 쌍 규모의 짧은 거리 이동이고 열화를 일으키는 것은 소자 두께 규모의 장거리 이동이므로, 요구되는 것은 이동성을 없애는 설계가 아니라 두 거리 척도를 분리하는 설계다.

둘째는 상 안정성이다. 유기 양이온 MA^+와 FA^+는 결합이 약해 열과 진공, 자외선에 분해되기 쉽고 한번 분해된 산물은 다시 조립되지 않으므로, 유기물을 뺀 전무기 조성이 반복 회복에 유리할 가능성이 있다. 그러나 3.2절에서 본 대로 $CsPbI_3$는 t가 0.84까지 내려가 상온에서 δ상으로 전이하므로 상 자체가 유지되지 않는다 (**?**). Br을 섞거나 표면을 처리해 안정화하는 방법이 연구되고 있으나, Br을 넣으면 밴드갭이 넓어져 단일접합 효율이 낮아진다.

2018년 성층권 기구 시험에서 손실의 주범이 방사선이 아니라 저온 상전이였다는 사실은 이 문제가 이론적 우려에 그치지 않음을 보여준다. 정리하면 회복 능력, 이온 이동에 의한 열화, 상 안정성은 부드러운 이온 격자라는 하나의 원인에서 함께 나오므로, 하나를 조정하면 나머지가 따라 움직인다. 2.2절에서 III-V족의 세 가지 제약이 단결정 기판이라는 하나의 조건에서 파생되었던 것과 같은 구조다.

5.2 검증: 무엇을 어떻게 재야 하는가

새로운 소재가 III-V족을 대체하려면 성능뿐만 아니라 예측 가능성을 확보해야 한다고 2.1절에서 말했다. 현재의 시험 체계로는 이것이 어렵다.

첫째, 반복 시험이 없다. 지금까지의 조사 실험은 대부분 조사와 회복을 한 번씩만 본다. 그런데 저궤도의 실제 조건이 바로 반복이다. 일조 구간에서는 온도가 높아 회복이 진행되고 식 구간에서는 온도가 낮아 회복이 멈추며, 그동안에도 방사선은 계속 쏟아진다. 즉 열 사이클과 방사선 손상은 따로 노는 변수가 아니라 서로 맞물려 있다. 두 시험을 따로 수행하는 기존 방식으로는 이 맞물림을 확인할 수 없으므로, 두 응력을 동시에 가하고 조사와 회복을 여러 차례 반복하는 시험 체계가 필요하다 (**?**). 국제우주정거장 시험이 귀중한 이유가 여기에 있는데, 이 반복 조건을 실제로 겪은 거의 유일한 자료이기 때문이다. 다만 4.3절에서 지적했듯 궤도에서 실시간 측정이 이루어지지 않았으므로, 반복 도중에 무슨 일이 일어났는지는 여전히 알 수 없다.

둘째, 보고 방식이 통일되지 않았다. 많은 논문이 양성자의 에너지와 개수만 적는데, 그러면 3.4절에서 설명한 변위손상선량으로 환산할 수 없다. 환산이 안 되면 논문끼리 비교할 수도, 궤도 수명을 추정할 수도 없다 (**?**). 페로브스카이트가 III-V족만큼 예측 가능한 소재가 되려면 이 부분이 먼저 정리되어야 한다.

셋째, "회복"이라는 용어 자체를 정밀하게 써야 한다. 표 4에서 보았듯 밀리초 단위의 프렌켈 쌍 재결합과 수개월 단위의 이온 재혼합은 전혀 다른 현상인데 같은 이름으로 불린다. 앞으로의 보고에서는 어느 단계의 회복을 말하는지 명시할 필요가 있다.

5.3 시스템: 흡수층 밖의 문제들

수송층과 전극. 흡수층 위아래에는 전하수송층(Charge transport layer)과 전극이 있는데, 유기 수송층은 자외선과 방사선에 분자 사슬이 끊어지고 금속 전극은 흡수층에서 흘러나온 아이오딘과 반응하거나 반대로 흡수층 쪽으로 파고든다(Metal migration) (**?**). 이런 변화에는 자가 회복 기전이 적용되지 않는다. 흡수층이 온전히 돌아와도 그 위아래가 망가져 있으면 소자는 작동하지 않으며, 최근 연구는 실제 소자 수명을 제한하는 것이 흡수층이 아닐 가능성을 지적한다 (**?**).

원자상 산소와 열 사이클. 저궤도에는 원자 상태의 산소가 존재해 위성 진행 방향 면에 부딪혀 표면을 깎아내며, 보호층이 없으면 유기층과 금속 전극이 빠르게 손상된다. 무기 배리어(Inorganic barrier)를 적용하면 이 손상을 크게 줄일 수 있다는 보고가 있다 (**?**). 열 사이클도 문제다. 층마다 열팽창계수가 다르므로 온도가 오르내릴 때 계면에 응력이 걸리는데, 페로브스카이트 박막의 계면 파괴 인성(Fracture toughness)은 반도체 박막 중 낮은 편이라 층이 벗겨지기 쉽다 (**?**).

봉지: 해답이자 새로운 제약. 위 두 문제의 해답은 봉지지만, 봉지는 곧바로 출력 대 질량비 문제로 되돌아온다. 문헌에 보고된 페로브스카이트의 출력 대 질량비는 23 W/g (**?**), 실험실 수준에서는 50 W/g까지인데 (**?**), 이는 대체로 봉지하지 않은 소자 기준이다. 두꺼운 고분자 필름을 양면에 붙이면 그 무게가 그대로 분모에 더해져 III-V족의 5.5 W/g와 차이가 없어진다. 반면 원자층 증착(Atomic Layer Deposition, ALD)으로 만든 수십 nm 두께의 무기 배리어는 무게 기여가 거의 없지만 잘 갈라진다. 두 종류를 교대로 쌓는 구조가 검토되고 있으며, 결국 문헌의 W/g 수치를 인용할 때는 봉지 조건을 함께 확인해야 한다.

국제우주정거장 시험 시료가 봉지된 상태였다는 점은 이 맥락에서 다시 읽힌다. 봉지 덕분에 10개월을 버텼지만, 동시에 그 봉지의 무게가 실제 임무에서 얼마의 대가를 치르는지는 그 시험이 답해 주지 않는다. 이 세 축이 함께 갖추어질 때 페로브스카이트는 비로소 궤도 인프라의 전력원 후보가 될 수 있다.

결론

본 탐구를 통해 할라이드 페로브스카이트가 차세대 우주용 전력원으로서 현실적인 후보임을 확인하였다. 인공지능과 클라우드 수요가 늘면서 데이터 인프라의 일부를 궤도로 옮기려는 시도가 나타나고 있으나, 궤도에는 전력망이 없어 전원을 현지에서 확보해야 한다. 현재 쓰이는 III-V족 다중접합 전지

는 성능과 신뢰성이 검증되어 있지만, 단결정 기판 위에서만 만들 수 있다는 조건 하나가 단가와 질량, 형상을 동시에 묶고 있어 위성을 다수 배치하거나 대면적 어레이를 펼치는 임무에서는 제약이 된다.

페로브스카이트는 이 제약을 우회할 수 있는 후보다. 단결정 기판이 필요 없고, 300 nm 두께로 태양광을 충분히 흡수하며, 저온 용액 공정으로 만들 수 있다. 특히 주목할 점은 방사선 내구성이 별도의 특수한 성질이 아니라 결함 수용성이 그대로 이어진 결과라는 것이다. 가전자대가 Pb 6s와 I 5p의 반결합성 혼성으로 이루어지므로 점결함이 밴드갭 깊은 곳이 아니라 밴드 가장자리에 준위를 만드는데, 방사선이 만드는 손상도 결국 점결함이기 때문이다. 여기에 자가 회복이 더해진다. 부드러운 이온 격자에서 요오드는 낮은 에너지로 움직이므로 밀려난 이온이 공공으로 되돌아가면 결함 두 개가 함께 사라지고, 방사선의 전리 성분이 만드는 국소 격자 진동이 이 과정을 돕는다.

특히 국제우주정거장 10개월 노출 시험은 궤도마다 200 °C가 넘는 온도 변화와 방사선을 반복해서 겪은 뒤에도 회복 능력이 남아 있음을 보여주었다는 점에서 의미가 있다. 다만 회복이 무조건적이지 않다는 점도 함께 확인되었다. 온도에 지수적으로 의존하므로 저궤도의 식 구간에서는 사실상 멈추고, 반복 사이클에서는 가역 성분이 사라지는 동안 결정성 손실 같은 비가역 성분이 조금씩 쌓인다. 따라서 자가 회복은 소자를 지켜 주는 안전장치가 아니라 특정 조건에서만 작동하는 기전으로 다루어야 하며, 설계의 목표는 그 조건 안에 소자를 붙들어 두는 것이다.

앞으로 페로브스카이트가 궤도 인프라의 전력원으로 자리 잡기 위해서는 소재·검증·시스템 세 측면의 노력이 함께 이루어져야 한다. 회복 능력을 유지하면서 장거리 이온 이동만 억제하는 조성 설계, 열 사이클과 방사선을 동시에 인가하고 조사와 회복을 반복하는 시험 체계, 그리고 가벼우면서도 원자상 산소와 자외선을 막아내는 봉지 기술이 그것이다. 이러한 기반이 마련된다면 페로브스카이트는 기존 III-V족 전지를 보완하며, 대규모 궤도 인프라 시대의 전력 문제를 해결하는 데 중요한 역할을 할 것으로 기대된다.

Example Questions

해당 문항을 제외한 자유주제로 연구해도 무관함.

Q1. **소재 설계*:* 페로브스카이트의 자가 회복은 요오드 이온이 쉽게 움직이기 때문에 가능하지만, 같은 성질이 이온 이동에 의한 성능 저하도 일으킨다. 격자를 딱딱하게 만들면 두 성질을 함께 잃는다는 점을 고려하여, 회복 능력은 유지하면서 장거리 이온 이동만 줄일 수 있는 설계 방안을 제안하시오.**

Q2. **반복 손상*:* 조사와 회복이 반복될 때 가역 성분은 사라지지만 결정성 손실 같은 비가역 성분은 축적된다. 이 축적 속도를 늦추려면 조성, 계면, 운용 조건 중 무엇을 어떻게 조절해야 하는지 근거를 들어 제안하시오.**

Q3. **시험 설계*:* 기존 연구는 대부분 조사와 회복을 한 차례씩만 관찰한다. 저궤도의 실제 조건인 *90*분 주기 반복을 모사하려면 어떤 시험을 설계해야 하는지, 인가할 응력과 측정 항목, 판정 기준을 함께 제시하시오.**

Q4. **시스템 적용*:* 페로브스카이트 태양전지를 우주 임무에 가장 먼저 적용한다면 어떤 임무가 적절한가*?* 궤도와 요구 수명, 전력 규모, 비용 조건을 근거로 설명하시오.**

Vol. 07 (2026)

Perovskite Solar Cells as Power Sources for Space Applications: Material Properties and Radiation Tolerance

Abstract

As the spread of artificial intelligence and cloud services drives a rapid increase in data-processing demand, efforts are emerging to move part of the computing infrastructure into Earth orbit in order to overcome the power-supply and cooling constraints faced by terrestrial data centers. Orbital facilities cannot rely on the terrestrial power grid and therefore must generate electricity locally, making solar cells a critical power source. However, the III-V multijunction cells currently used in space are expensive and can only be fabricated on heavy single-crystal substrates, making them ill-suited to new missions that require satellite constellations numbering in the thousands or very large-area arrays.

This report introduces perovskites as a promising alternative. It first summarizes how the requirements for space solar cells differ from those for terrestrial applications and analyzes the origins of the limitations of current materials. It then examines the structural origin of perovskites, the principles of compositional design, their outstanding optoelectronic properties, and their defect tolerance, before explaining how this defect tolerance leads to radiation resistance and 'self-healing.' It further reviews orbital demonstration cases ranging from stratospheric balloon experiments to long-term exposure on the International Space Station and proposes the material-, validation-, and system-level challenges that must be addressed for practical deployment.

Introduction

The rapid expansion of generative AI model training and inference, together with cloud services, is causing terrestrial data-processing demand and electricity consumption to surge. Large data centers require hundreds of megawatts of power and cooling capacity, yet further expansion on the ground is increasingly constrained by grid-connection delays, shortages of cooling water, and land-use regulations. As an alternative, the concept of 'orbital computing infrastructure (orbital data centers),' which would exploit abundant solar radiation and the vacuum environment of space, has emerged.

Orbital computing infrastructure has recently attracted renewed attention because terrestrial constraints have intensified while the advantages of orbit have become increasingly quantifiable. These advantages can be

summarized in four points.

First, solar energy is highly available. Unlike on Earth, orbit is free from night, clouds, and atmospheric attenuation. In particular, a dawn-dusk sun-synchronous orbit, which follows the boundary between day and night and avoids entering Earth's shadow, can provide nearly continuous illumination and potentially yield up to eight times more solar energy than on the ground (Google Research, 2025).

Second, the cooling mechanism is different. In vacuum, there is no convection or conduction, so heat must be rejected entirely by radiation, requiring large radiator areas. However, no cooling water is needed. In other words, the thermal problem in orbit shifts from a 'water problem' to the same 'area-and-mass problem' faced by solar arrays.

Third, orbit is free from land-use and many siting constraints. Obstacles to terrestrial expansion such as land acquisition, local approval, and grid-connection queues do not exist in the same form.

Fourth, launch costs have fallen. The spread of reusable launch vehicles has reduced transportation costs per kilogram, making large-scale delivery to orbit increasingly realistic.

Concrete developments are already appearing. Low Earth Orbit (LEO) communications constellations containing thousands of satellites are already in operation (Magliarisi et al., 2025), and in November 2025 Google announced a research plan to place clusters of satellites in sun-synchronous orbit carrying computing semiconductors (Google Research, 2025). However, because many issues—including thermal management and radiation tolerance—remain unverified, the concept remains largely at the proof-of-concept stage.

Orbital infrastructure must secure its power locally. RTGs (Radioisotope Thermoelectric Generators) and small nuclear reactors face limitations such as relatively low output and substantial safety-validation requirements. Thus, for continuous power delivery at the MW-to-GW scale required by large computing facilities, solar cells are the principal practical option (Song et al., 2026). LEO constellations also require launching thousands of satellites, so both launch-vehicle fairing volume and mass constraints become critical (Magliarisi et al., 2025). Yet conventional space solar cells were optimized for the long-term operation of individual satellites rather than for such large-scale missions, meaning that the choice of core photovoltaic material must be reconsidered from the outset.

This report first summarizes the requirements that space solar cells must satisfy (Section 1), then analyzes the origins of the limitations of currently used III-V multijunction cells (Section 2). It next examines what perovskites are and why they exhibit strong radiation tolerance (Section 3), reviews demonstration tests conducted in actual orbital environments (Section 4), and finally organizes the remaining challenges for practical deployment along three axes: materials, validation, and systems (Section 5).

1. Generating Electricity in Space

1.1 Solar Irradiation Conditions: AM0

To evaluate solar-cell performance and understand the operating environment, the first step is to characterize the illumination conditions. The common standard used to describe differences in the path that sunlight travels before reaching a device, whether on Earth or in space, is 'Air Mass (AM).'

AM is the ratio describing how long a path sunlight has traveled through Earth's atmosphere. The path length for the Sun directly overhead is defined as 1 (AM1), and the value increases as the solar angle becomes more oblique and the optical path through the atmosphere grows longer. Using the solar zenith angle θ, AM can be approximated as follows.

$$\mathrm{AM} = \frac{1}{\cos\theta}$$

The percentage efficiencies commonly quoted for terrestrial solar cells are measured under the standard ground condition AM1.5G. This condition represents the average solar position at mid-latitudes, corresponding to a zenith angle of about $48.2°$, with the irradiance standardized to $1000\ \mathrm{W/m^2}$. By contrast, space uses AM0 as the standard because sunlight passes through no atmosphere at all. In unobstructed space, solar irradiance is approximately $1361–1366\ \mathrm{W/m^2}$, substantially higher than at the Earth's surface (AST, 2019). As shown in Figure 1, AM1 corresponds to the Sun directly overhead; as the Sun moves toward the horizon, the geometric path through the atmosphere becomes longer and the value increases to AM1.5 and AM2. An orbital environment with no atmosphere corresponds to AM0.

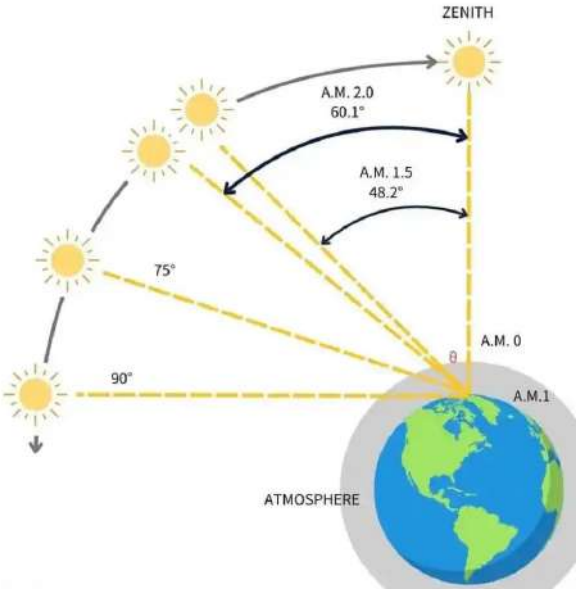


Figure 1. Definition of air mass as a function of the solar zenith angle

The two conditions differ not only in the intensity of sunlight but also substantially in its spectral composition. On Earth, atmospheric ozone blocks ultraviolet (UV) radiation, while water vapor and carbon dioxide absorb portions of the near-infrared (NIR) spectrum. In the atmosphere-free AM0 environment, these filtering effects are absent. Figure 2, which overlays the two spectra, shows this difference and reveals three major consequences for space solar-cell design.

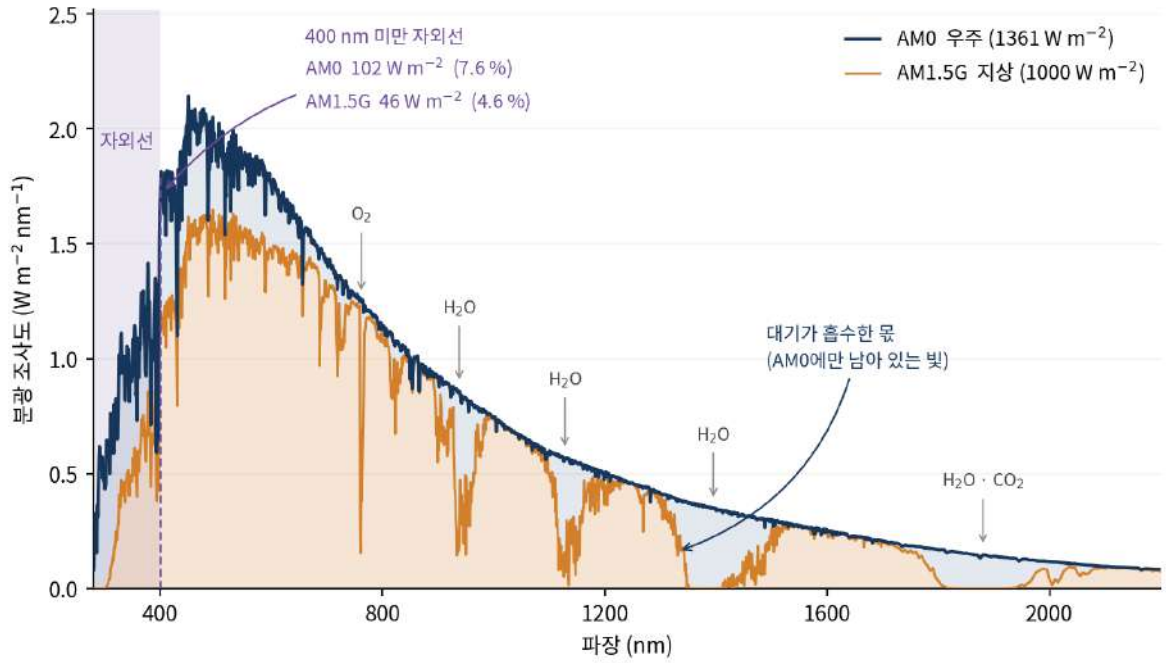


Figure 2. Comparison of the AM0 and AM1.5G solar spectra (AST, 2019)

First, the higher ultraviolet fraction increases material degradation. Space contains a noticeably larger proportion of high-energy ultraviolet photons with wavelengths below 400 nm. Because these photons carry sufficient energy to break chemical bonds in organic materials within the solar cell, organic charge-transport layers and protective coatings degrade much faster than under terrestrial conditions.

Second, the near-infrared region that is partially absorbed by the terrestrial atmosphere becomes fully available. In the AM1.5G spectrum, atmospheric water vapor removes portions of the near-infrared band, whereas these regions are restored under AM0 conditions. As a result, the absolute number of photons available to narrow-bandgap absorber materials increases.

Third, there is no guarantee that the material with the highest performance on Earth will also be optimal in space (Romano et al., 2022). Because the spectral composition changes, the position of the **'optimal bandgap (the energy range for light absorption)'** also shifts. When the same solar cell is measured under terrestrial and space conditions, its efficiency is generally lower under AM0. Therefore, terrestrial efficiency

values cannot be taken directly as measures of space performance; space devices must be redesigned and evaluated with their absorption ranges optimized specifically for AM0 conditions.

1.2 Four Key Requirements for Space Solar Cells

In addition to illumination conditions, space solar cells must satisfy several key requirements that are rarely major considerations for terrestrial photovoltaics.

Radiation resistance. The orbital environment includes the Van Allen belts, where high-energy protons and electrons are trapped by Earth's magnetic field, as well as continuous exposure to Solar Particle Events (SPEs) and cosmic rays. When these particles strike a semiconductor lattice, they can displace atoms from their sites and create defects. Because this damage accumulates over time and is irreversible in most semiconductors (Messenger et al., 2001), the practical lifetime of a space solar cell is often determined not simply by its initial efficiency but by its degradation rate.

Power-to-mass ratio. This metric indicates how much electrical power can be generated per unit mass and is particularly important in space missions where launch mass is tightly constrained. Importantly, the denominator includes not only the absorber layer but also the substrate, encapsulation, electrode wiring, and deployment structure. Making the light-absorbing layer extremely thin provides little advantage if the supporting substrate remains heavy.

Stowed volume and deployability. In missions deploying large numbers of satellites, volume can become an even stricter constraint than mass. The key question is how compactly and densely the array can be folded into the limited volume inside the launch-vehicle fairing. This is why flexible structures that can be folded or rolled for launch and then deployed in orbit are advantageous.

Thermal-cycle tolerance. A LEO satellite alternates between sunlight and eclipse, when it passes through Earth's shadow, roughly every 90 minutes. During this cycle, device temperature can swing dramatically from approximately $-150\,^{\circ}$C to $+120\,^{\circ}$C, producing about 16 cycles per day and more than 5,000 cycles per year. Because each layer in a thin-film device has a different coefficient of thermal expansion (CTE), repeated mechanical stress develops at the interfaces. Over time, this can cause delamination and eventual device failure.

These four requirements are not independent but instead form strong trade-offs. A thicker protective layer improves radiation shielding but reduces the power-to-mass ratio, while a thin polymer substrate improves flexibility and packing density but becomes more vulnerable to atomic oxygen and ultraviolet attack. Designing a space solar cell therefore requires identifying the optimal balance among these competing factors, a problem revisited in Section 5.

2. Current Space Solar Cells and Their Limitations

2.1 III-V Multijunction Solar Cells

Most solar cells currently used in orbit are III-V compound multijunction devices, consisting of multiple layers of compound semiconductors made from Group 13 and Group 15 elements of the periodic table. A representative three-junction combination is InGaP, InGaAs, and Ge.

The purpose of stacking multiple layers is to reduce energy loss as heat. With a single absorber, photons with energy below the bandgap pass through unabsorbed, while photons with energy above the bandgap are absorbed but lose their excess energy as heat. By stacking materials with different bandgaps, the upper layers absorb short-wavelength, high-energy photons while the lower layers capture longer-wavelength, lower-energy photons. This spectral division enables efficiencies above 30% under AM0 conditions.

These devices also have an extensive reliability record. They have accumulated decades of orbital operating experience, and models for predicting radiation-induced degradation are well established (Messenger et al., 2001). As a result, the reduction in power output over the mission lifetime can be estimated quantitatively during the design stage. Since spacecraft cannot normally be repaired after launch, this predictability is far from trivial. Any new material seeking to replace III-V cells must therefore provide not only high performance but also comparable predictability.

2.2 One Single-Crystal Substrate, Three Interconnected Limitations

The limitations of III-V cells are commonly summarized as high cost, high mass, and poor flexibility or rigidity. However, these are not three independent drawbacks. They all originate from one underlying requirement: the cells must be fabricated on single-crystal substrates.

To understand why, consider why a single-crystal substrate is necessary. Multijunction cells stack semiconductors with different properties, and if the atomic spacing, or lattice constant, of one layer does not match that of the layer below, dislocations form at the interface. A dislocation is a line defect in which the crystal arrangement is misaligned, and it can trap and annihilate photogenerated charge carriers. Growing successive layers while maintaining lattice matching therefore requires a Ge or GaAs single-crystal wafer with highly ordered atomic alignment. The growth process also relies on precision equipment operating at high temperature and high vacuum, such as Metal-Organic Chemical Vapor Deposition (MOCVD) or Molecular Beam Epitaxy (MBE).

High cost. Single-crystal wafers are themselves expensive, and Ge is relatively scarce and difficult to refine. In addition, epitaxial growth is slow, has low precursor utilization, and requires costly equipment, making the manufacturing cost per unit area two to three orders of magnitude higher than that of terrestrial silicon cells

(Hoang et al., 2022). This may be acceptable for a few satellites, but the economics change dramatically when building infrastructure at the scale of thousands of spacecraft.

High mass. A wafer must be thick enough—typically around 150 μm—to avoid bending or breaking during high-temperature processing. Because Ge has a density of $5.32\ \mathrm{g/cm^3}$, such a substrate has an areal mass of roughly $800\ \mathrm{g/m^2}$. Yet the combined thickness of the active layers that actually absorb light and generate electricity is only a few μm. Thus, most of the mass is merely structural support that does not directly contribute to power generation. Figure 3 compares this architecture side by side with a perovskite device. Epitaxial lift-off techniques that remove and reuse the substrate after growth are also being studied, but the additional processing further increases cost.

Poor flexibility. Ge and GaAs are brittle materials that fracture rather than plastically deform under stress, and thicker substrates experience larger surface strain when bent. As a result, they cannot be readily folded or rolled. Real satellites typically connect flat panels with hinges and fold them like an accordion, but this occupies substantial fairing volume and introduces mechanically complex deployment systems that themselves create failure risks.

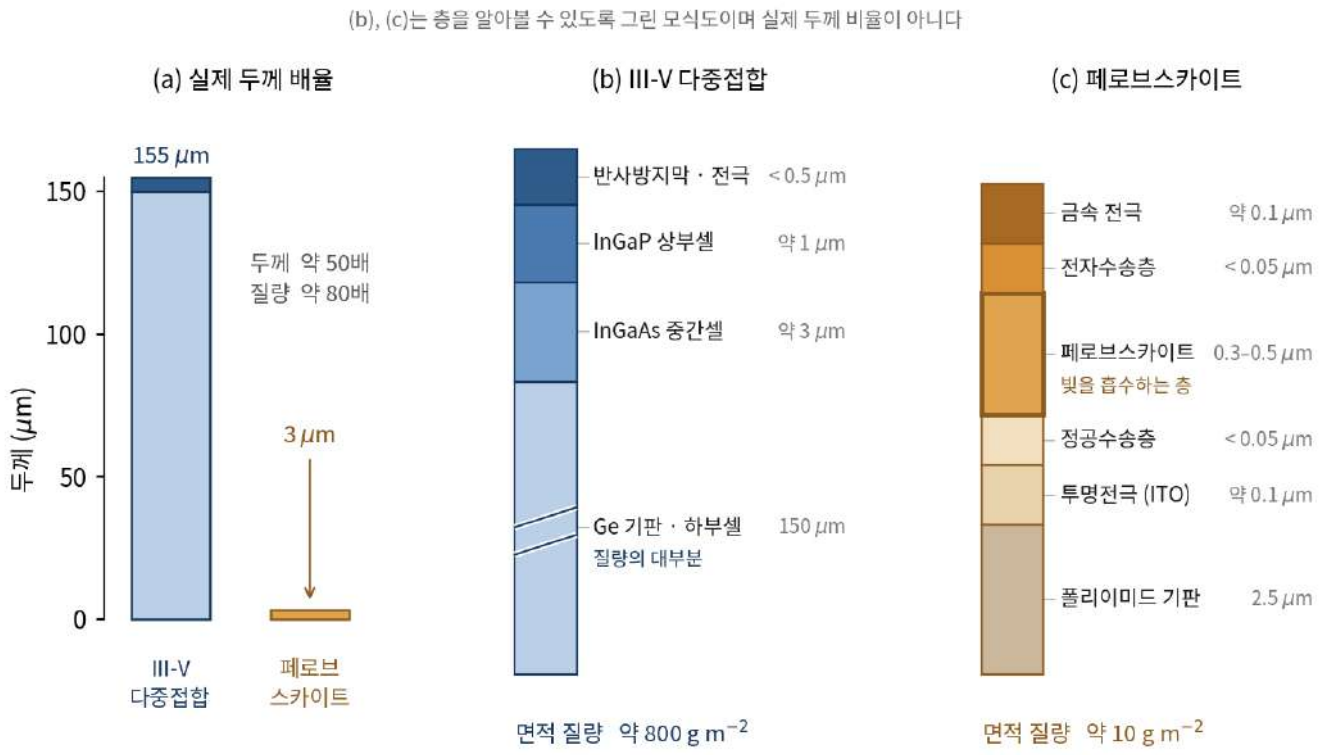


Figure 3. Comparison of the cross-sectional structures of III-V multijunction and perovskite solar cells. (a) Relative thicknesses of the two devices, (b) layer structure of the III-V device, and (c) layer structure of the perovskite device

In summary, the single requirement for a single-crystal substrate simultaneously constrains cost, mass, and form factor or flexibility. Overcoming these limitations therefore requires more than incremental process improvement; it requires a fundamentally different material platform. The properties of major candidate materials are summarized in Table 1.

Table 1. Comparison of candidate photovoltaic technologies for space applications (Song et al., 2026)

Property	III-V multijunction	Crystalline Si	Perovskite
AM0 efficiency	Above 30%	15–18%	Under development
Absorber thickness	Several μm	Above 150 μm	0.3–0.5 μm
Substrate	Ge/GaAs single crystal	Si wafer	Polyimide, etc.
Power-to-mass ratio	Approx. 5.5 W/g	0.5–2 W/g	23–30 W/g
Flexibility	None	Limited	High
Processing temperature	Above 600 °C	Above 1400 °C	Below 150 °C
Space validation	Extensive	Limited	Early stage

2.3 Why Perovskites Stand Out Among Alternative Technologies

Several thin-film photovoltaic technologies other than perovskites can be fabricated without single-crystal substrates and can support flexible devices. Why, then, are perovskites receiving especially strong attention as next-generation space photovoltaic materials? The answer becomes clear when the major competing thin-film candidates are examined one by one.

CIGS. This compound semiconductor, composed of copper, indium, gallium, and selenium, can be fabricated on flexible substrates and exhibits strong radiation resistance, so it has been considered for space applications. However, its efficiency improvement has largely stagnated, and high-vacuum deposition equipment is still required, limiting the potential reduction in manufacturing cost. It also faces a raw-material supply constraint because indium is a scarce metal with limited reserves.

CdTe. CdTe is a representative thin-film photovoltaic technology that has achieved terrestrial commercialization, but it is usually fabricated on thick glass substrates, making flexible implementation difficult. In addition, environmental regulation associated with the use of toxic cadmium becomes a concern for large-scale production intended for space use.

Organic solar cells. These are extremely light and flexible, making them highly attractive from a mass perspective. However, their absolute power conversion efficiency (PCE) is lower, and they are particularly vulnerable to ultraviolet radiation. As discussed above, the large UV fraction under AM0 conditions makes this a

serious weakness in the space environment.

Perovskites. Perovskites can simultaneously satisfy several requirements that competing candidates meet only individually. First, their efficiency has already reached a practical level: certified single-junction efficiencies are in the 27% range, while tandem structures combined with silicon have reached the 34% range, exceeding what other thin-film technologies have achieved. Second, their absorption coefficient is extremely high, allowing sufficient solar absorption with an absorber only about 300 nm thick. At this thickness, a thin polymer film only a few μm thick can provide adequate mechanical support without a single-crystal substrate. This is the basis for reported power-to-mass ratios more than an order of magnitude higher than those of conventional III-V cells (Song et al., 2026). Third, perovskites can be fabricated by low-temperature solution processes below 150°C. Flexible polymer substrates that would be damaged by high temperatures can therefore be used directly, and expensive high-vacuum equipment is not essential, allowing substantial reductions in manufacturing cost. Fourth, perovskites possess intrinsic structural resistance to radiation damage and self-healing behavior. Unlike the comparatively simple radiation hardness of materials such as CIGS, these properties arise from the distinctive electronic structure of perovskites and are examined in detail in Section 3.

Most importantly, the reason perovskites are regarded as a 'next-generation' technology is not that they dominate a single metric, but that they show the potential to satisfy several competing requirements—efficiency, mass, cost, and radiation resistance—at the same time.

3. What Is a Perovskite?

3.1 A Name for a Structure, Not a Single Material

One point must be clarified first. Perovskite is not the name of a specific compound but the name of a crystal structure. Any material whose atoms are arranged in this particular way can be called a perovskite. Without this distinction, the discussion that follows can easily become confusing.

The name originates from a mineral. In 1839, German mineralogist Gustav Rose discovered calcium titanate with the chemical formula $CaTiO_3$ in Russia's Ural Mountains and named it 'perovskite' after the Russian mineralogist Lev Perovski. As more compounds with the same structural arrangement were discovered, the term came to describe a broader family of materials.

Today, the term perovskite refers to the entire class of materials with an ABX_3-type crystal arrangement such as $CaTiO_3$; the original mineral is simply one member of this family. The structure had long been studied in oxide materials, in which oxygen occupies the X site. $BaTiO_3$ is a representative ferroelectric, and some copper oxides exhibit high-temperature superconductivity. Thus, the perovskite structure was not newly discovered for photovoltaics; it was already a familiar concept in materials science.

The materials used in solar cells are 'halide perovskites,' in which halide ions (I, Br, or Cl) occupy the X site instead of oxygen. Organic-inorganic hybrid compounds in this family were first synthesized and reported by Dieter Weber in 1978 (Weber, 1978), although the work was then purely crystallographic and unrelated to photovoltaics. The first use of such a material in a solar cell came in 2009, when Tsutomu Miyasaka and colleagues employed it as a sensitizer in a Dye-Sensitized Solar Cell (DSSC) and reported an efficiency of 3.8% (Kojima et al., 2009). Device architectures then improved rapidly: efficiencies exceeded 10% after the introduction of a solid Hole Transport Layer (HTL) in 2012, and by 2026 certified single-junction efficiencies had reached the 27% range, while silicon-perovskite tandem structures had reached the 34% range (National Renewable Energy Laboratory, 2026). This progress took less than one-third of the time required for conventional crystalline silicon to reach comparable efficiencies. Figure 4 shows the annual trend in certified efficiencies for major photovoltaic materials.

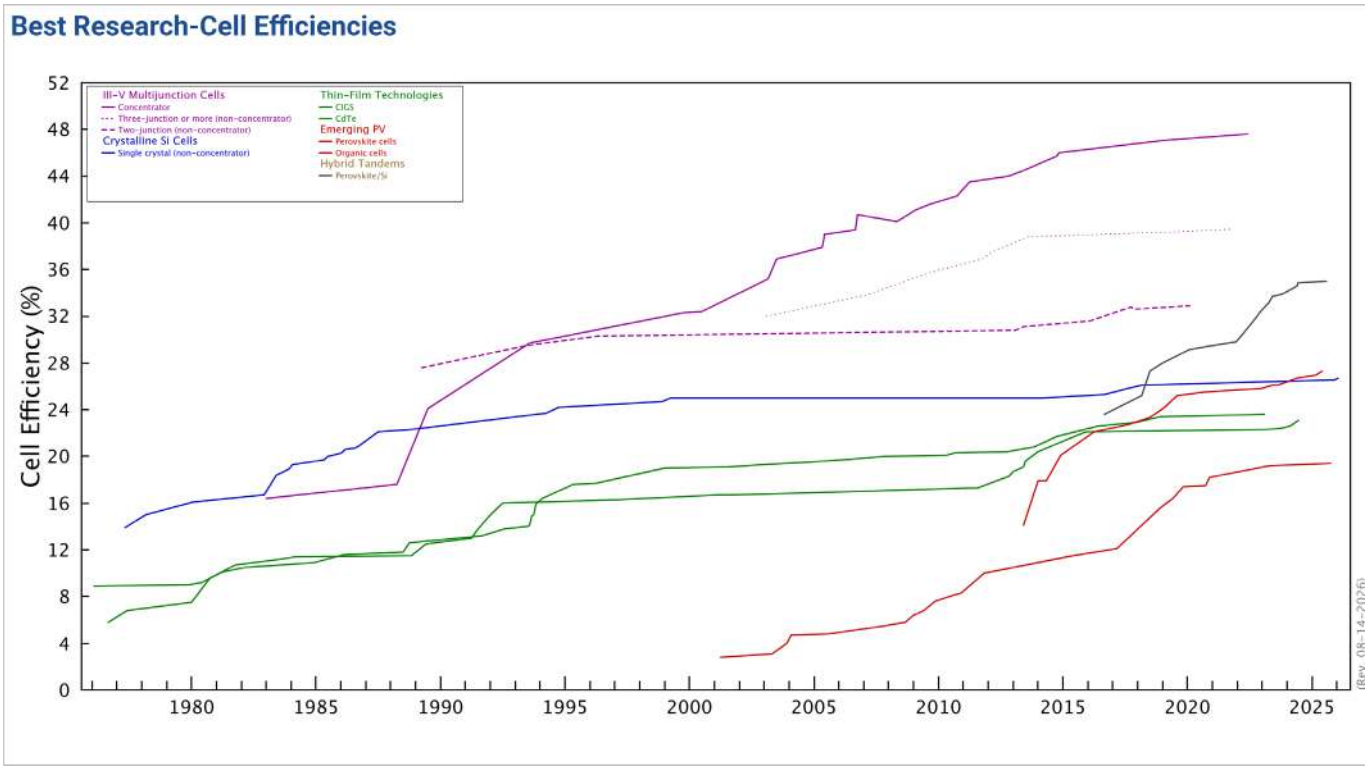


Figure 4. Historical trend in certified photovoltaic efficiency by material system

The important point here is not merely the efficiency value but the manufacturing method. The material can be fabricated simply by coating a liquid precursor solution onto a substrate and drying it at a relatively low temperature near 100℃. The resulting film is polycrystalline, with grain sizes of only a few hundred nanometers. In other words, it contains a very high density of grain boundaries.

This behavior is difficult to reconcile with conventional semiconductor intuition. In silicon and III-V semiconductors, crystal perfection strongly determines performance; as discussed in Section 2.2, even a single dislocation, a tiny line defect, can annihilate charge carriers and sharply reduce efficiency. Grain boundaries are far larger disruptions of crystal order than dislocations, yet perovskites continue to operate at high efficiency despite being full of them. The reason is not that the defects are absent, but that 'even when defects exist, they cause relatively little harm to device performance.' The fundamental explanation lies in the unusual electronic structure of these materials, discussed in detail in Section 3.3.

3.2 ABX_3 Structure and Compositional Design

The chemical formula of a halide perovskite is ABX_3. Each of the three letters corresponds to a different type of ion.

The A site is occupied by a relatively large cation. Common choices include the organic cations methylammo-

nium (MA^+) and formamidinium (FA^+), as well as the inorganic cation cesium (Cs^+). The B site is occupied by a divalent metal cation, usually lead (Pb^{2+}); tin (Sn^{2+}) is often used in efforts to avoid lead. The X site is occupied by a halide anion such as iodide (I^-), bromide (Br^-), or chloride (Cl^-).

These ions form the following structure. One B ion is surrounded by six X ions to form an octahedron. The octahedra connect by sharing corners to build a three-dimensional framework, while the A ions occupy cavities within that framework (Figure 5). Thus, the A ion is not a framework-forming element; rather, it fills the cavity and helps prevent the structure from collapsing.

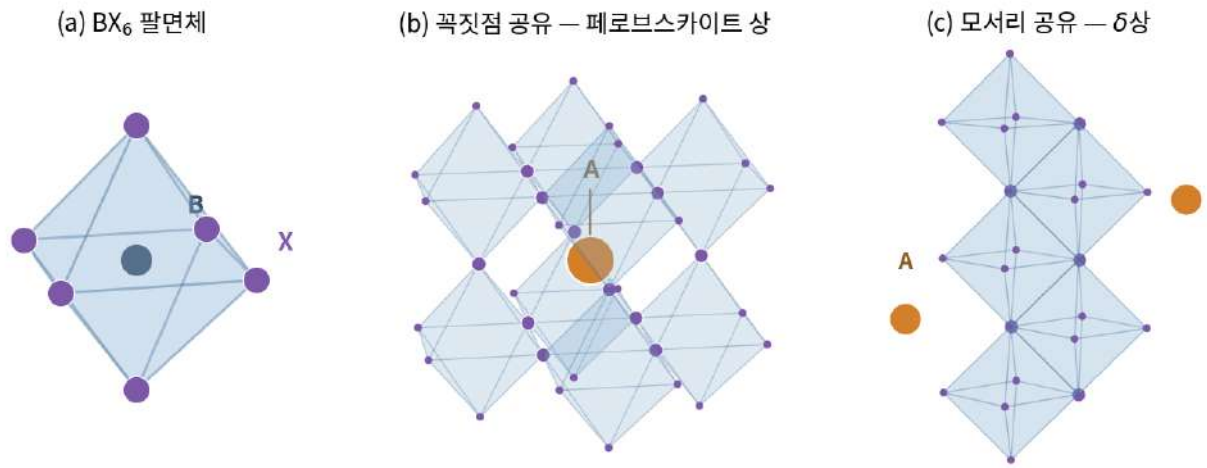


Figure 5. Crystal structure of an ABX_3 perovskite. (a) A single BX_6 octahedron, (b) a three-dimensional corner-sharing octahedral framework with A-site ions occupying the cavities, (c) the δ phase, in which the octahedra rearrange into one-dimensional edge-sharing chains

A problem arises when the size of the A ion does not match the cavity. If the A ion is too small, it cannot adequately fill the site, causing the framework to tilt inward. As the octahedra tilt, the atomic arrangement can eventually change, altering the material's light-absorption properties. Conversely, if the A ion is too large, it expands the framework and again destabilizes the structure. The sizes of all three ions therefore need to be mutually compatible for the perovskite structure to remain stable.

The Goldschmidt tolerance factor provides a single numerical measure of this size compatibility. It is denoted by t and is defined as follows.

$$t = \frac{r_A + r_X}{\sqrt{2}\,(r_B + r_X)}$$

r_A, r_B, and r_X are the ionic radii of the species occupying the respective sites. The numerator represents the distance associated with the A-X pair, while the denominator is a geometrically scaled B-X distance. When

the two are well matched, t approaches 1. In other words, t is a score indicating how well the ionic sizes fit together.

One caution is necessary. MA^+ and FA^+ are molecular ions composed of several atoms rather than single spherical atoms, so their apparent size depends on orientation. At room temperature, however, they rotate rapidly and behave approximately like spherical particles on average. Their effective radii are therefore estimated from the volume swept out during rotation (Kieslich et al., 2014).

Table 2. Tolerance factors and phase stability of representative compositions

Composition	Tolerance factor t	Phase stability at room temperature	Key feature
$MAPbI_3$	Approx. 0.91	Stable	Vulnerable to heat and vacuum
$FAPbI_3$	Approx. 0.99	Metastable	Favorable narrow bandgap
$CsPbI_3$	Approx. 0.84	Unstable	Thermally robust but prone to phase transition
$Cs_{0.2}FA_{0.8}PbI_3$	Approx. 0.96	Stable	t tuned through cation mixing

Three observations can be drawn from Table 2.

First, the instability of $CsPbI_3$ becomes clear numerically. Because Cs^+ is relatively small, it does not fully occupy the A-site cavity, lowering t to about 0.84. At room temperature, the arrangement changes: the three-dimensional corner-sharing framework collapses and the octahedra reorganize into a one-dimensional edge-sharing chain structure. This state is called the δ phase, in which the bandgap becomes wider and light absorption deteriorates severely, making the material unsuitable for solar cells (Steele et al., 2019). Thus, although all-inorganic compositions are thermally robust, their inability to preserve the desired phase is a fundamental challenge, revisited in Section 5.1.

Second, mixing multiple cations can shift t toward a desired value. Combining small Cs with larger FA adjusts the average effective size and can raise t to around 0.96. This means that compositional design can be treated as a quantitative target rather than as intuition alone; in fact, most recent high-efficiency devices use mixtures of two or three A-site cations.

Third, the X site can also be tuned. Replacing I with Br decreases the ionic radius, thereby increasing t, while simultaneously widening the bandgap. A single change in X-site composition can therefore tune both structural stability and optical properties.

However, t should not be treated as a universal rule. The commonly cited range $0.8 < t < 1.0$ originates from oxide perovskites and cannot be applied directly to halide systems. In halide perovskites, the experimentally observed range for a stable cubic phase is much narrower, approximately $t = 0.9$ to 1.0 (Travis et al., 2016).

The tolerance factor is therefore best used as an initial screening tool for a large compositional space, with final judgments confirmed by computation or experiment.

3.3 Optoelectronic Properties and Defect Tolerance

We now examine what properties this structure provides as a photovoltaic absorber and return to the question raised in Section 3.1.

It absorbs light extremely efficiently. Perovskites are direct-bandgap semiconductors, meaning that electrons can be promoted directly across the bandgap when photons are absorbed. This makes optical absorption highly efficient. Their absorption coefficient is much larger than that of silicon, so a layer only about 300 nm thick can absorb roughly 95% of visible light—about one three-hundredth the thickness of a human hair. Silicon, by contrast, is an indirect-bandgap semiconductor, so electronic transitions require assistance from lattice vibrations or phonons. Absorption is therefore weaker, and hundreds of μm are needed to achieve the same function. This nearly three-orders-of-magnitude difference in absorber thickness underlies the low-mass advantage discussed in Section 2.3.

Charge carriers can travel long distances. Photogenerated charges contribute to current only if they reach the electrodes. The average distance they can travel before disappearing is called the diffusion length. Even in solution-processed polycrystalline films, this distance can exceed 1 μm (Stranks et al., 2013), and values above 175 μm have been reported in large crystals (Dong et al., 2015).

The significance becomes clear when these two properties are considered together. An absorber thickness of 300 nm is sufficient, while charge carriers can travel more than three times farther than that. In other words, charge generated throughout the absorber can, in principle, reach an electrode before recombining. These two properties rarely coexist. Materials that absorb light strongly often contain many defects that rapidly eliminate carriers, while materials in which carriers survive for long periods often absorb more weakly. Absorbers that satisfy both conditions are uncommon.

Charge pairs do not require a special dissociation mechanism. When light is absorbed, electrons and holes are generated as bound pairs known as excitons. If the pair remains bound, it cannot contribute efficiently to electrical current, so the electron and hole must separate. The energy required for this separation—the exciton binding energy—depends on the material. In perovskites, this value is lower than the thermal energy already available at room temperature (Romano et al., 2022). Thus, thermal motion alone can spontaneously dissociate the pair into freely mobile carriers. This contrasts with organic solar cells, which require finely intermixed donor and acceptor materials to separate excitons.

The bandgap can be tuned to a desired value. As discussed in Section 3.2, changing the X-site composition

modifies the bandgap. Replacing I with Br progressively widens the bandgap across a range of roughly 1.2 to 2.3 eV, which includes the 1.7–1.8 eV range required for the top cell in tandem structures combined with silicon or another perovskite.

Defects are not necessarily harmful. We can now return to the question from Section 3.1: why does the material work despite containing so many defects?

Consider silicon and III-V semiconductors first. When atoms bond, the available electronic states divide into two broad categories: bonding states, which stabilize the bond, and antibonding states, which destabilize it. Collections of bonding states form the valence band, while collections of antibonding states form the conduction band. If an atom is removed and a bond is broken, an electron left without a normal bonding environment can create a state between the two bands, often deep within the bandgap. This is called a deep trap. A deep trap can capture both an electron and a hole, allowing the two to recombine and disappear. A single defect can therefore eliminate many photogenerated carrier pairs.

Perovskites reverse this conventional arrangement. The top of the valence band is not primarily bonding in character but instead contains antibonding hybridization between Pb 6s and I 5p orbitals, placing it at unusually high energy (Yin et al., 2014). The bottom of the conduction band is dominated by Pb 6p orbitals.

This inverted arrangement is crucial. The defects that occur most frequently—iodine vacancies, A-site vacancies, and interstitial ions—tend to create states near the band edges rather than in the middle of the bandgap. These are called shallow levels, and thermal energy at room temperature is sufficient to release trapped charges quickly (Walsh and Zunger, 2017). Because the charge is not retained, it is not readily lost to recombination. Figure 6 illustrates the difference between the two cases.

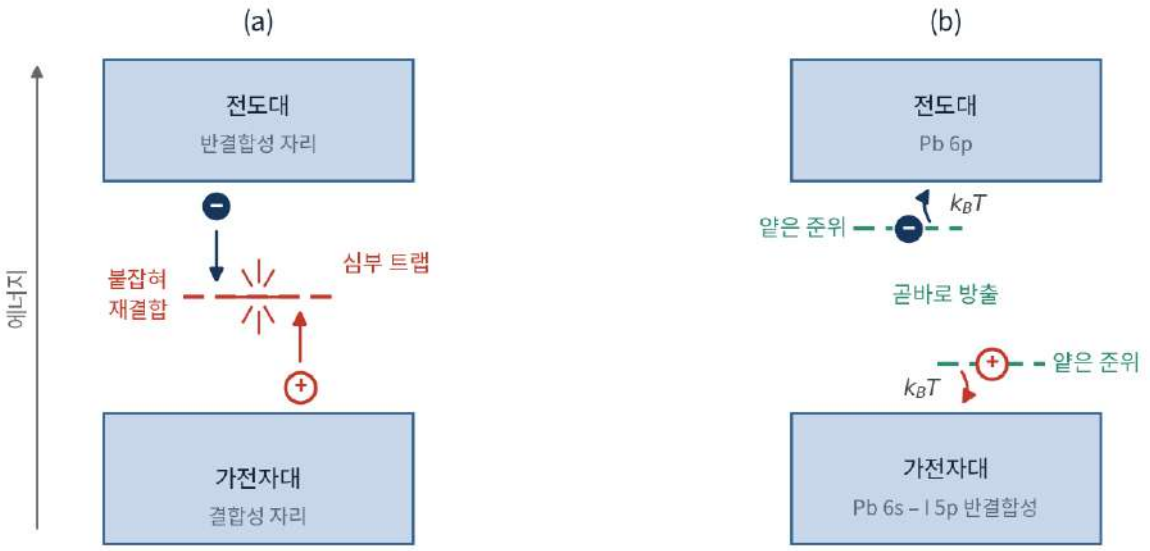


Figure 6. Comparison of defect-induced energy levels. (a) Deep traps and carrier recombination in silicon and III-V semiconductors, and (b) shallow levels in perovskites with carrier release driven by thermal energy (k_BT)

Two additional effects reinforce this behavior. First, perovskites have a relatively large dielectric constant. Charged defects create local electric fields that can attract passing carriers, but a large dielectric constant screens these fields strongly and reduces their influence. Second, the carriers are relatively light. Their small effective masses allow them to move rapidly past defect sites (Brandt et al., 2015).

The combined result of these three features is called defect tolerance. It does not mean that the material contains no defects; rather, it means that defects cause comparatively little loss in performance. This is why a rough, solution-processed polycrystalline film can still reach efficiencies around 27%.

3.4 Why Perovskites Are Radiation Tolerant

This property is also the starting point for understanding the material's suitability for space, because radiation damage ultimately produces the same type of point defects.

High-energy charged particles travel through the orbital environment. The main components are protons and electrons trapped by Earth's magnetic field, supplemented by particles emitted by the Sun and by cosmic rays. When these particles strike a semiconductor, they lose energy in two fundamentally different ways, and this distinction is central to the discussion that follows.

Ionizing Energy Loss (IEL). As a particle passes through the material, it excites electrons. Large numbers of electron-hole pairs are generated, but the atoms remain in their original lattice sites. The crystal structure itself

is therefore largely unchanged.

Non-Ionizing Energy Loss (NIEL). The particle collides directly with atoms in the lattice. If the collision transfers sufficient energy, an atom is displaced from its site, leaving a vacancy behind while the displaced atom occupies an interstitial position. This vacancy-interstitial pair is called a Frenkel pair. In this case, the crystal structure is physically altered.

In conventional space solar cells, NIEL is the principal source of permanent damage (Messenger et al., 2001), because Frenkel pairs become persistent deep traps. IEL is generally treated as less important for long-term structural damage because it primarily creates transient electronic excitations without displacing atoms. Based on this framework, the lifetime of a space solar cell is often predicted using the Displacement Damage Dose (DDD). Although incoming particles differ in type and energy, their ability to displace lattice atoms can be calculated and summed into a single damage metric, which can then be used to extrapolate orbital lifetime from ground-based irradiation experiments.

We can now apply the defect tolerance discussed in Section 3.3. Both the vacancies and interstitials produced by NIEL are point defects. In silicon and III-V semiconductors, these defects create deep traps, so efficiency declines progressively as the irradiation dose accumulates. In perovskites, however, the same point defects predominantly produce shallow levels (Yin et al., 2014; Walsh and Zunger, 2017). The material does not prevent defects from forming; rather, many of the defects that form are less effective at trapping charge carriers. Radiation resistance is therefore not an unrelated special property; it is a direct consequence of defect tolerance.

Experiments support this interpretation. Lang and colleagues irradiated devices with 68 MeV protons to doses roughly one thousand times higher than the damage threshold of crystalline silicon (Lang et al., 2016). Only the short-circuit current decreased substantially, while the open-circuit voltage and fill factor remained almost unchanged. Because these latter two parameters are sensitive to carrier recombination, their stability indicates that newly formed defects did not act as efficient pathways for charge loss. Electron irradiation produced similar results: more than 92% of the initial efficiency was retained after substantial exposure to 1 MeV electrons (Song et al., 2026), and proton tolerance has also been reported in all-perovskite multijunction architectures (Lang et al., 2021).

The damage can partially disappear on its own. An even more striking phenomenon occurs after irradiation stops: some of the lost performance returns. This is not merely a consequence of reduced initial damage; rather, some damage that has already formed can be partially reversed, a behavior rarely observed in conventional space photovoltaic materials.

In the experiments by Lang and colleagues, the short-circuit current recovered after irradiation ended (Lang

et al., 2016), and similar behavior was later reported by several groups. Afshari and colleagues irradiated a triple-halide device with protons and observed an immediate reduction in performance, but after two months the device returned to approximately its pre-irradiation level (Afshari et al., 2023). Durant and colleagues treated recovery as a controlled experimental variable rather than an accidental observation: they concentrated proton damage at a specific depth in the device and then intentionally induced electronic ionization to test whether recovery could be triggered (Durant et al., 2021). The most direct evidence comes from the dual-irradiation experiments of Kirmani and colleagues (Kirmani et al., 2024). They first exposed the device to slow 0.06 MeV protons. Because slow protons stop near the surface, a relatively large fraction of their energy is deposited as NIEL, reducing the efficiency to 74% of its initial value. They then added 1.0 MeV protons, keeping the displacement component relatively similar while increasing IEL by tens of times. The efficiency rose again, recovering to 83% in the NIP architecture and 85% in the PIN architecture. In other words, one component of radiation helped reverse damage caused by another. The major reports to date are summarized in Table 3.

Table 3. Major reports of recovery behavior following proton irradiation

Study	Irradiation condition	Damage	Recovery
(Lang et al., 2016)	68 MeV protons	Reduced short-circuit current; open-circuit voltage and fill factor maintained	Short-circuit current recovered after irradiation ended
(Afshari et al., 2023)	Proton irradiation; triple-halide device	Reduced photovoltaic performance	Returned to the initial level after two months
(Durant et al., 2021)	Depth-targeted proton irradiation	Localized displacement damage	Recovery induced by electronic ionization
(Kirmani et al., 2024)	Dual irradiation with 0.06 MeV followed by 1.0 MeV protons	Efficiency decreased to 74%	Recovered to 83–85%

Two mechanisms appear to act together to produce this recovery (Figure 7).

First, the lattice is sufficiently flexible for displaced ions to return to their original sites. Perovskites possess a relatively soft ionic crystal lattice. Because the bonding is weak, the activation energy required for ions to move is also small; the energy required for an iodide ion to migrate through vacancies in the lattice is several times lower than in conventional silicon-based semiconductors (Eames et al., 2015). At this energy scale, ions can move spontaneously at room temperature without high-temperature annealing. Consider a Frenkel pair created by radiation: the displaced iodine ion and the vacancy left behind are initially close to one another. Only

a short migration distance is needed for the iodine to refill the vacancy, at which point both defects disappear simultaneously. Calculations indicate that recombination of the pair is energetically much more favorable than allowing the two defects to separate farther apart (ACS, 2026).

Second, ionizing damage can paradoxically assist recovery. Kirmani and colleagues proposed that, in addition to the displacement damage created by non-ionizing energy loss, ionizing energy loss induces localized lattice vibrations, and the vibrational energy can help drive displaced atoms back toward their original sites (Kirmani et al., 2024). In effect, the radiation itself effectively provides highly localized annealing at damaged regions. This mechanism is especially pronounced in materials with soft lattices and strong electron-phonon coupling. Perovskites also have low thermal conductivity, so locally generated heat is not rapidly dissipated and can contribute efficiently to microscale annealing (Kirmani and Sellers, 2025). Such behavior is not typically observed in rigid covalently bonded semiconductors.

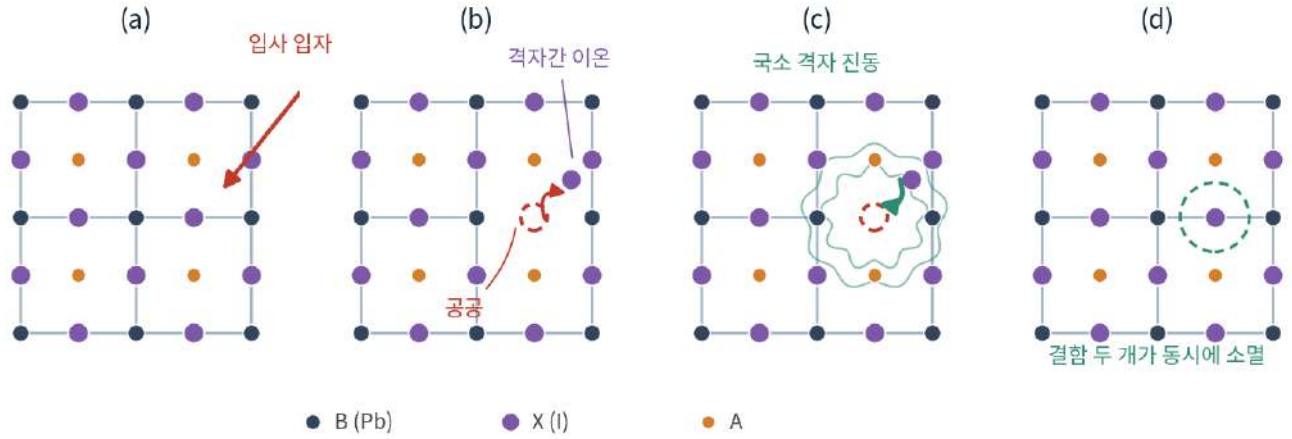


Figure 7. Formation and self-healing of a Frenkel pair. (a) Incidence of a high-energy particle, (b) formation of a vacancy and an interstitial ion, (c) short-range ion migration assisted by local lattice vibrations, (d) simultaneous annihilation of the two defects

However, healing occurs only under certain conditions. Self-healing requires both an environment in which ions can move and a driving force that promotes that motion; temperature, illumination, and the ionizing component of radiation can provide this driving force. In other words, self-healing is not a fixed property that the material expresses under all circumstances, but a condition-dependent phenomenon, and it is particularly sensitive to temperature. To refill a vacancy, an ion must overcome an activation-energy barrier, and as temperature decreases, the fraction of ions able to cross that barrier falls exponentially. Indeed, comparing a sunlit interval at about 390 K with an eclipse interval at about 120 K yields a difference in recovery rate spanning more than ten orders of magnitude. During the cryogenic eclipse interval, recovery can therefore be considered

effectively halted.

The same limitation appears under repeated damage. In experiments alternating light and darkness, lattice expansion and bandgap shifts were reversible, but the width of the diffraction peak continued to broaden with each cycle (Alberti et al., 2025). This indicates that the crystal gradually loses order and leaves behind increasingly disordered regions. Other studies have reported that when damage is highly concentrated by large irradiation doses, efficiency does not fully recover even when both heat and light are applied (Sharma et al., 2026). Thus, each cycle generates both reversible and irreversible components: the former can disappear, while the latter gradually accumulates. Table 4 summarizes this distinction.

Table 4. Reversible and irreversible components under repeated damage

Process	Time scale	Behavior under repetition
Frenkel-pair recombination	Very fast (milliseconds–seconds)	Remains active over repeated cycles
Ion redistribution and remixing	Slow (tens of hours–months)	Effectively stops at low temperature
Loss of crystallinity and amorphization	Irreversible	Accumulates with each cycle
Organic-cation decomposition	Irreversible	Increases monotonically
Transport-layer and electrode degradation	Irreversible	Increases monotonically

This distinction also explains why reported 'recovery' times in the literature range from milliseconds to as long as two months, spanning nearly ten orders of magnitude (Song et al., 2026; Alberti et al., 2025; Afshari et al., 2023). These measurements are not necessarily contradictory; rather, physically distinct processes are often grouped under the same term. A displaced ion returning to a nearby vacancy may occur almost instantaneously, whereas redistributing and remixing ions across larger distances takes much longer, and restoring a crystal that has already become structurally disordered is essentially impossible at room temperature.

Self-healing should therefore be understood not as an automatic safeguard that always protects the device, but as a mechanism that functions only under specific conditions. The design goal should not be to rely on healing itself, but to keep the device within an operating regime in which healing remains active. This also means that thermal management becomes part of the radiation-mitigation strategy (Kirmani and Sellers, 2025).

4. Case Studies: Orbital Demonstrations

4.1 What Ground Tests Alone Cannot Reveal

All irradiation experiments discussed in Section 3 were performed using ground-based accelerators that exposed samples to protons or electrons. In actual orbit, however, vacuum, extreme temperature variations, atomic oxygen, ultraviolet radiation, and energetic particles act simultaneously. In particular, if self-healing depends exponentially on temperature as discussed in Section 3.4, then in LEO it repeatedly starts and stops as the spacecraft alternates between sunlight and eclipse every 90 minutes. Such coupled effects cannot be reproduced by ground tests that apply one stress at a time, which is why direct orbital demonstrations remain essential.

4.2 From the Stratosphere to Low Earth Orbit

The earliest demonstrations began in the stratosphere. In 2018, $MAPbI_3$ devices were carried by balloon to an altitude of 32 km, where 64% of their initial efficiency remained after more than three hours of exposure (Song et al., 2026). The cause of the loss was particularly revealing: it was attributed not primarily to radiation, but to low-temperature phase transitions and interface recombination. This result shows that the phase-stability problem discussed in Section 3.2 can emerge before radiation damage becomes the dominant failure mode.

In 2020, a sounding-rocket flight carried both perovskite and organic solar cells to an altitude of approximately 240 km (Reb et al., 2020). Electrical characteristics measured during flight showed no detectable radiation-induced damage over the 360-second measurement interval, with only a slight reduction in power density. Although brief, this was one of the first demonstrations of actual operation in the space environment.

From 2020 to 2021, a ten-month exposure experiment was conducted on the International Space Station (ISS) (Song et al., 2026). This experiment is discussed separately in the following subsection.

A CubeSat demonstration followed in 2024. Big Red Sat-1, a 1U CubeSat designed to evaluate perovskite solar cells in LEO, was launched in March 2024 on a resupply mission to the International Space Station. In October 2025, perovskite solar cells were also carried aboard HTV-X1, an uncrewed cargo spacecraft of the Japan Aerospace Exploration Agency (JAXA), as part of an ongoing plan to expose them to the space environment for roughly two months while measuring current-voltage characteristics.

4.3 Ten-Month Exposure Test on the International Space Station

Among orbital tests conducted to date, the International Space Station experiment most directly addresses the questions raised in Section 3 (Song et al., 2026).

Conditions. Encapsulated $MAPbI_3$ samples were exposed for ten months at an altitude of 408 km. At this orbit, sunlight and eclipse alternate every 90 minutes, so the samples experienced temperature swings exceeding 200 °C during each orbit while remaining continuously exposed to radiation. In other words, the temperature-sensitive self-healing conditions discussed in Section 3.4 were repeated thousands of times.

Results. After retrieval, the samples showed partial degradation, including yellowing of the glass and formation of PbI_2. However, when they were returned to Earth and illuminated again, much of the structure recovered, defect density decreased, and carrier lifetime increased.

Significance. The results indicate that the capacity for recovery itself remained after ten months of repeated damage-and-healing cycles. Using the terminology of Table 4, irreversible components such as PbI_2 formation and glass discoloration clearly accumulated, yet the mechanisms responsible for restoring the reversible components did not completely disappear. Thus, for the question raised in Section 3.4—whether recovery eventually collapses under repetition—the experiment provides a partially positive answer, at least over a ten-month period in LEO.

Limitations. These results should not be overinterpreted. First, recovery was observed after the samples were returned to Earth and illuminated at room temperature; it was not measured in real time in orbit. Second, the samples were encapsulated, which reduced their direct exposure to atomic oxygen and ultraviolet radiation. Third, the sample areas were laboratory-scale and cannot be assumed to represent the behavior of full-sized arrays.

These limitations are shared by other orbital demonstrations. Tests conducted so far have generally been small in scale, short in duration, and limited to research-sized devices. They should therefore be viewed as early-stage validation rather than proof of practical readiness. Nevertheless, they are important because data are beginning to accumulate that allow direct comparison between ground-based testing and actual orbital measurements, and substantially more such experiments are needed.

5. The Future of Perovskites for Space Power

For perovskites to become established as a space power source, progress must occur simultaneously in materials, validation, and system engineering. The International Space Station case discussed in Section 4 provides a compact illustration of how these three axes interact in practice.

5.1 Materials: Trade-Offs Associated with Self-Healing

At the material level, the most difficult issue is that the very property enabling recovery can also drive degradation.

The first issue is ion migration. Section 3.4 identified ionic motion as the origin of healing, but the same mobility can also cause problems. Under an applied voltage, ions drift toward one side of the device, causing the current-voltage curve to depend on scan direction (hysteresis). In mixed-halide compositions, I and Br can also separate under illumination, producing photo-induced phase segregation (Hoke et al., 2015). This creates local variations in bandgap and is particularly problematic for compositions intended for tandem top cells. Yet making the lattice more rigid to suppress ion migration also sacrifices defect tolerance and self-healing. Healing requires only short-range motion on the scale of a Frenkel pair, whereas degradation is driven by long-range migration across a significant fraction of the device thickness. The goal, therefore, is not to eliminate mobility entirely but to separate these two length scales.

The second issue is phase stability. The organic cations MA^+ and FA^+ have relatively weak bonds and are prone to decomposition under heat, vacuum, and ultraviolet exposure; once decomposition products form, they do not spontaneously reassemble. All-inorganic compositions may therefore be advantageous for repeated recovery. However, as discussed in Section 3.2, $CsPbI_3$ has a tolerance factor as low as 0.84 and transforms into the δ phase at room temperature, so the desired phase itself is unstable (Steele et al., 2019). Strategies such as Br alloying and surface treatment are being explored to stabilize the phase, but adding Br widens the bandgap and reduces the efficiency potential of a single-junction cell.

The fact that low-temperature phase transition rather than radiation was the dominant loss mechanism in the 2018 stratospheric balloon experiment shows that this concern is not merely theoretical. In summary, recovery capability, ion-migration-induced degradation, and phase stability all emerge from the same underlying feature—a soft ionic lattice—so changing one tends to shift the others as well. This mirrors the situation in Section 2.2, where three limitations of III-V cells all originated from the single requirement of a single-crystal substrate.

5.2 Validation: What Should Be Measured and How?

Section 2.1 argued that any new material replacing III-V technology must provide not only high performance but also predictability. The current testing framework does not yet provide this.

First, repeated-cycle testing is largely absent. Most irradiation experiments to date examine only one irradiation event followed by one recovery period. Actual LEO operation, however, is inherently repetitive. In sunlight, the temperature rises and recovery proceeds; during eclipse, the temperature falls and recovery slows or stops, while radiation exposure continues throughout both phases. Thermal cycling and radiation damage are therefore coupled variables rather than independent stresses. Conventional tests that apply the two separately cannot reveal this interaction, so testing protocols must apply them simultaneously and repeat irradiation-recovery cycles multiple times (Kirmani et al., 2022). This is precisely why the International Space Station experiment is so valuable: it is one of the few datasets to have experienced the full repeated orbital condition.

However, as noted in Section 4.3, no real-time measurements were taken in orbit, so the detailed evolution during the cycles remains unknown.

Second, reporting methods are not standardized. Many papers report only proton energy and particle fluence, which is insufficient for direct conversion to the Displacement Damage Dose described in Section 3.4. Without that conversion, results cannot be compared consistently across studies or used reliably to estimate orbital lifetime (Kirmani et al., 2022). For perovskites to become as predictable as III-V materials, this reporting framework must be standardized.

Third, the term 'recovery' itself must be used more precisely. As shown in Table 4, millisecond-scale Frenkel-pair recombination and ion remixing over periods of months are physically distinct phenomena but are often described using the same word. Future reports should explicitly state which stage or mechanism of recovery is being measured.

5.3 Systems: Challenges Beyond the Absorber Layer

Transport layers and electrodes. Above and below the absorber are charge-transport layers and electrodes. Organic transport layers can undergo molecular-chain scission under ultraviolet radiation and energetic particles, while metal electrodes can react with iodine migrating out of the absorber or diffuse into the absorber themselves through metal migration (Domanski et al., 2016). The self-healing mechanisms of the perovskite absorber do not repair these changes. Even if the absorber fully recovers, the device cannot function if the surrounding layers have degraded. Recent studies therefore suggest that the absorber itself may not be the component that ultimately limits operational lifetime (Byun et al., 2026).

Atomic oxygen and thermal cycling. LEO contains atomic oxygen, which erodes surfaces when it strikes the spacecraft's ram-facing side. Without protective coatings, organic layers and metal electrodes can be damaged rapidly. Inorganic barriers have been reported to reduce this erosion substantially (Seid et al., 2024). Thermal cycling is another concern. Because each layer has a different coefficient of thermal expansion, temperature changes generate stress at interfaces. The interfacial fracture toughness of perovskite thin films is relatively low among semiconductor films, making them prone to delamination (Rolston et al., 2016).

Encapsulation: both a solution and a new constraint. Encapsulation addresses the two problems above, but immediately reintroduces the power-to-mass issue. Reported power-to-mass ratios for perovskite photovoltaics reach 23 W/g (Kaltenbrunner et al., 2015), and laboratory demonstrations have reported values as high as 50 W/g (Song et al., 2026), but these values generally refer to unencapsulated devices. If thick polymer films are attached to both sides, their mass is added directly to the denominator, potentially erasing much of the advantage over III-V cells at about 5.5 W/g. By contrast, inorganic barriers only tens of nanometers thick can be deposited by Atomic Layer Deposition (ALD) with negligible mass penalty, but such layers are brittle

and prone to cracking. Hybrid structures alternating these two types of barriers are therefore being investigated. Consequently, any literature value for W/g should always be interpreted together with the encapsulation condition.

The fact that the International Space Station samples were encapsulated should be reconsidered in this context. Encapsulation helped them survive ten months, but the experiment did not answer how much mass penalty that encapsulation would impose in a real mission. Only when materials, validation, and system engineering advance together can perovskites become credible power-source candidates for orbital infrastructure.

Conclusion

This study confirms that halide perovskites are promising candidates for next-generation space photovoltaic power sources. As demand for artificial intelligence and cloud computing grows, efforts are emerging to move part of data infrastructure into orbit. Because orbital systems cannot rely on a terrestrial grid, they must generate power locally. The III-V multijunction cells currently used in space have proven performance and reliability, but the single requirement that they be fabricated on single-crystal substrates simultaneously constrains cost, mass, and form factor, making them less suitable for missions involving large satellite constellations or very large deployable arrays.

Perovskites offer a way around these constraints. They do not require single-crystal substrates, can absorb sunlight effectively in a layer only about 300 nm thick, and can be fabricated by low-temperature solution processing. Most importantly, their radiation resistance is not an unrelated special property but a direct extension of their defect tolerance. Because the valence band is formed by antibonding hybridization between Pb 6s and I 5p orbitals, point defects tend to create electronic levels near the band edges rather than deep inside the bandgap—and radiation damage also ultimately produces point defects. Self-healing adds another layer of resilience. In the soft ionic lattice, iodide ions can move with low activation energy, so a displaced ion can return to a vacancy and eliminate both defects at once, while localized lattice vibrations generated by the ionizing component of radiation can assist this process.

The ten-month International Space Station exposure experiment is especially significant because it showed that recovery capability persisted even after repeated exposure to temperature swings exceeding 200 °C per orbit together with continuous radiation. At the same time, it confirmed that recovery is not unconditional. Because the process depends exponentially on temperature, it effectively stops during the cold eclipse phase of LEO, and under repeated cycling irreversible components such as loss of crystallinity gradually accumulate even while reversible components are removed. Self-healing should therefore be treated not as a permanent safeguard but as a mechanism that operates only within a specific range of conditions. The design goal is to keep the device within that operating window.

For perovskites to become practical power sources for orbital infrastructure, progress must be made simultaneously in materials, validation, and system engineering. Required advances include compositional designs that preserve short-range healing while suppressing long-range ion migration, test protocols that simultaneously apply thermal cycling and radiation while repeatedly alternating damage and recovery, and lightweight encapsulation technologies that block atomic oxygen and ultraviolet radiation. If these foundations can be established, perovskites could complement conventional III-V cells and play an important role in addressing the power challenge of large-scale orbital infrastructure.

Example Questions

Students may also pursue a research topic of their choice beyond the questions listed below.

Q1. Materials Design: Self-healing in perovskites is enabled by the high mobility of iodide ions, yet the same property can also cause performance degradation through ion migration. Considering that rigidifying the lattice may suppress both effects, propose a design strategy that preserves self-healing while selectively reducing long-range ion migration.

Q2. Repeated Damage: During repeated irradiation and recovery cycles, reversible damage can heal whereas irreversible changes such as loss of crystallinity accumulate. Propose, with justification, how composition, interfaces, or operating conditions should be controlled to slow the accumulation of irreversible damage.

Q3. Test Design: Most previous studies observe only a single irradiation event followed by a single recovery period. Design an experiment that reproduces the approximately 90-minute repeated cycles encountered in LEO, specifying the applied stresses, measured quantities, and criteria for evaluating recovery and degradation.

Q4. System Application: If perovskite solar cells were to be introduced first in a space mission, what type of mission would be most appropriate? Justify your choice based on orbit, required mission lifetime, power scale, and cost constraints.

References

2019. ASTM E490-00a: Standard Solar Constant and Zero Air Mass Solar Spectral Irradiance Tables. ASTM International, West Conshohocken, PA.

2026. Finite-Temperature Evolution of Frenkel Defects in Hybrid Perovskites: Healing and Lead-Methylammonium Antisite Pairs. ACS Applied Materials & Interfaces **18**:22071.

Afshari, H., S. A. Chacon, S. Sourabh, T. A. Byers, V. R. Whiteside, R. Crawford, B. Rout, G. E. Eperon, and I. R. Sellers. 2023. Radiation Tolerance and Self-Healing in Triple Halide Perovskite Solar Cells. APL Energy **1**:026105.

Alberti, A., S. Valastro, E. Nonni, F. Matteocci, L. Cinà, A. Di Carlo, and A. La Magna. 2025. Resilience to Demixing and Phase Segregation in Perovskite Solar Cells under Light–Dark Cycles and Temperature. ACS Energy Letters **10**:2259–2267.

Brandt, R. E., V. Stevanović, D. S. Ginley, and T. Buonassisi. 2015. Identifying Defect-Tolerant Semiconductors with High Minority-Carrier Lifetimes: Beyond Hybrid Lead Halide Perovskites. MRS Communications **5**:265–275.

Byun, J., S.-P. Yoon, B. Park, et al. 2026. Verification and Mitigation of Proton-Induced Non-Ionizing Damage in Perovskite Solar Cells for Space Applications. Advanced Energy Materials .

Domanski, K., J.-P. Correa-Baena, N. Mine, et al. 2016. Not All That Glitters Is Gold: Metal-Migration-Induced Degradation in Perovskite Solar Cells. ACS Nano **10**:6306–6314.

Dong, Q., Y. Fang, Y. Shao, et al. 2015. Electron-Hole Diffusion Lengths Greater than 175 Micrometers in Solution-Grown $CH_3NH_3PbI_3$ Single Crystals. Science **347**:967–970.

Durant, B. K., H. Afshari, S. Singh, B. Rout, G. E. Eperon, and I. R. Sellers. 2021. Tolerance of Perovskite Solar Cells to Targeted Proton Irradiation and Electronic Ionization Induced Healing. ACS Energy Letters **6**:2362–2368.

Eames, C., J. M. Frost, P. R. F. Barnes, B. C. O'Regan, A. Walsh, and M. S. Islam. 2015. Ionic Transport in Hybrid Lead Iodide Perovskite Solar Cells. Nature Communications **6**:7497.

Google Research, 2025. Project Suncatcher: Scaling Machine Learning Compute in Space.

Hoang, M. T., Y. Yang, B. Tuten, and H. Wang. 2022. Are Metal Halide Perovskite Solar Cells Ready for Space Applications? The Journal of Physical Chemistry Letters **13**:2908–2920.

Hoke, E. T., D. J. Slotcavage, E. R. Dohner, et al. 2015. Reversible Photo-Induced Trap Formation in Mixed-Halide Hybrid Perovskites for Photovoltaics. Chemical Science **6**:613–617.

Kaltenbrunner, M., G. Adam, E. D. Głowacki, et al. 2015. Flexible High Power-per-Weight Perovskite Solar Cells with Chromium Oxide–Metal Contacts for Improved Stability in Air. Nature Materials **14**:1032–1039.

Kieslich, G., S. Sun, and A. K. Cheetham. 2014. Solid-State Principles Applied to Organic–Inorganic Perovskites: New Tricks for an Old Dog. Chemical Science **5**:4712–4715.

Kirmani, A. R., T. A. Byers, Z. Ni, et al. 2024. Unraveling Radiation Damage and Healing Mechanisms in Halide Perovskites Using Energy-Tuned Dual Irradiation Dosing. Nature Communications **15**:696.

Kirmani, A. R., B. K. Durant, J. Grandidier, et al. 2022. Countdown to Perovskite Space Launch: Guidelines to Performing Relevant Radiation-Hardness Experiments. Joule **6**:1015–1031.

Kirmani, A. R., and I. R. Sellers. 2025. Are Metal-Halide Perovskite Solar Cells Really Radiation Tolerant? Joule **9**:101852.

Kojima, A., K. Teshima, Y. Shirai, and T. Miyasaka. 2009. Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells. Journal of the American Chemical Society **131**:6050–6051.

Lang, F., N. H. Nickel, J. Bundesmann, et al. 2016. Radiation Hardness and Self-Healing of Perovskite Solar Cells. Advanced Materials **28**:8726–8731.

Lang, F., et al. 2021. Proton-Radiation Tolerant All-Perovskite Multijunction Solar Cells. Advanced Energy Materials **11**:2102246.

Magliarisi, D., E. Casalicchio, and V. Salvatore. 2025. Orbit-Aware Task Scheduling in Satellite Edge Computing. Cluster Computing **28**:1013.

Messenger, S. R., G. P. Summers, E. A. Burke, R. J. Walters, and M. A. Xapsos. 2001. Modeling Solar Cell Degradation in Space: A Comparison of the NRL Displacement Damage Dose and the JPL Equivalent Fluence Approaches. Progress in Photovoltaics: Research and Applications **9**:103–121.

National Renewable Energy Laboratory, 2026. Best Research-Cell Efficiency Chart. https://www.nrel.gov/pv/cell-efficiency.html.

Reb, L. K., M. Böhmer, B. Predeschly, et al. 2020. Perovskite and Organic Solar Cells on a Rocket Flight. Joule **4**:1054–1062.

Rolston, N., B. L. Watson, C. D. Bailie, et al. 2016. Mechanical Integrity of Solution-Processed Perovskite Solar Cells. Extreme Mechanics Letters **9**:353–358.

Romano, V., A. Agresti, R. Verduci, and G. D'Angelo. 2022. Advances in Perovskites for Photovoltaic Applications in Space. ACS Energy Letters **7**:2490–2514.

Seid, B. A., et al. 2024. Understanding and Mitigating Atomic Oxygen-Induced Degradation of Perovskite Solar Cells for Near-Earth Space Applications. Small **20**:2311097.

Sharma, M., M. Parashar, D. K. Saini, et al. 2026. Probing Self-Healing Dynamics in Perovskite Solar Cells under Proton Irradiation via In Situ and In Operando Characterization. Newton **2**:100332.

Song, S., H. W. Cho, and G.-H. Kim. 2026. Perovskite Solar Cells for Low Earth Orbit Space Applications. Energy Materials **6**:600019.

Steele, J. A., et al. 2019. Thermal Unequilibrium of Strained Black $CsPbI_3$ Thin Films. Science **365**:679–684.

Stranks, S. D., G. E. Eperon, G. Grancini, et al. 2013. Electron-Hole Diffusion Lengths Exceeding 1 Micrometer in an Organometal Trihalide Perovskite Absorber. Science **342**:341–344.

Travis, W., E. N. K. Glover, H. Bronstein, D. O. Scanlon, and R. G. Palgrave. 2016. On the Application of the Tolerance Factor to Inorganic and Hybrid Halide Perovskites: A Revised System. Chemical Science **7**:4548–4556.

Walsh, A., and A. Zunger. 2017. Instilling Defect Tolerance in New Compounds. Nature Materials **16**:964–967.

Weber, D. 1978. $CH_3NH_3PbX_3$, ein Pb(II)-System mit kubischer Perowskitstruktur. Zeitschrift für Naturforschung B **33**:1443–1445.

Yin, W.-J., T. Shi, and Y. Yan. 2014. Unusual Defect Physics in $CH_3NH_3PbI_3$ Perovskite Solar Cell Absorber. Applied Physics Letters **104**:063903.

Contributors



Minseong Eom
ehjung@kentech.ac.kr

Environmental climate technology and hydrogen energy trucks are the main areas of interest. There is a high interest in the idea of solving various environmental problems through the combination of biotechnology and chemistry. I am also interested in KENTECH's energy grid track and environmental issues education status.



Yunseok Choi
mcbenz@kentech.ac.kr

Currently an undergraduate student of KENTECH, interested in Environmental and Climate Technology. As a researcher of the Oh Group Lab, I study quantum dots to explore their potential in developing sustainable materials for next-generation energy and environmental applications.



Suhyeon Oh
suhnluz@kentech.ac.kr

Editor. Undergraduate student of KENTECH, with background Grid modernization. My board interests are understanding complex energy systems, exploring innovative grid solutions, and applying computational modeling. My goal is to simplify complex phenomena through scientific exploration, helping students think independently, ask questions and expand their investigations.



Eui Hyuk Jung
ehjung@kentech.ac.kr

Dr. Eui Hyuk Jung is Associate Professor in the Department of Energy Engineering at the Korea Institute of Energy Technology (KENTECH) in Naju, Republic of Korea. He received his combined B.S. and Ph.D. in Materials Science and Engineering from Seoul National University in 2016 under the supervision of Prof. Won Ho Jo. Prior to joining KENTECH in 2021, he held postdoctoral research appointments at the Korea Research Institute of Chemical Technology (KRICT, 2016–2019), Korea University under Prof. Jun Hong Noh (2019), and the University of Toronto under Prof. Edward H. Sargent (2019–2021). He has contributed significantly to the development of next-generation energy materials and photovoltaics, with notable recognition including the POSCO Science Fellowship and the Young Scientist Award from The Polymer Society of Korea in 2022. His current research focuses on the synthesis, structural engineering, and optoelectronic properties of advanced functional materials for high-efficiency energy harvesting and optoelectronic devices.

KENTECH
Korea Institute of Energy Technology

The Road Not Taken

And that has made all the difference

한국에너지공과대학교 입학기획운영팀
061-320-9655