

스마트 팩토리 내 트윈 모델 성능 향상을 위한 물리 모델 고도화 연구

유동건, 박선아, 서지윤, 이상금*

*한밭대학교

20211917@edu.hanbat.ac.kr, ssunaa0322@gmail.com, sillou0821@gmail.com,

*sangkeum@hanbat.ac.kr

Research on Advanced Physics-Based Models for Improving Twin Model Performance in Smart Factories

Donggeon Yu, Suna Park, Jiyun Seo, and Sangkeum Lee*

*Hanbat National University.

요 약

제조공정은 복잡한 열/유체/기계 시스템으로 구성되어 스팀 사용량 예측이 어렵고, 이에 따라 물리 모델과 AI 모델을 결합한 트윈 모델이 주목받고 있다. 본 연구는 국내 제조공정에서 수집된 1 분 간격의 1 년치 스팀 사용량 시계열 데이터를 기반으로, 전처리와 영향인자의 상관관계 기반 변수 선별 과정을 거쳐 해석력과 신뢰도가 높은 6 개 변수를 PySR(Python Symbolic Regression)입력으로 활용한 다. 트윈 모델은 LSTM(Long Short-Term Memory)기반 예측과 물리 모델 결과를 가중합하여 구성되며, 물리 모델의 예측 오차에 대해 PySR 기반 보정 수식을 적용함으로써 해석 가능성과 성능을 동시에 확보한다. 그 결과, 기존 물리 모델 대비 약 19%의 상관관계 향상이 나타나며, 트윈 모델의 실용성과 확장 가능성을 입증한다.

I. 서 론

제조공정은 스팀 소비가 많은 에너지 집약 공정으로, 과잉/부족사용 시 에너지 낭비, 탄소 배출 증가/제품 품질 저하를 초래한다[1]. 스팀 사용량의 정확한 예측과 최적화는 공정 효율 향상, 비용 절감, 품질 유지, 탄소 저감을 위한 핵심 과제이다[1]. 하지만 제조공정은 열, 유체, 기계 시스템이 복합된 비정형적 구조로, 계절, 설비 상태, 원료 특성 등 다양한 요인에 따라 공정 조건이 수시로 변동된다. 이로 인해 스팀 사용량 예측이 더욱 어려워지며, 이에 대한 대응으로 트윈 모델이 주목받는다[2].

트윈 모델은 물리 모델의 해석 능력과 AI 모델의 데이터 기반 예측력을 결합함으로써, 제조공정 내 다양한 운전 조건에 대한 시뮬레이션과 예측이 가능하다[3]. 예를 들어, 스팀 압력, 설비 가동률, 외기 온도, 원료 수분 함량 등이 변화할 경우, AI 모델은 과거 데이터를 기반으로 스팀 소비 패턴을 정확하게 예측하며, 물리 모델은 공정 원리를

바탕으로 데이터 부족이나 이상 조건에서도 예측 안정성을 제공한다. 다만, 실제 적용에서는 두 모델 간 성능 불균형이 문제로 작용한다. 최근 수년간 인공지능 모델은 빠르게 고도화된 반면, 물리 모델은 제한적인 개선에 그쳤다. 따라서 두 모델 간 균형과 협력 구조의 개선은 트윈 모델 성능 향상의 핵심 요소로 작용한다[4].

II. 제조공정 데이터 분석

2.1 데이터 및 환경 구성

국내 제조 공정에서 수집된 스팀 사용량 시계열 데이터를 활용한다. 해당 데이터는 1 분 간격으로 기록된 1 년치 자료를 포함하며, 결측값 보간, 이상치 제거, 정규화 등의 전처리 과정을 거쳐 분석에 사용한다. 사용된 값들은 스팀사용량, 재료농도, 이력타입, 평량, 생산 길이, 속도 표준, 속도 계산이다. 스팀사용량은 최적화하기 위한 목표값이다. Reel 평균, Size 평균, Reel 수분량, stockFlow 는

물리 모델에 사용되는 변수이다. 재료 농도, 이력 타입, 평량, 생산 길이, 속도 표준, 속도 계산은 스팀사용량과 상관관계가 높은 변수이며, PySR(Python Symbolic Regression)의 입력에 활용한다.

2.3 영향인자의 상관계수

그림 1 에서 데이터분석의 신뢰성을 확보하기 위해 결측률이 6% 미만인 변수만을 선별하여 PySR 의 입력에 활용하고, 입력 종류 중 6 개가 선정된 상관관계 히트맵이다.

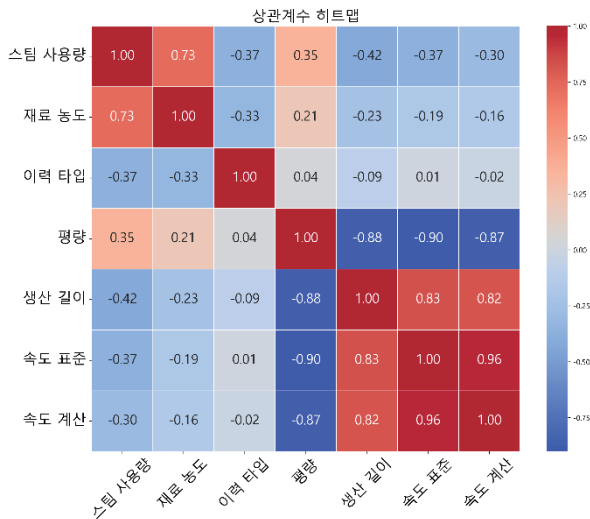


그림 1. 영향인자 간의 상관계수 히트맵

III. 방법론

3.1 트윈 모델의 전체 구조

계절, 설비 상태, 원료 특성 등 다양한 요인에 따라 열, 유체, 기계 시스템이 복합된 비정형적 구조인 팩토리에서 공정 조건이 수시로 변동된다. 이로 인해 스팀 사용량 예측이 어려워져, 트윈 모델의 구조를 제안한다. 물리 모델과 인공지능 모델을 가중합 형태로 결합하여 스팀 사용량을 예측하는 모델이다. 연구의 핵심 목적은 물리 모델의 고도화에 있으며, 트윈 모델은 다음과 같다:

$$Steam_{트윈} = \alpha \times LSTM(x) + \underbrace{\beta \times Steam_{계산}}_{\text{고도화 부분}}$$

여기서 기존의 $Steam_{계산}$ 를 고도화한다. $LSTM(x)$ 는 LSTM(Long Short-Term Memory) 기반의 예측값이고, $Steam_{계산}$ 은 물리 모델 예측값이며, α 와 β 는 모델 학습 이후 성능 개선을 위해 후처리 단계에서 조정 가능한 가중치이다.

3.2 물리 모델의 성능 향상

본 연구는 기존 물리 모델인 $Steam_{계산}$ 의 구조를 고도화하는 데 목적이 있다. 기존 물리식은 다음과 같이 정의된다:

$$Steam_{계산} = W_1 \times Reel_{평균} + W_2 \times Size_{평균} + W_3 \times Reel_{수분량} + W_4 \times stockFlow + B$$

W_1, W_2, W_3, W_4 는 물리식의 가중치이며, B 는 상수항이다. 본 연구에서는 기존식에 PySR 을 이용해 조정항 $A(x)$ 와 상수항 B_{adj} 를 추가하여 $Steam_{잔차}$ 을 적용한다. 이는 단순한 오차 보정이 아니라, 잔차의 구조를 수학적 물리적으로 해석 가능한 형태로 표현하기 위함이다. 입력 변수와 출력 간의 관계를 기호 기반 수식으로 추론하는 symbolic regression 기법으로, 해석력과 데이터 기반 보정을 자연스럽게 결합할 수 있는 효과적인 방식이다. PySR 모델은 다음과 같은 함수 집합 $A(x)$ 을 허용한다:

$$A(x) = \left\{ \log(x), a^x, \sqrt{x}, \sin(x), \cos(x), \tan(x), x^n, \frac{1}{x}, Wx \right\}$$

상관계수로 재료농도 및 선형관계로 Wx 를 선정한다. $Steam_{잔차}$ 에 대해 PySR 을 통해 다음과 같이 정의된다:

$$\begin{aligned} Steam_{잔차} &= Steam_{실제} - Steam_{계산} \\ &\approx \underbrace{A(x) + B_{adj}}_{PySR \text{ 모델}} \\ &= W_5 \times \text{재료_농도} + B_{adj} \end{aligned}$$

실제값에 PySR 을 직접 적용하면 기존 물리 모델과 무관한 함수가 생성되어 통합이 어려우나, 잔차에 적용하면 기존 물리 모델을 유지한 채 부족한 부분만 보정할 수 있다. 이는 최종적으로 개선된 물리 모델과 인공지능 모델이 결합 트윈 모델은 다음과 같이 정리된다:

$$\begin{aligned} Steam_{트윈}^* &= \alpha \times LSTM(x) + \beta \times Steam_{계산}^* \\ &= \alpha \times LSTM(x) + \beta \times \left(Steam_{계산} + \underbrace{A(x) + B_{adj}}_{PySR \text{ 모델}} \right) \\ &= \alpha \times LSTM(x) + \beta \times (W_1 \times Reel_{평균} + W_2 \times Size_{평균} + W_3 \times Reel_{수분량} + W_4 \times stockFlow + B + W_5 \times \text{재료_농도} + B_{adj}) \end{aligned}$$

$Steam_{실제}$ 에 대한 $Steam_{계산}$, $Steam_{계산}^*$ 간의 상관계수는 각각 0.5427, 0.6464 로, 약 19%의 향상된다.

IV. 결론

본 연구는 스팀 사용량 예측 정확도를 향상시키기 위한 트윈 모델 구조를 제안하고, 이를 위해 물리 모델의

고도화를 중심으로 분석한다. 특히 기존 물리 모델의 한계를 보완하기 위해 PySR 기반의 symbolic regression 기법을 도입하고, 단순한 오차 보정이 아닌 물리적으로 해석 가능한 수식을 도출함으로써 모델의 해석력과 예측력을 동시에 확보한다. 제안하는 모델은 데이터 신뢰도를 기준으로 6 개의 변수를 선별하여 적용하며, 그 결과 상관관계수가 기존 물리 모델 대비 유의미하게(약 19%) 향상된다. 향후에는 제안한 트윈 모델 구조를 실제 제지 설비 환경에 적용하고, 운전 조건 변화에 따라 α , β 가중치를 유연하게 조정할 수 있는 동적 가중치 구조를 도입한다. 또한 시간에 따른 LSTM 과 물리 모델의 학습을 자동화하는 강화학습 기반의 최적화 연구를 수행한다.

ACKNOWLEDGMENT

“본 연구는 2025 년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 연구결과로 수행되었음”(2022-0-01068)

참 고 문 헌

- [1] 이상금, 2022. 제지공정 건조 스팀에너지 모델링 방법론. 2022 년도 한국통신학회 하계종합학술대회, page 458-459.
- [2] Zhang, Jiwei, et al. "An intelligent digital twin system for paper manufacturing in the paper industry." *Expert Systems with Applications* 230 (2023): 120614.
- [3] Sun, Chao, and Victor G. Shi. "PhysiNet: A combination of physics-based model and neural network model for digital twins." *International journal of intelligent systems* 37.8 (2022): 5443-5456.
- [4] Sun, Chao, and Victor G. Shi. "PhysiNet: A combination of physics-based model and neural network model for digital twins." *International journal of intelligent systems* 37.8 (2022): 5443-5456
- [5] Cranmer, Miles. "Interpretable machine learning for science with PySR and SymbolicRegression. jl." *arXiv preprint arXiv:2305.01582* (2023).
- [6] Lee, Sangkeum, et al. "Anomaly detection of smart metering system for power management with battery storage system/electric vehicle." *ETRI Journal* 45.4 (2023): 650-665.