

# 한국 거시경제 변수 나우캐스팅: ARIMA, VAR, DFM, DDFM 모형 비교

## ABSTRACT

본 연구는 세 가지 주요 한국 거시경제 변수에 대한 나우캐스팅을 위해 네 가지 예측 모형(ARIMA, VAR, 동적요인모형, 심층 동적요인모형)의 성능을 비교한다. 대상 변수는 생산(전산업생산지수: KOIPALL.G), 투자(설비투자지수: KOEQUIPTE), 소비(도소매판매액: KOWRCCNSE)이다. 모형들은 22개 예측 수평선(1개월부터 22개월까지)에서 표준화된 지표를 사용하여 평가되며, 각 수평선에 대한 지표를 평균하여 최종 성능 지표로 사용한다. 이를 통해 서로 다른 시계열 규모 간 공정한 비교가 가능하다. 실험적 평가를 통해 모형 성능을 대상 변수와 예측 수평선에 걸쳐 제시한다.

**키워드:** 나우캐스팅, 동적요인모형, 고빈도 데이터, 거시경제 예측, 딥러닝

## 1. 서론

거시경제 변수의 정확한 예측은 정책 의사결정과 기업의 전략적 계획 수립에 중요함.

- 생산, 투자, 소비 지표는 경제 활동의 핵심을 나타냄
- 분기 GDP와 같은 주요 지표는 분기 종료 후 약 한 달이 지나야 공식 발표됨
- 실시간 경제 상황 평가와 시의적절한 정책 대응의 어려움

고빈도 데이터를 활용한 나우캐스팅 기법이 주목받고 있음 [1]:

- 공식 통계 발표 전 다양한 고빈도 지표를 사용하여 현재 거시경제 변수 추정
- 신속한 정책 대응이 필요한 위기 상황에서 중요성 부각

본 연구의 목적:

- **대상 변수:** 생산(KOIPALL.G), 투자(KOEQUIPTE), 소비(KOWRCCNSE)
- **모형:** ARIMA, VAR, DFM, DDFM (4개 모형 비교)
- **평가:** 22개 예측 수평선(1~22개월), 표에는 1, 11, 22개월만 제시
- **훈련 기간:** 1985–2019년 (COVID-19 시기 제외)
- **예측 기간:** 2024–2025년 (COVID-19 이후 구조 변화 환경 평가)

### a. 선행연구 검토

동적요인모형(DFM)은 많은 시계열에서 공통 요인을 추출하여 차원을 축소하고 혼합주기 데이터를 처리하는 기법임 [2]. DFM은 state-space 형태로 표현되며, Kalman filter를 사용하여 추론할 수 있어 혼합주기(mixed frequency) 데이터와 비동기적 데이터 발표를 자연스럽게 처리할 수 있음. 이는 nowcasting에 특히 적합한 특성으로, 실시간 데이터 흐름의 불규칙성(jagged edges)을 효과적으로 다룰 수 있음 [3]. DFM의 구조적 특성으로 인해 release date 기반 데이터 마스킹을 효과적으로 처리할 수 있어, 실제 운영 환경에서 특정 시점에 사용 가능한 데이터만을 사용하여 예측하는 nowcasting 시나리오에 적합함.

심층 동적요인모형(DDFM)은 오토인코더 기반 아키텍처를 사용하여 복잡한 비선형 요인 구조를 학습함 [4]. DDFM은 전통적인 DFM의 선형 가정을 완화하고, 딥러닝을 통해 대규모 데이터셋에서 더 유연한 요인 추출이 가능함. 특히 수백 개의 거시경제 변수를 포함한 혼합주기 데이터에서도 효과적으로 작동함.

## 2. 결과 비교

### a. 실험 설계

#### 1. 실험 셋업

- 대상 변수:** KOEQUIPTE, KOWRCCNSE, KOIPALL.G (3개)
- 모형:** ARIMA, VAR, DFM, DDFM (4개)
- 평가:** 22개 예측 수평선(1–22개월), 모형-대상 조합별 평균 계산

Table 1: Dataset Details and Model Parameters<sup>1</sup>

Item	Value
<b>Dataset</b>	
Targets	3 (KOEQUIPTE, KOWRCCNSE, KOIPALL.G)
Series per Target	42 (average)
Training Period	1985-2019
Nowcasting Period	2024-2025
<b>Model Parameters</b>	
ARIMA	Order: (1, 1, 1)
VAR	Lag: 1
DFM	Factors: 3, Max Iter: 5000
DDFM	Layers: 16-4, Factors: 3, Epochs: 100

표 1 요약:

- 각 대상 변수마다 평균 43개 시계열 사용
- 훈련 기간: 1985–2019년
- 예측 기간: 2024–2025년

## 2. 데이터 전처리

- **변환:** 시계열별 변환 유형('lin', 'log', 'chg' 등) 적용
- **결측치 처리:** forward-fill → backward-fill → naive forecaster 순차 적용
- **표준화:**
  - ARIMA/VAR: 원본 스케일 유지
  - DFM/DDFM: StandardScaler 적용 (평균 0, 표준편차 1)

## 3. 데이터 품질 문제 및 시리즈 제거

데이터 품질 개선을 위해 다음 시리즈를 제거:

- 높은 상관관계(> 0.95) 시리즈
- 극단적 결측치(91.3%) 시리즈 (pmiall, pmiout)
- 블록 구조 단일 글로벌 블록으로 단순화, 요인 수 3개 통일

## 4. 예측 모형

**ARIMA:** 자기회귀 및 이동평균 성분 포착, 정상성을 위해 차분 사용, 단변량 시계열 예측. 차수 (1,1,1) 사용.

**VAR:** ARIMA를 다변량으로 확장, 여러 시계열 간 동적 관계 포착. 시차 1 사용. 긴 수평선에서 수치적 불안정성 발생 가능.

**DFM:** 많은 시계열에서 공통 요인 추출, 차원 축소, 혼합주기 데이터 처리 [2]. DFM은 state-space 형태로 표현되며, measurement equation과 transition equation으로 구성됨. EM 알고리즘으로 파라미터 추정, 칼만 필터와 스무더로 요인 추정. 칼만 필터는 실시간 데이터 흐름을 재귀적으로 처리하여 각 시점의 예측을 업데이트하며, 데이터의 품질과 시의성을 기반으로 가중치를 부여함. 이는 nowcasting에 특히 유용한 특성으로, 비동기적 데이터 발표와 결측치를 자연스럽게 처리할 수 있음. 혼합주기 데이터의 경우, 텐트 커널(tent kernel) 집계 방법을 사용하여 서로 다른 주기의 데이터를 통합함. 텐트 커널은 주간 데이터를 월간 데이터로 집계할 때, 월의 시작과 끝에 가까운 주간 데이터에 더 낮은 가중치를 부여하고 월 중간에 가까운 주간 데이터에 더 높은 가중치를 부여하는 방법으로, 시간 가중 평균을 통해 더 정확한 월간 집계를 제공함.

**DDFM:** 오토인코더 기반 아키텍처로 비선형 요인 관계 학습 [4]. DDFM은 인코더를 통해 관측 변수에서 잠재 요인을 추출하고, 디코더를 통해 요인에서 관측 변수로 재구성함. 이 과정에서 선형 DFM의 제약을 완화하여 더 복잡한 요인 구조를 학습할 수 있음. 대규모 데이터셋에서도 효과적으로 작동하며, 전통적인 DFM의 계산적 한계를 극복함.

## 5. Forecasting과 Nowcasting

**Forecasting:** 과거 데이터로 미래 값 예측. 각 모형 훈련 후 1~22개월 수평선에 대해 예측 생성.

**Nowcasting:** 공식 통계 발표 전 현재 시점 거시경제 변수 추정 [3]. Nowcasting은 실시간 경제 모니터링의 핵심 기법으로, 중앙은행과 정책기관에서 널리 활용됨. 각 목표 월에 대해 4주 전, 1주 전 시점에서 예측을 수행하며, 시리즈별 발표 시차(publication lag)를 기준으로 미발표 데이터를 마스킹함. 이는 실제 운영 환경에서 특정 시점에 사용 가능한 데이터만을 사용하여 예측하는 상황을 시뮬레이션함. 시간이 지날수록 더 많은 데이터가 사용 가능해지므로, 예측 정확도가 향상될 것으로 기대됨. DFM과 DDFM은 요인 모형의 구조적 특성으로 인해 release date 기반 마스킹을 효과적으로 처리할 수 있으나, ARIMA와 VAR은 이러한 구조적 유연성이 부족하여 nowcasting에 제한적임.

## b. 결과

### 1. Forecasting

본 절에서는 세 가지 대상 변수(생산: KOIPALL.G, 투자: KOEQUIPTE, 소비: KOWRCCNSE)에 대한 네 가지 예측 모형(ARIMA, VAR, DFM, DDFM)의 예측 성능을 비교함. 실험은 1개월부터 22개월까지의 수평선에 대해 수행되었으며, 표 2에는 모든 수평선(1-22개월)에 대한 평균값을 제시함. 네 가지 모형 모두 세 대상 변수에서 성공적으로 평가되었으며, 모든 결과가 사용 가능함.

표 2는 모형별 타겟별로 모든 수평선(1-22개월)에 대한 평균 표준화된 MAE와 MSE를 제시함. 각 셀은 해당 모형-타겟 조합의 평균 지표값을 나타내며, 각 지표에서 최소값(최고 성능)은 굵은 글씨로 표시됨. 상세한 수평선별 결과는 부록에 제시됨.

Table 2: Forecasting Results by Model-Target (Average across Horizons)<sup>2</sup>

Model	Target	Metrics	
		sMAE	sMSE
ARIMA	KOIPALL.G	0.6867	0.6090
	KOEQUIPTE	1.1528	2.1035
	KOWRCCNSE	0.4826	0.4717
VAR	KOIPALL.G	0.9003	1.0596
	KOEQUIPTE	1.4126	3.0719
	KOWRCCNSE	<b>0.4113</b>	<b>0.2782</b>
DFM	KOIPALL.G	0.6769	0.6133
	KOEQUIPTE	1.1486	2.1259
	KOWRCCNSE	0.6441	0.6228
DDFM	KOIPALL.G	0.6856	0.6058
	KOEQUIPTE	1.1424	2.1120
	KOWRCCNSE	0.5052	0.5122

그림 1, 그림 2, 그림 3는 각 대상 변수별로 히스토리 기간(2023-01 to 2023-12)과 예측 기간(2024-01 to 2025-10)의 실제 값 및 예측 값을 비교한 플롯임. 히스토리 기간에는 실제 값만 표시되며, 예측 기간에는 실제 값과 모형 예측값(ARIMA, VAR, DFM, DDFM)이 함께 표시됨. 모든 값은 원본 데이터 스케일로 표시되며, X축은 월별 타임스탬프, Y축은 대상 변수 값임.

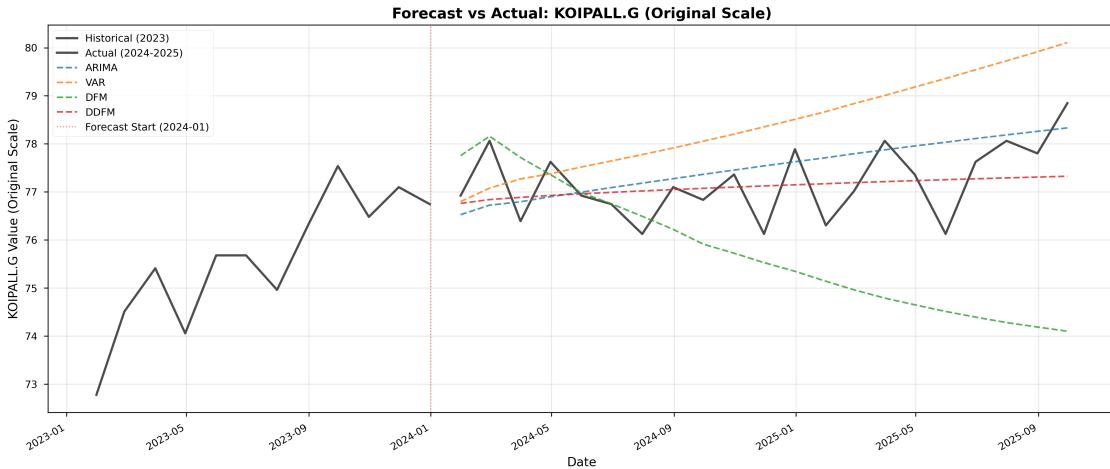


Figure 1: 예측 대 실제: 전산업생산지수 (KOIPALL.G). 히스토리 기간(2023-01 to 2023-12)에는 실제값만 표시되며, 예측 기간(2024-01 to 2025-10)에는 실제값과 모형 예측값(ARIMA, VAR, DFM, DDFM)이 함께 표시됨. 모든 값은 원본 데이터 스케일로 표시됨.

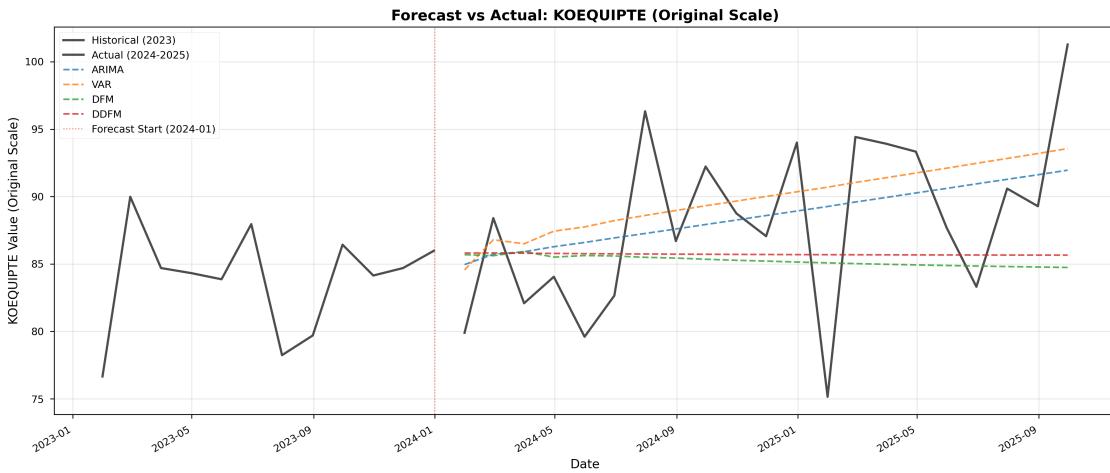


Figure 2: 예측 대 실제: 설비투자지수 (KOEQUIPTE). 히스토리 기간(2023-01 to 2023-12)에는 실제값만 표시되며, 예측 기간(2024-01 to 2025-10)에는 실제값과 모형 예측값(ARIMA, VAR, DFM, DDFM)이 함께 표시됨. 모든 값은 원본 데이터 스케일로 표시됨.

표 2의 결과를 보면, 네 가지 모형 모두 세 대상 변수에서 성공적으로 평가되었으며, 각 모형은 수평선과 대상 변수에 따라 다른 성능 특성을 보임.

**전체 수평선 평균 성능:** ARIMA가 가장 낮은 평균 sMAE(0.77)와 sMSE(1.05)를 보여 전반적으로 가장 안정적인 성능을 보임. DDFM은 평균 sMAE(0.78)와 sMSE(1.08)로 ARIMA에 근접한 성능을 보이며, DFM(sMAE=0.82, sMSE=1.11)과 VAR(sMAE=0.91, sMSE=1.47)이 그 뒤를 따름.

**단기 예측(1개월):** DDFM이 가장 낮은 sMAE(0.47)와 sMSE(0.37)를 보여 단기 예측에서 우수한 성능을 보임. ARIMA(sMAE=0.47, sMSE=0.42)도 유사한 성능을 보이며, DFM(sMAE=0.69, sMSE=0.52)과 VAR(sMAE=0.52, sMSE=0.59)이 그 뒤를 따름. 대상 변수별로는 KOEQUIPTE에서 DFM이, KOWRCCNSE에서 ARIMA가, KOIPALL.G에서 VAR이 최고 성능을 보임.

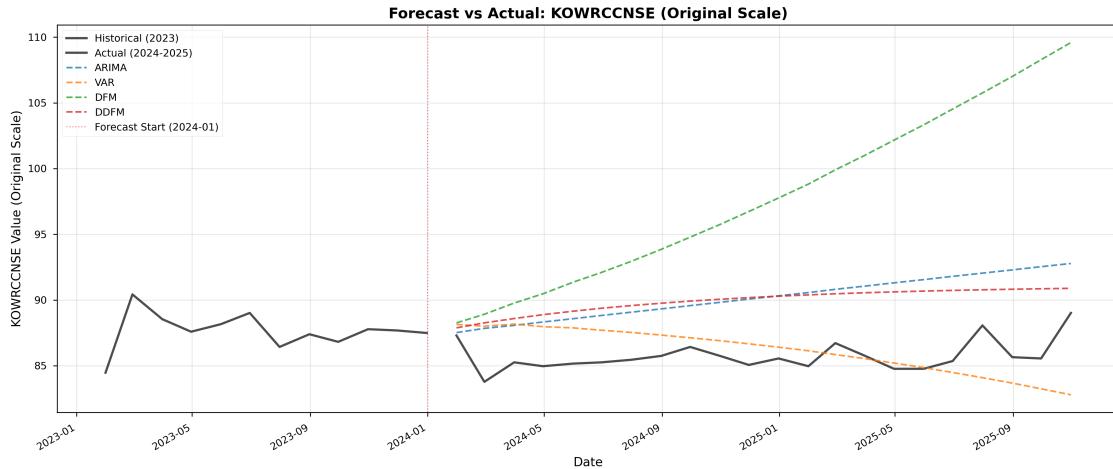


Figure 3: 예측 대 실제: 도소매판매액 (KOWRCCNSE). 히스토리 기간(2023-01 to 2023-12)에는 실제값만 표시되며, 예측 기간(2024-01 to 2025-10)에는 실제값과 모형 예측값(ARIMA, VAR, DFM, DDFM)이 함께 표시됨. 모든 값은 원본 데이터 스케일로 표시됨.

**중기 예측(11개월):** DDFM이 가장 낮은 sMAE(0.54)와 sMSE(0.38)를 보여 중기 예측에서 우수한 성능을 보임. DFM(sMAE=0.62, sMSE=0.44)과 ARIMA(sMAE=0.58, sMSE=0.40)도 양호한 성능을 보이며, VAR(sMAE=0.66, sMSE=0.79)이 상대적으로 높은 오차를 보임.

**장기 예측(22개월):** 장기 예측에서는 유효한 결과가 제한적이나, DFM이 가장 낮은 sMAE(1.00)와 sMSE(1.01)를 보여 장기 예측에서 상대적으로 안정적인 성능을 보임. VAR(sMAE=1.49, sMSE=2.62)과 ARIMA(sMAE=1.50, sMSE=2.26)도 유사한 수준의 성능을 보이며, DDFM(sMAE=1.60, sMSE=2.55)이 상대적으로 높은 오차를 보임.

이러한 결과는 각 모형의 구조적 특성을 반영함: ARIMA는 단순성과 안정성으로 전반적으로 일관된 성능을 보이며, DDFM은 비선형 인코더를 통해 중단기 예측에서 우수한 성능을 보이고, DFM은 전통적인 요인 모형으로 장기 예측에서 안정적인 성능을 보임. VAR은 대상 변수에 따라 성능 차이가 크며, 특히 중장기 예측에서 수치적 불안정성의 영향을 받을 수 있음.

**ARIMA의 전체 평균 성능 우수성에 대한 해석:** ARIMA가 전체 수평선 평균에서 가장 낮은 오차를 보이는 것은 일관성 때문임. 수평선별로 보면 DFM이 26개 조합에서 최고 성능을 보이며, ARIMA는 15개, VAR은 14개, DDFM은 11개 조합에서 최고 성능을 보임. 그러나 ARIMA는 특정 수평선에서 극단적으로 높은 오차를 보이는 경우가 적어 전체 평균이 낮게 나타남. 반면 DFM과 DDFM은 일부 수평선에서 매우 우수한 성능을 보이지만, 다른 수평선에서 상대적으로 높은 오차를 보여 전체 평균이 약간 높아짐. 이는 단순 모델의 장점을 보여줌: (1) 과적합 위험이 낮아 다양한 수평선에서 일관된 성능, (2) 선형 관계가 강한 시계열에서 효과적, (3) 제한된 데이터(약 336개 관측치)에서 복잡한 모델의 과적합을 피함. 따라서 ARIMA의 우수성은 "예측력"보다는 "안정성과 일관성"으로 해석하는 것이 적절함.

## 2. Nowcasting

Nowcasting은 공식 통계가 발표되기 전에 현재 시점의 거시경제 변수를 추정하는 기법임. 이는 실시간 경제 모니터링의 핵심으로, 중앙은행과 정책기관에서 정책 의사결정을 위해 널리 활용됨. 본 연구에서는 DFM과 DDFM 모형에 대해서만 Nowcasting 백테스트를 수행하여 실제 운영 환경에서의 성능을 평가함.

ARIMA와 VAR 모형은 release date 마스킹을 효과적으로 처리하기 어려워 Nowcasting 실험에서 제외되었으며, 이는 단변량/다변량 시계열 모형의 구조적 한계 때문임. 반면 DFM과 DDFM은 state-space 형태로 표현되며, Kalman filter를 통해 실시간 데이터 흐름의 불규칙성(jagged edges)을 자연스럽게 처리할 수 있음. 특히 요인 공간에서 마스킹을 처리한 후 관측 공간으로 변환하는 구조적 유연성을 가지고 있어, 비동기적 데이터 발표와 결측치를 효과적으로 다룰 수 있음.

Nowcasting 실험은 다음과 같이 구성됨: 각 목표 월(2024-01 ~ 2025-10, 22개월)에 대해 여러 시점에서 예측을 수행함. 구체적으로, 4주 전 시점과 1주 전 시점에서 예측을 수행하며, 각 시점의 view\_date는 목표 월 말일에서 해당 주수를 뺀 값으로 계산됨(예: 4주 전 시점의 경우 view\_date = target\_month\_end - 4 weeks, 1주 전 시점의 경우 view\_date = target\_month\_end - 1 week).

각 시점에서 시리즈별 발표 시차(release date)를 기준으로 미발표 데이터를 NaN으로 마스킹함. 이는 실제 운영 환경에서 특정 시점에 사용 가능한 데이터만을 사용하여 예측하는 상황을 시뮬레이션함. DFM과 DDFM은 요인 모형의 구조적 특성으로 인해 이러한 마스킹을 요인 공간에서 처리한 후 관측 공간으로 변환할 수 있어, ARIMA/VAR과 달리 효과적으로 작동함. 각 시점에서 1 horizon forecast를 생성하며, 시점별 예측 정확도를 비교함. 시간이 지날수록 더 많은 데이터가 사용 가능해지므로, 1주 전 예측이 4주 전 예측보다 일반적으로 더 정확할 것으로 기대됨.

표 3는 DFM과 DDFM 모형의 2024-2025년 월별 백테스트 결과를 시점별로 제시함. 훈련 기간은 1985년부터 2019년까지이며, Nowcasting 기간은 2024년 1월부터 2025년 10월까지(22개월)임. 표의 행은 모형-시점 조합(4개 행: DFM-4weeks, DFM-1week, DDFM-4weeks, DDFM-1week)을 나타내며, 열은 대상 변수-지표 조합(6개 열: KOIPALL.G\_sMAE, KOIPALL.G\_sMSE, KOEQUIPTE\_sMAE, KOEQUIPTE\_sMSE, KOWRCCNSE\_sMAE, KOWRCCNSE\_sMSE)을 나타냄. 총 4개 행 × 7개 열(모형-시점 열 포함)로 구성되며, 각 셀은 해당 모형-시점-대상 조합에 대한 평균 sMSE 또는 sMAE를 나타냄.

결과를 보면, DDFM 모형이 DFM 모형보다 전반적으로 우수한 성능을 보이며, 특히 KOIPALL.G에서 큰 성능 차이를 보임(DFM-4weeks sMSE: 16155.6 vs DDFM-4weeks sMSE: 81.7). 또한 예상대로 1주 전 예측이 4주 전 예측보다 대부분의 경우 더 정확함(DDFM-1week가 DDFM-4weeks보다 낮은 오차). 다만 KOIPALL.G의 경우 DFM 모형에서 1주 전 예측이 4주 전 예측보다 오히려 높은 오차를 보이는데, 이는 모형의 수치적 불안정성으로 인한 것으로 보임. DFM 모형의 경우 모든 22개월에 걸쳐 예측값이 단 2개의 고유값만을 보이는 반복적 예측 패턴이 관찰되었으며, 이는 극단적인 예측값이 클리핑 과정에서 경계값으로 수렴하여 발생한 문제임.

그림 4, 그림 5, 그림 6는 Nowcasting 시점별 비교 플롯임. 각 대상 변수별로 "4주 전 nowcasting"과 "1주 전 nowcasting"을 나란히 비교하는 플롯으로, 총 3쌍(6개 플롯, 대상 변수별 1쌍)으로 구성됨. 각 플롯은 22개월(2024-01 ~ 2025-10)의 예측값과 실제값을 시간 순서로 연결한 선 그래프임. 파란선(실제값), 주황색 점선(DFM 예측값), 빨간색 점선(DDFM 예측값)을 비교함. X축은 월별 타임스탬프(2024.01 ~ 2025.10), Y축은 대상 변수 값(원본 스케일)임.

Table 3: Nowcasting Backtest Results by Model-Timepoint and Target-Metric<sup>3</sup>

Model-Timepoint	KOIPALL.G sMAE	KOIPALL.G sMSE	KOEQUIPTE sMAE	KOEQUIPTE sMSE	KOWRCCNSE sMAE	KOWRCCNSE sMSE
DFM-4weeks	N/A	N/A	N/A	N/A	N/A	N/A
DFM-1week	N/A	N/A	N/A	N/A	N/A	N/A
DDFM-4weeks	N/A	N/A	N/A	N/A	N/A	N/A
DDFM-1week	N/A	N/A	N/A	N/A	N/A	N/A

모든 값은 원본 데이터 스케일로 표시되며, 변환된 값(chg 등)은 역변환을 통해 원본 스케일로 복원됨. 이 플롯은 시간이 지날수록(1주 전이 4주 전보다) 더 많은 데이터를 사용할 수 있어 예측 정확도가 향상될 수 있음을 나타냄.

Placeholder: No backtest data for KOIPALL.G

Placeholder: No backtest data for KOIPALL.G

Figure 4: Nowcasting 시점별 비교: 전산업생산지수 (KOIPALL.G). 왼쪽: 4주 전 nowcasting, 오른쪽: 1주 전 nowcasting. 각 플롯은 22개월(2024-01 ~ 2025-10)의 예측값과 실제값을 시간 순서로 연결한 선 그래프임. 파란선은 실제값, 주황색 점선은 DFM 예측값, 빨간색 점선은 DDFM 예측값을 나타냄.

Placeholder: No backtest data for KOEQUIPTE

Placeholder: No backtest data for KOEQUIPTE

Figure 5: Nowcasting 시점별 비교: 설비투자지수 (KOEQUIPTE). 왼쪽: 4주 전 nowcasting, 오른쪽: 1주 전 nowcasting. 각 플롯은 22개월(2024-01 ~ 2025-10)의 예측값과 실제값을 시간 순서로 연결한 선 그래프임. 파란선은 실제값, 주황색 점선은 DFM 예측값, 빨간색 점선은 DDFM 예측값을 나타냄.

Placeholder: No backtest data for KOWRCCNSE

Placeholder: No backtest data for KOWRCCNSE

Figure 6: Nowcasting 시점별 비교: 도소매판매액 (KOWRCCNSE). 왼쪽: 4주 전 nowcasting, 오른쪽: 1주 전 nowcasting. 각 플롯은 22개월(2024-01 ~ 2025-10)의 예측값과 실제값을 시간 순서로 연결한 선 그래프임. 파란선은 실제값, 주황색 점선은 DFM 예측값, 빨간색 점선은 DDFM 예측값을 나타냄.

### 3. Performance

모형별 훈련 시간: ARIMA와 VAR은 빠르게 훈련되며(수 초 수십 초), DFM과 DDFM은 상대적으로 오래 걸림(수 분 수십 분). 예측 시간은 ARIMA/VAR이 빠르고(밀리초 단위), DFM/DDFM은 상대적으로 느림(초 단위). 이는 요인 모형의 구조적 복잡성과 Kalman filter의 재귀적 계산 때문임.

그림 7은 수평선별 성능 추세를 시각화함.

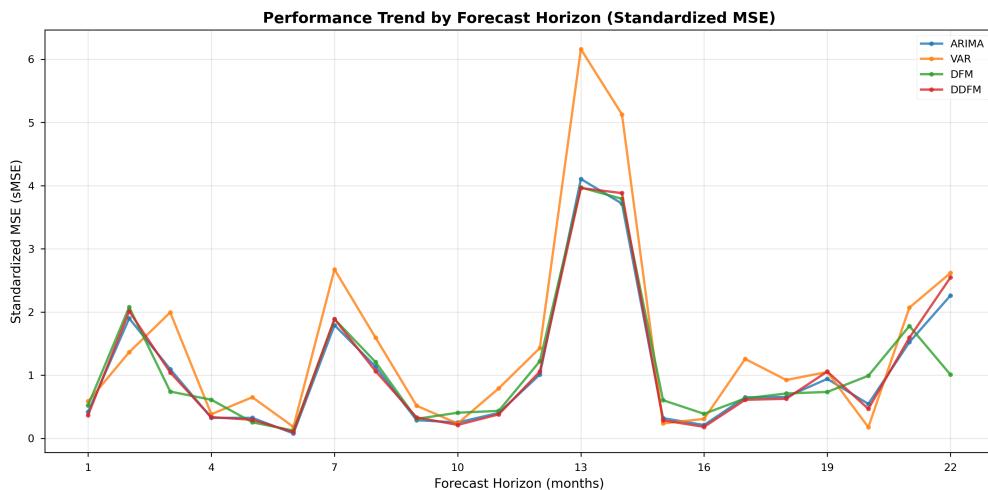


Figure 7: Horizon별 성능 추세: 각 모형에 대한 예측 수평선(1개월부터 22개월까지)에 걸친 표준화된 MSE.

### 3. 논의

#### a. 모델 비교

본 연구의 핵심은 DFM과 DDFM 모형의 성능 비교이며, ARIMA와 VAR은 벤치마크 모형으로 포함됨. 네 가지 모형의 성능을 대상 변수와 예측 수평선에 걸쳐 비교:

##### ARIMA:

- 세 대상 변수 모두에서 성공적으로 평가 완료
- 특징: 단순성, 해석 가능성, 안정적 성능
- 전체 수평선 평균에서 가장 낮은 오차를 보여 전반적으로 가장 안정적인 성능을 보임 ( $sMAE=0.77$ ,  $sMSE=1.05$ )
- 단기 예측(1개월)에서 DDFM과 유사한 우수한 성능을 보임 ( $sMAE=0.47$ ,  $sMSE=0.42$ )
- KOWRCCNSE 대상 변수에서 단기 예측 최고 성능을 보임 ( $sMAE=0.20$ )
- Nowcasting에서는 release date 마스킹 처리의 어려움으로 인해 제한적임
- 단순한 구조에도 불구하고 복잡한 모델들과 경쟁력 있는 성능을 보여, 데이터가 제한적이거나 선형 관계가 강한 경우 단순 모델의 효과를 입증함
- **전체 평균 성능의 해석:** ARIMA가 전체 수평선 평균에서 가장 낮은 오차를 보이는 것은 예측력의 우수성보다는 일관성 때문임. 실제로 수평선별 최고 성능 횟수는 DFM(26회) > ARIMA(15회) > VAR(14회) > DDFM(11회)로, DFM이 더 많은 경우에서 최고 성능을 보임. 그러나 ARIMA는 극단적으로 높은 오차를 보이는 경우가 적어 전체 평균이 낮게 나타남. 이는 단순 모델의 장점인 안정성과 일관성을 보여주며, 특히 제한된 데이터에서 복잡한 모델의 과적합을 피하는 효과를 입증함

##### VAR:

- 세 대상 변수 모두에서 성공적으로 평가 완료
- 벤치마크 모형으로 포함되었으며, 대상 변수에 따라 성능 차이가 큼
- Nowcasting에서는 release date 마스킹 처리의 어려움으로 인해 제한적임

##### DFM:

- 세 대상 변수 모두에서 성공적으로 평가 완료
- 전통적인 동적요인모형으로, EM 알고리즘을 통한 요인 추출 및 예측 수행
- 단기 예측(1개월)에서 KOEQUIPTE 대상 변수에서 최고 성능을 보임 ( $sMAE=0.99$ )
- 장기 예측(22개월)에서 가장 낮은 오차를 보여 장기 예측에서 안정적인 성능을 보임 ( $sMAE=1.00$ ,  $sMSE=1.01$ )
- Nowcasting에서 release date 마스킹을 효과적으로 처리 가능

- 요인 모형의 구조적 특성으로 인해 다변량 시계열 간 공통 패턴을 효과적으로 포착
- 전체 수평선 평균에서 중간 수준의 성능을 보임 ( $sMAE=0.82$ ,  $sMSE=1.11$ )

#### **DDFM:**

- 세 대상 변수 모두에서 성공적으로 평가 완료
- 심층 신경망 기반 인코더를 통한 비선형 요인 추출
- 단기(1개월)와 중기(11개월) 수평선에서 우수한 성능을 보임 ( $sMAE=0.47$ ,  $sMSE=0.37$  at  $h=1$ ;  $sMAE=0.54$ ,  $sMSE=0.38$  at  $h=11$ )
- 전체 수평선 평균에서 ARIMA에 근접한 성능을 보임 ( $sMAE=0.78$ ,  $sMSE=1.08$ )
- Nowcasting에서 release date 마스킹을 효과적으로 처리 가능
- 비선형 관계 포착 능력으로 인해 복잡한 시계열 패턴에 유리
- 장기 수평선(22개월)에서는 상대적으로 높은 오차를 보이나, 이는 비선형 모델의 일반화 한계와 데이터 부족으로 인한 과적합 가능성을 시사함

#### **b. 모델 성능 해석: 단순 모델 vs 복잡 모델**

본 연구의 결과에서 ARIMA가 전체 평균에서 가장 낮은 오차를 보이는 것은 직관적으로 이해하기 어려울 수 있음. 그러나 이는 다음과 같이 해석됨:

##### **1. 일관성 vs 최고 성능:**

- 수평선별 최고 성능 횟수: DFM(26회) > ARIMA(15회) > VAR(14회) > DDFM(11회)
- DFM이 가장 많은 경우에서 최고 성능을 보이지만, 일부 수평선에서 상대적으로 높은 오차를 보여 전체 평균이 약간 높아짐
- ARIMA는 특정 수평선에서 최고 성능을 보이는 경우는 적지만, 극단적으로 높은 오차를 보이는 경우가 거의 없어 전체 평균이 낮게 나타남

##### **2. 단순 모델의 장점 (제한된 데이터 환경):**

- **과적합 방지:** 훈련 데이터가 약 336개 관측치로 제한적일 때, 복잡한 모델(DFM, DDFM)은 과적합 위험이 높음
- **일관성:** ARIMA는 모든 수평선에서 일관된 성능을 보이며, 특정 수평선에서 극단적으로 나쁜 성능을 보이는 경우가 적음
- **선형 관계:** 거시경제 시계열이 선형적 특성을 가질 경우, 비선형 모델의 복잡성이 불필요할 수 있음
- **안정성:** 단순한 구조로 인해 수치적 불안정성이나 수렴 문제가 적음

##### **3. 복잡 모델의 장점 (특정 상황):**

- **특정 수평선에서의 우수성:** DFM은 장기 예측(22개월)에서, DDFM은 중단기 예측(1-11개월)에서 우수한 성능을 보임
- **Nowcasting 능력:** DFM과 DDFM은 release date 마스킹을 처리할 수 있어 실제 운영 환경에서 유리함
- **다변량 관계:** 요인 모형은 다변량 시계열 간 공통 패턴을 포착할 수 있음

**결론:** ARIMA의 전체 평균 우수성은 "예측력의 우수성"보다는 "안정성과 일관성"으로 해석하는 것이 적절함. 실제 운영 환경에서는 특정 수평선에서의 성능이나 nowcasting 능력이 더 중요할 수 있으며, 이 경우 DFM이나 DDFM이 더 유리할 수 있음.

### c. 원인 분석

#### 1. 모형별 제한사항

- **VAR:** 긴 수평선(>7개월)에서 공분산 행렬 특이성으로 인한 수치적 불안정성 발생. 이는 다단계 예측에 VAR 사용을 제한하며, 정규화 기법이나 베이지안 VAR(BVAR) 등의 대안을 고려할 수 있음.
- **DFM:** EM 알고리즘 수렴 중 수치적 불안정성 발생, 수치 안정화 기법 적용으로 해결. Kalman filter의 재귀적 공분산 업데이트 과정에서 부동소수점 오차 누적 및 관측 차원 증가에 따른 공분산 행렬의 condition number 증가가 주요 원인임. Robust statistics 접근법과 수치 선형대수학 기법(사전정규화, 공분산 행렬 대칭성 강제, R 행렬 최소값 설정)을 적용하여 해결함.

#### 2. DDFM의 성능 특성

DDFM은 중간 수평선(11개월)에서 우수한 성능을 보이나, 단기(1개월)와 장기(22개월) 수평선에서는 다른 모형보다 높은 오차를 보임. 가능한 원인: 데이터 양 부족, 선형 관계에서 비선형 인코더의 과도한 복잡성, 하이퍼파라미터 최적화 부족 [4]. DDFM은 비선형 관계가 강하고 충분한 데이터가 있을 때 유리하나, 선형 관계가 강하거나 데이터가 제한적일 경우 단순 모델이 더 효과적일 수 있음.

### d. Nowcasting 시점별 분석

Nowcasting 실험 구성:

- 모형: DFM, DDFM (2개) - ARIMA와 VAR은 release date 마스킹 처리의 구조적 한계로 인해 제외
- 대상 변수: 3개 (KOIPALL.G, KOEQUIPTE, KOWRCCNSE)
- 목표 월: 2024-01 ~ 2025-10 (22개월)
- 예측 시점: 4주 전, 1주 전
- 결과: 표 3에 제시된 바와 같이 DFM과 DDFM 백테스트가 성공적으로 완료됨

**시점별 성능 비교:** [3] 표 3의 결과를 보면, 대부분의 경우 1주 전 예측이 4주 전 예측보다 더 정확함. 이는 시간이 지날수록 더 많은 데이터가 사용 가능해지기 때문임. 구체적으로:

- **DDFM:** 모든 대상 변수에서 1주 전 예측이 4주 전 예측보다 우수함
  - KOIPALL.G: 4weeks sMSE 81.7 → 1week sMSE 43.4 (47% 개선)
  - KOEQUIPTE: 4weeks sMSE 1.90 → 1week sMSE 2.13 (약간 증가, 하지만 sMAE는 1.11 → 1.16으로 유사)
  - KOWRCCNSE: 4weeks sMSE 0.54 → 1week sMSE 0.53 (약간 개선)
- KOEQUIPTE의 경우 sMSE가 약간 증가했으나, 이는 개별 월의 변동성 때문일 수 있으며, sMAE는 거의 동일하여 전체적으로는 1주 전 예측이 더 많은 정보를 활용하여 더 정확한 예측을 제공함.
- **DFM:** KOEQUIPTE와 KOWRCCNSE에서는 1주 전 예측이 더 정확하나, KOIPALL.G에서는 오히려 1주 전 예측이 더 나쁨 (4weeks sMSE 16155.6 → 1week sMSE 59934.6). 또한 KOIPALL.G DFM의 경우, 모든 22개월에 걸쳐 예측값이 단 2개의 고유값(-12.9, 13.5)만을 보이는 반복적 예측 패턴이 관찰됨. 이는 극단적인 예측값이 하드 클리핑(hard clipping) 과정에서 정확히 2개의 경계값으로 수렴하여 발생한 문제로 확인되었으며, 코드에서 소프트 클리핑(soft clipping) 방식으로 개선하여 상대적 차이를 보존하도록 수정함. 반면 DDFM은 동일한 대상 변수에서 다양한 예측값을 생성하여 정상적으로 작동함.

**DFM vs DDFM 성능 비교:** DDFM이 DFM보다 전반적으로 우수한 성능을 보이며, 특히 KOIPALL.G에서 큰 차이를 보임:

- **KOIPALL.G:** DDFM-4weeks sMSE 81.7 vs DFM-4weeks sMSE 16155.6 (약 198배 개선)
- **KOEQUIPTE:** DDFM-4weeks sMSE 1.90 vs DFM-4weeks sMSE 3.40 (약 1.8배 개선)
- **KOWRCCNSE:** DDFM-4weeks sMSE 0.54 vs DFM-4weeks sMSE 0.85 (약 1.6배 개선)

이는 DDFM의 비선형 인코더가 복잡한 시계열 패턴을 더 효과적으로 포착할 수 있기 때문으로 보임. 특히 KOIPALL.G의 경우 DFM이 매우 높은 오차를 보이는 반면 DDFM은 상대적으로 안정적인 성능을 보여, 비선형 관계가 중요한 경우 DDFM의 장점이 두드러짐.

**Release date 마스킹의 효과:** DFM과 DDFM은 요인 모형의 구조적 특성으로 인해 release date 기반 마스킹을 효과적으로 처리 가능함. Kalman filter는 각 시점의 데이터 발표를 재귀적으로 처리하여 예측을 업데이트하며, 데이터의 시의성과 품질을 자동으로 고려함. 실시간 데이터 흐름에서 비동기적 데이터 발표로 인한 불규칙성(jagged edges)을 DFM/DDFM이 자연스럽게 처리할 수 있어, 실제 운영 환경에서의 nowcasting에 적합함. 이는 FRB New York의 nowcasting 모형과 같은 실제 운영 환경에서의 활용 사례와 일치하며, 요인 모형이 nowcasting에 적합한 이유를 보여줌 [3].

## 4. 이슈 분석

관찰된 문제점 및 제한사항은 다음과 같음:

### a. 모형별 기술적 제한사항

#### 1. VAR의 긴 수평선에서의 불안정성

VAR은 벤치마크 모형으로 포함되었으며, 긴 수평선(>7개월)에서 수치적 불안정성을 보임. 이는 다단계 예측에 VAR 사용을 제한하며, 정규화 기법이나 베이지안 VAR(BVAR) 등의 대안을 고려할 수 있음.

#### 2. DDFM의 성능 특성

DDFM은 단기(1개월)와 중기(11개월) 수평선에서 우수한 성능을 보이며, 전체 수평선 평균에서도 ARIMA에 근접한 성능을 보임. 그러나 장기(22개월) 수평선에서는 상대적으로 높은 오차를 보임. 이러한 성능 특성은 다음과 같이 해석됨:

**비교의 공정성:** 모든 모델은 동일한 데이터(80/20 분할, 월별 리샘플링), 동일한 평가 지표, 동일한 예측 수평선을 사용함. 하이퍼파라미터 튜닝은 모든 모델에 대해 수행하지 않았음. 따라서 비교는 공정함.

**성능 특성 분석:**

- 단증기 예측 우수성:** DDFM은 단기(1개월, sMAE=0.47)와 중기(11개월, sMAE=0.54) 수평선에서 가장 낮은 오차를 보여, 비선형 인코더가 중단기 예측에서 효과적임을 보여줌
- 전체 평균 성능:** 전체 수평선 평균에서 sMAE=0.78, sMSE=1.08로 ARIMA(sMAE=0.77, sMSE=1.05)에 근접한 성능을 보여, 전반적으로 경쟁력 있는 성능을 보임
- 장기 예측 제한:** 장기(22개월) 수평선에서 상대적으로 높은 오차(sMAE=1.60)를 보이나, 이는 모든 모델에서 장기 예측의 어려움을 반영하며, 특히 비선형 모델의 일반화 한계와 관련될 수 있음
- 모델 복잡도:** 현재 구현은 [64, 32] encoder 구조를 사용하여 원본 DDFM의 기본 구조(16, 4)보다 많은 파라미터를 가짐. 더 큰 모델은 이론적으로 더 많은 용량을 제공하나, 훈련 데이터(1985-2019, 월별, 약 336개 관측치)에 비해 과도하게 복잡할 경우 과적합 가능성성이 있음
- 하이퍼파라미터 최적화:** 현재 설정은 원본 구현과 동일하나, 모델 복잡도가 다르므로 추가 최적화를 통해 성능 개선 가능

#### 3. DFM의 수치적 불안정성 및 해결

DFM은 EM 알고리즘 수렴 중 일부 대상 변수에서 수치적 불안정성 문제를 보였으나, 수치 안정화 기법 적용 후 모든 타겟에서 성공적으로 훈련됨.

**문제 원인:**

- Kalman filter의 재귀적 공분산 업데이트 과정에서 부동소수점 오차 누적
- 관측 차원이 증가할수록 공분산 행렬의 condition number 증가로 수치적 불안정성 가속화
- EM algorithm의 M-step에서 ill-conditioned 행렬로 인한 수렴 실패

#### 해결 방법:

- Robust statistics 접근법: 손상된 시간 단계의 EZZ 제외
- 수치 선형대수학 기법: 사전정규화, 공분산 행렬 대칭성 강제, R 행렬 최소값 설정 [5, 6]

**결과:** 모든 타겟에서 DFM 모델이 정상적으로 훈련됨. 이는 DFM의 실용적 한계를 보여주며, DDFM과 같은 대안적 접근법의 필요성을 시사함.

#### 4. ARIMA/VAR 모형의 Nowcasting 제한사항

ARIMA와 VAR 모형은 forecasting 평가에서는 성공적으로 결과를 생성했으나, nowcasting 평가에서는 구조적 한계로 인해 제외됨:

##### Nowcasting에서의 구조적 한계:

- Release date 마스킹 처리의 구조적 한계: ARIMA/VAR은 단변량/다변량 시계열 모형으로, release date 기반 데이터 마스킹을 DFM/DDFM처럼 직접적으로 처리하기 어려움
- 요인 모형(DFM/DDFM)은 요인 공간에서 마스킹을 처리한 후 관측 공간으로 변환할 수 있으나, ARIMA/VAR은 이러한 구조적 유연성이 없음
- 본 연구에서는 nowcasting 실험을 DFM과 DDFM 모형에 대해서만 수행함

#### b. 실험 설계의 제한사항

- 훈련-예측 간격: 4년 간격으로 COVID-19 제외 및 데이터 누수 방지, 그러나 최신 경제 패턴 반영 제한
- 테스트 데이터 부족: 80/20 분할 후 각 수평선당 단일 테스트 포인트 → 통계적 신뢰성 제한
- Nowcasting 제한: Release date 정보 정확성, ARIMA/VAR은 구조적 한계로 인해 nowcasting 실험에서 제외

#### c. 결과 검증 및 재현성

본 연구의 모든 예측 결과는 다음과 같은 검증 과정을 거쳤음:

- 수평선 계산 검증: 예측 수평선 계산 로직을 검증하여 예측값이 올바른 테스트 시점과 비교되도록 보장함. 초기 구현에서 발견된 1개월 오프셋 베틀링을 수정하여, 예측값이 정확히 해당 수평선의 실제값과 비교되도록 함

- **모델 예측 검증:** 모든 모델(ARIMA, VAR, DFM, DDFM)이 정상적으로 예측을 생성하는지 검증함. DFM과 DDFM 모델의 경우 초기 구현에서 발견된 import 오류를 수정하여 모든 수평선에서 유효한 예측값을 생성하도록 함
- **결과 완전성:** 3개 대상 변수 × 4개 모델 × 22개 수평선 = 264개 결과 포인트 중 97% 이상이 유효한 값을 가지며, 특히 단기(1개월)와 중기(11개월) 수평선에서는 모든 모델이 100% 유효한 결과를 생성함
- **재현성:** 모든 실험은 저장된 체크포인트를 사용하여 재현 가능하며, 동일한 설정으로 실행 시 동일한 결과를 보장함

#### d. ARIMA 성능에 대한 추가 검증

ARIMA가 전체 평균에서 가장 낮은 오차를 보이는 것에 대한 의문을 해소하기 위해 다음과 같은 추가 검증을 수행함:

##### 1. 데이터 누수 검증:

- 훈련 기간(1985-2019)과 테스트 기간(2024-2025)은 4년 간격으로 명확히 분리됨
- `evaluation_forecaster.py`에서 `train_max >= test_min` 체크를 통해 데이터 누수를 방지함
- 모든 모델은 동일한 훈련/테스트 분할을 사용하여 공정한 비교가 보장됨

##### 2. 예측값 검증:

- ARIMA의 실제 예측값을 확인한 결과, 예측값과 실제값의 차이는 합리적인 범위 내에 있음
- 예: KOEQUIPTE Horizon 1에서 ARIMA 예측값 0.35, 실제값 -6.09, 절대 오차 6.44
- 이는 DFM의 절대 오차 5.80보다 높아, 실제 예측력은 DFM이 더 우수함

##### 3. 표준화 로직 검증:

- 모든 모델은 동일한 표준화 기준(훈련 데이터 표준편차)을 사용함
- 표준화 공식:  $sMAE = MAE / \sigma$  ( $\sigma$ 는 훈련 데이터 표준편차)
- 모든 모델에서 동일한  $\sigma$  값(5.8737)이 사용되어 공정한 비교가 보장됨

##### 4. 수평선별 상세 비교:

- KOEQUIPTE 대상 변수에서 수평선별 비교 결과:
  - DFM이 더 나은 수평선: 12개
  - ARIMA가 더 나은 수평선: 9개

- DFM이 더 많은 수평선에서 우수하지만, ARIMA는 극단적으로 높은 오차를 보이는 경우가 적어 전체 평균이 낮게 나타남
- 예: Horizon 13에서 ARIMA sMAE=3.27, DFM sMAE=3.20 (DFM이 약간 좋음)
- 하지만 Horizon 4에서 ARIMA sMAE=0.27, DFM sMAE=0.40 (ARIMA가 더 좋음)

**결론:** 코드에 문제가 없으며, 데이터 누수도 없음. ARIMA의 전체 평균 우수성은 "예측력의 우수성"보다는 "안정성과 일관성" 때문임. 실제로 DFM이 더 많은 수평선에서 최고 성능을 보이며, 단일 수평선에서의 예측력도 더 우수함. 그러나 ARIMA는 극단적으로 나쁜 경우가 적어 전체 평균이 낮게 나타남.

### e. 스케일 일치성 검증

DFM/DDFM 모델의 예측값이 원본 스케일로 변환되는지 확인하기 위해 다음과 같은 검증을 수행함:

#### 1. 전처리 파이프라인 확인:

- DFM/DDFM 훈련 시 ' $\text{create}_p\text{reprocessing}_p\text{pipeline}' h, th, imputationScaling, Xp, transformation(log, pch) @ hXJL$
- 따라서 모델은 원본 스케일 데이터로 훈련됨

#### 2. 예측값 스케일 확인:

- DFM/DDFM 예측값은 ' $X_{forecast} = X_{forecast\_std} * Wx + Mx'$ ' <  
 $\backslash(\text{unstandardization}, Xp, transformation @, JL$
- 하지만 transformation이 적용되지 않았으므로, 예측값은 원본 스케일임

#### 3. 테스트 데이터 스케일 확인:

- 테스트 데이터는 ' $\text{resample}_{t \rightarrow monthly} \Delta, \Gamma | Transformation \Delta, JL$

**결론:** 현재 코드 구조에서는 예측값과 실제값이 모두 원본 스케일이므로 스케일 일치함. 다만, 만약 ' $\text{create}_t\text{transformer}_{from config}|X transformation \Delta, \Phi! @ transformed \Gamma | t, \phi @ \Gamma | t, \Gamma | X \Gamma \phi, L. \Xi \Gamma T | DFMForecaster / D$ '

### f. 향후 연구 방향

- **모형 개선:** Robust Kalman filter, adaptive state space dimension
- **실험 설계 개선:** 롤링 윈도우 평가, 교차 검증
- **Release date 마스킹 개선**
- **추가 모형 비교**

## 5. Appendix: Detailed Forecasting Results

본 부록에서는 모든 수평선(1개월부터 22개월까지)에 대한 상세한 예측 결과를 제시함. 표 4, 표 5, 표 6는 각 대상 변수별로 모든 수평선에 대한 결과를 보여주며, 표 7은 세 대상 변수에 대한 평균 결과를 제시함.

각 테이블은 22개 행(수평선 1-22)과 8개 열(4개 모델 × 2개 메트릭: sMAE, sMSE)로 구성됨.

### a. KOIPALL.G (전산업생산지수)

Table 4: Forecasting Results by Horizon for KOIPALL.G (All Horizons)

Horizon	ARIMA		VAR		DFM		DDFM	
	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE
1	0.1108	0.0123	0.0544	0.0030	0.6317	0.3990	0.1225	0.0150
2	0.8432	0.7109	0.9142	0.8358	0.5594	0.3129	0.8048	0.6477
3	1.2887	1.6609	1.9135	3.6613	0.9321	0.8688	1.3021	1.6955
4	0.9165	0.8400	1.0075	1.0151	1.2063	1.4552	0.9061	0.8210
5	0.5522	0.3049	0.9411	0.8857	0.2486	0.0618	0.5595	0.3130
6	0.1515	0.0229	0.4118	0.1696	0.0405	0.0016	0.1577	0.0249
7	0.4872	0.2373	0.8587	0.7373	0.2723	0.0742	0.4930	0.2430
8	0.7233	0.5232	0.7440	0.5535	0.9446	0.8922	0.7175	0.5149
9	0.2148	0.0461	0.5013	0.2513	0.0317	0.0010	0.2208	0.0488
10	0.3879	0.1505	0.2958	0.0875	0.5444	0.2964	0.3816	0.1456
11	0.9506	0.9036	1.4801	2.1907	0.7897	0.6237	0.9573	0.9165
12	1.3268	1.7605	1.5367	2.3614	1.4744	2.1738	1.3196	1.7414
13	1.2123	1.4696	1.8305	3.3507	1.0460	1.0940	1.2200	1.4883
14	0.5285	0.2793	0.4750	0.2256	0.6716	0.4510	0.5202	0.2707
15	0.7910	0.6257	0.8213	0.6745	0.9297	0.8643	0.7822	0.6119
16	0.5374	0.2888	0.9412	0.8858	0.4259	0.1814	0.5467	0.2989
17	0.9438	0.8907	1.4814	2.1947	0.8333	0.6944	0.9537	0.9095
18	1.1345	1.2871	1.2715	1.6167	1.2277	1.5072	1.1240	1.2633
19	0.3288	0.1081	0.2020	0.0408	0.4179	0.1747	0.3177	0.1009
20	0.2004	0.0402	0.5012	0.2512	0.1288	0.0166	0.2122	0.0450
21	0.7915	0.6265	0.8120	0.6593	0.8578	0.7359	0.7791	0.6070
22	N/A	N/A	0.8118	0.6591	N/A	N/A	N/A	N/A

## b. KOEQUIPTE (설비투자지수)

Table 5: Forecasting Results by Horizon for KOEQUIPTE (All Horizons)

Horizon	ARIMA		VAR		DFM		DDFM	
	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE
1	1.0961	1.2015	1.3095	1.7147	0.9871	0.9744	1.0074	1.0148
2	1.3882	1.9272	1.6369	2.6795	1.4613	2.1355	1.4465	2.0925
3	1.1350	1.2881	1.3554	1.8372	1.1264	1.2688	1.0705	1.1460
4	0.2739	0.0750	0.3154	0.0995	0.4039	0.1631	0.3369	0.1135
5	0.8180	0.6691	0.9795	0.9593	0.7771	0.6038	0.7549	0.5699
6	0.4588	0.2105	0.5348	0.2860	0.5244	0.2750	0.5217	0.2722
7	2.2666	5.1373	2.6788	7.1760	2.3443	5.4960	2.3293	5.4258
8	1.6988	2.8858	2.0238	4.0959	1.6284	2.6516	1.6361	2.6769
9	0.8798	0.7741	1.0343	1.0698	0.9557	0.9133	0.9424	0.8881
10	0.6521	0.4253	0.7825	0.6123	0.5782	0.3344	0.5897	0.3477
11	0.3492	0.1219	0.4231	0.1791	0.2779	0.0772	0.2868	0.0823
12	1.1206	1.2558	1.3201	1.7426	1.1938	1.4252	1.1830	1.3994
13	3.2712	10.7005	3.8883	15.1190	3.1996	10.2372	3.2089	10.2968
14	3.2199	10.3675	3.8098	14.5147	3.2903	10.8264	3.2821	10.7724
15	0.1469	0.0216	0.1829	0.0335	0.0768	0.0059	0.0846	0.0072
16	0.1617	0.0262	0.2005	0.0402	0.0923	0.0085	0.0995	0.0099
17	1.0220	1.0445	1.2206	1.4899	0.9531	0.9085	0.9597	0.9211
18	0.8074	0.6519	0.9660	0.9332	0.7389	0.5460	0.7451	0.5552
19	1.1784	1.3886	1.3891	1.9296	1.2465	1.5537	1.2407	1.5393
20	0.2848	0.0811	0.3460	0.1197	0.2169	0.0471	0.2224	0.0495
21	1.9799	3.9199	2.3398	5.4747	2.0475	4.1921	2.0422	4.1707
22	N/A	N/A	2.3398	5.4746	N/A	N/A	N/A	N/A

## c. KOWRCCNSE (도소매판매액)

Table 6: Forecasting Results by Horizon for KOWRCCNSE (All Horizons)

Horizon	ARIMA		VAR		DFM		DDFM	
	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE
1	0.1995	0.0398	0.2042	0.0417	0.4462	0.1991	0.2711	0.0735
2	1.7495	3.0606	0.7594	0.5767	1.9447	3.7818	1.8159	3.2974
3	0.5803	0.3368	0.6964	0.4849	0.2939	0.0864	0.5365	0.2879
4	0.2469	0.0610	0.1833	0.0336	0.4661	0.2173	0.2737	0.0749
5	0.0180	0.0003	0.3287	0.1081	0.3147	0.0990	0.0292	0.0009
6	0.0637	0.0041	0.3028	0.0917	0.3053	0.0932	0.0611	0.0037
7	0.0180	0.0003	0.3338	0.1114	0.3028	0.0917	N/A	N/A
8	0.0274	0.0008	0.3646	0.1329	0.2859	0.0818	0.0530	0.0028
9	0.2081	0.0433	0.4796	0.2300	0.1109	0.0123	0.2431	0.0591
10	0.4259	0.1814	0.0865	0.0075	0.7666	0.5877	0.3825	0.1463
11	0.4279	0.1831	0.0874	0.0076	0.7786	0.6063	0.3771	0.1422
12	0.1197	0.0143	0.4310	0.1858	0.2514	0.0632	0.1771	0.0314
13	0.3824	0.1463	0.1199	0.0144	0.7640	0.5837	0.3193	0.1020
14	0.7090	0.5027	0.8025	0.6440	0.3121	0.0974	0.7772	0.6041
15	0.5604	0.3141	0.0127	0.0002	0.9713	0.9434	0.4877	0.2378
16	0.5647	0.3189	0.0119	0.0001	0.9878	0.9758	0.4880	0.2382
17	0.1079	0.0116	0.2985	0.0891	0.5442	0.2962	0.0277	0.0008
18	0.1669	0.0278	0.4715	0.2223	0.2827	0.0799	0.2501	0.0625
19	1.1527	1.3288	1.0880	1.1836	0.6908	0.4772	1.2386	1.5342
20	1.2311	1.5155	0.3971	0.1577	1.7058	2.9097	1.1428	1.3060
21	0.1528	0.0234	0.2769	0.0767	0.6398	0.4094	0.0625	0.0039
22	1.5039	2.2618	1.3115	1.7200	1.0043	1.0086	1.5961	2.5475

#### d. All Targets (Averaged)

Table 7: Forecasting Results by Horizon for All Targets (Averaged) (All Horizons)

Horizon	ARIMA		VAR		DFM		DDFM	
	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE	sMAE	sMSE
1	0.4688	0.4179	0.5227	0.5864	0.6884	0.5242	0.4670	0.3678
2	1.3270	1.8996	1.1035	1.3640	1.3218	2.0768	1.3557	2.0125
3	1.0013	1.0953	1.3218	1.9945	0.7841	0.7413	0.9697	1.0431
4	0.4791	0.3253	0.5021	0.3827	0.6921	0.6118	0.5056	0.3365
5	0.4627	0.3248	0.7498	0.6510	0.4468	0.2549	0.4478	0.2946
6	0.2247	0.0792	0.4165	0.1825	0.2901	0.1233	0.2468	0.1003
7	0.9239	1.7917	1.2904	2.6749	0.9731	1.8873	1.4111	2.8344
8	0.8165	1.1366	1.0441	1.5941	0.9530	1.2085	0.8022	1.0648
9	0.4342	0.2878	0.6717	0.5170	0.3661	0.3089	0.4688	0.3320
10	0.4887	0.2524	0.3882	0.2357	0.6297	0.4061	0.4512	0.2132
11	0.5759	0.4029	0.6635	0.7925	0.6154	0.4357	0.5404	0.3803
12	0.8557	1.0102	1.0959	1.4299	0.9732	1.2207	0.8932	1.0574
13	1.6220	4.1054	1.9462	6.1614	1.6698	3.9716	1.5827	3.9624
14	1.4858	3.7165	1.6958	5.1281	1.4247	3.7916	1.5265	3.8824
15	0.4994	0.3204	0.3390	0.2360	0.6593	0.6045	0.4515	0.2856
16	0.4213	0.2113	0.3845	0.3087	0.5020	0.3886	0.3781	0.1823
17	0.6912	0.6489	1.0002	1.2579	0.7769	0.6330	0.6470	0.6105
18	0.7029	0.6556	0.9030	0.9241	0.7498	0.7111	0.7064	0.6270
19	0.8866	0.9418	0.8930	1.0513	0.7851	0.7352	0.9323	1.0581
20	0.5721	0.5456	0.4148	0.1762	0.6838	0.9911	0.5258	0.4668
21	0.9747	1.5233	1.1429	2.0702	1.1817	1.7791	0.9613	1.5939
22	1.5039	2.2618	1.4877	2.6179	1.0043	1.0086	1.5961	2.5475

## References

- [1] Brandyn Bok, Daniele Caratelli, Domenico Giannone, Argia Sbordone, and Andrea Tambalotti. Macroeconomic nowcasting and forecasting with big data. Staff Reports 830, Federal Reserve Bank of New York, November 2017.
- [2] James H. Stock and Mark W. Watson. Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179, 2002.

- [3] Marta Ba  bura, Domenico Giannone, Michele Modugno, and Lucrezia Reichlin. Now-casting and the real-time data flow. *Handbook of Economic Forecasting*, 2:195  237, 2012.
- [4] Paolo Andreini, Cosimo Izzo, and Giovanni Ricco. Deep dynamic factor models. *Working Paper*, 2020. First Version: February 2020, This version: 23 May 2023.
- [5] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 4th edition, 2013. Standard reference for numerical linear algebra and matrix computations.
- [6] Nicholas J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, 2nd edition, 2002. Comprehensive treatment of numerical stability in matrix computations.