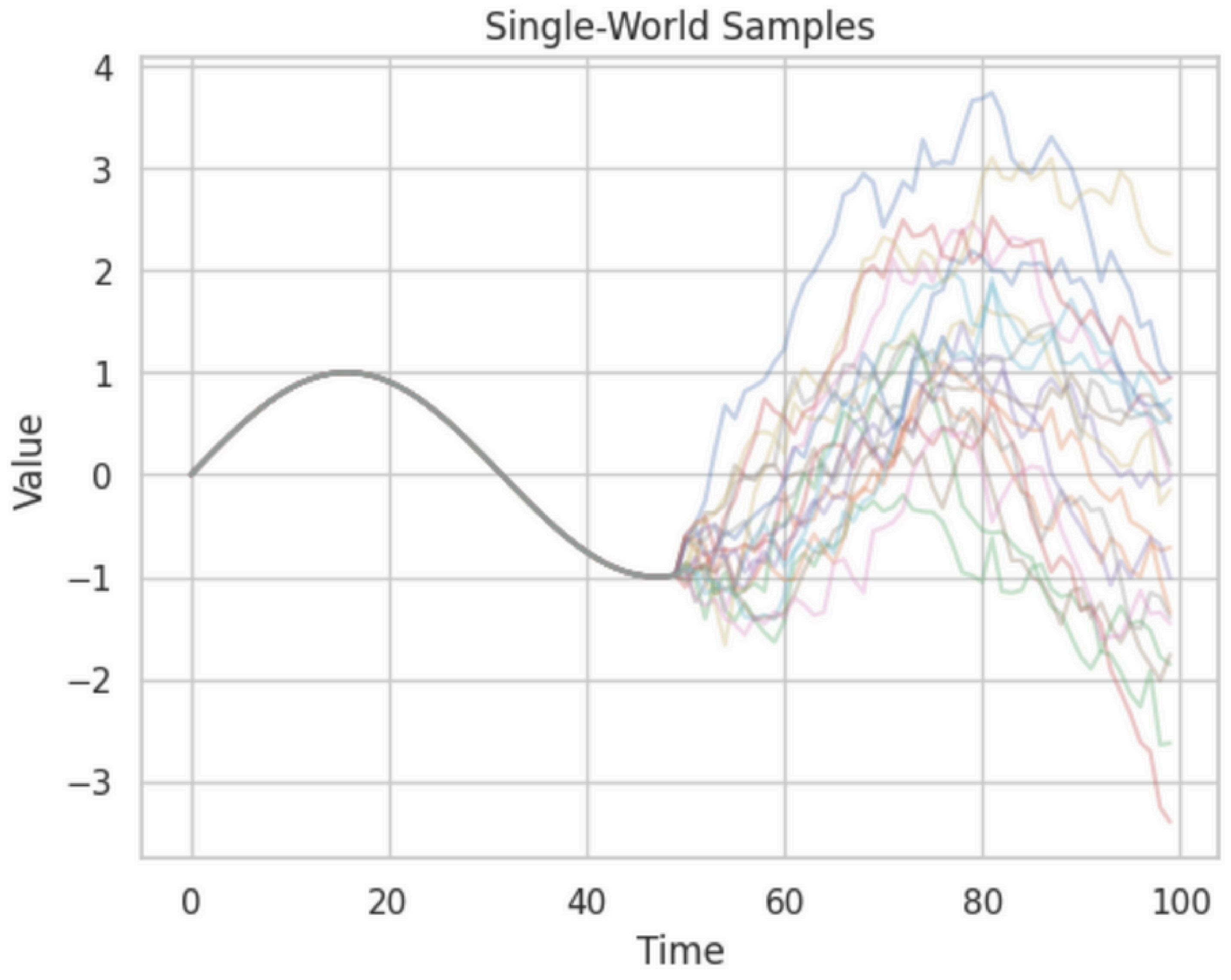
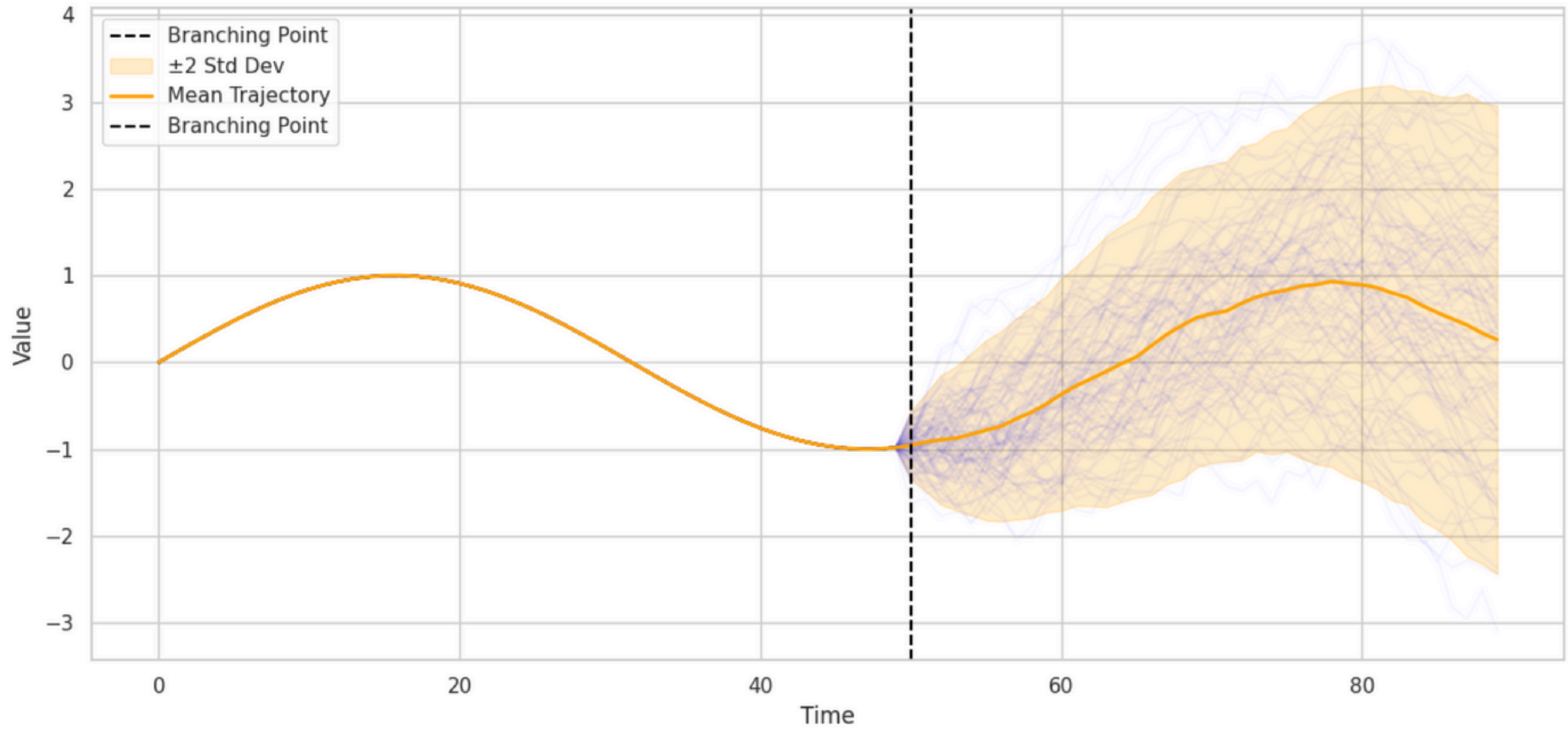
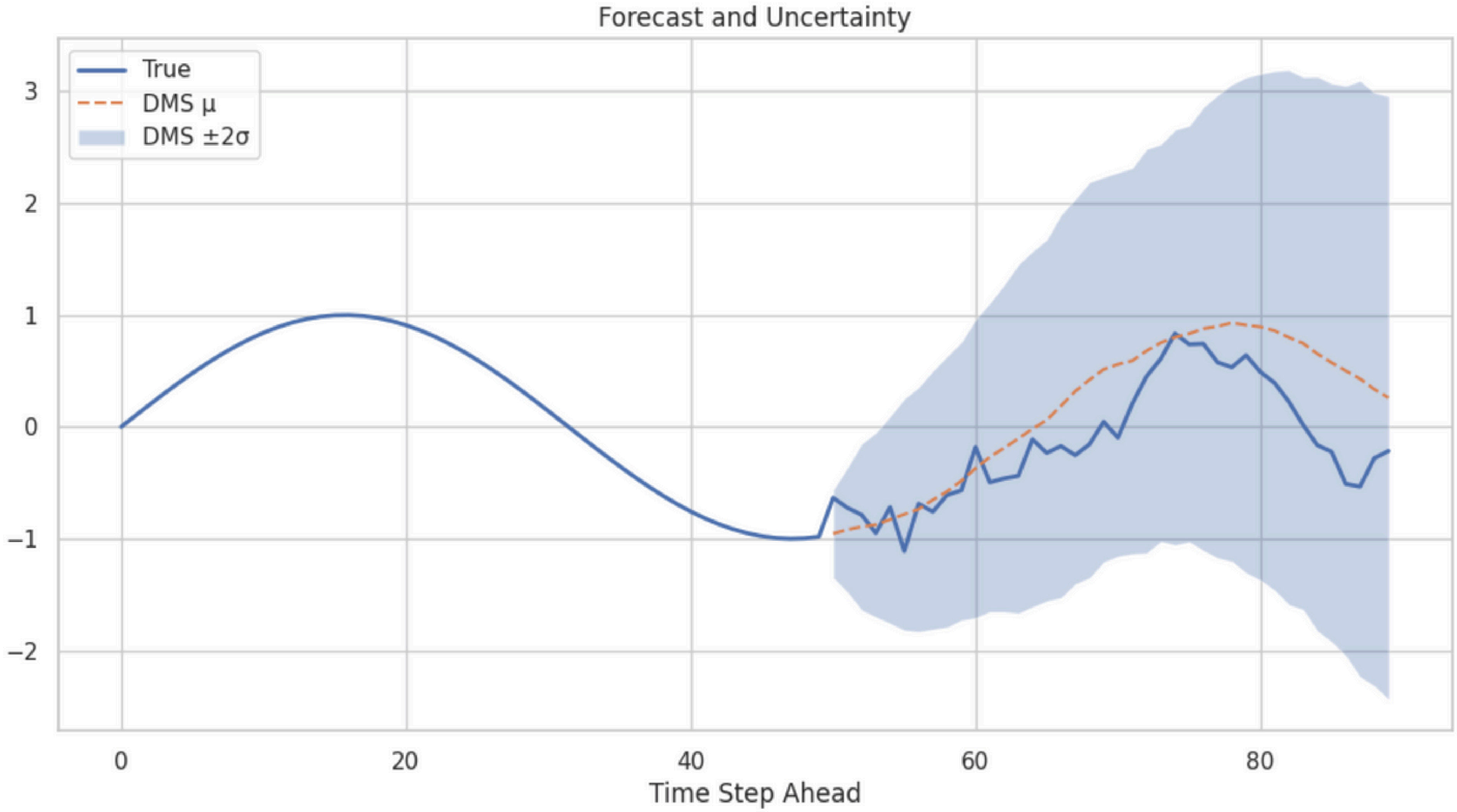
Milestone Meeting 3

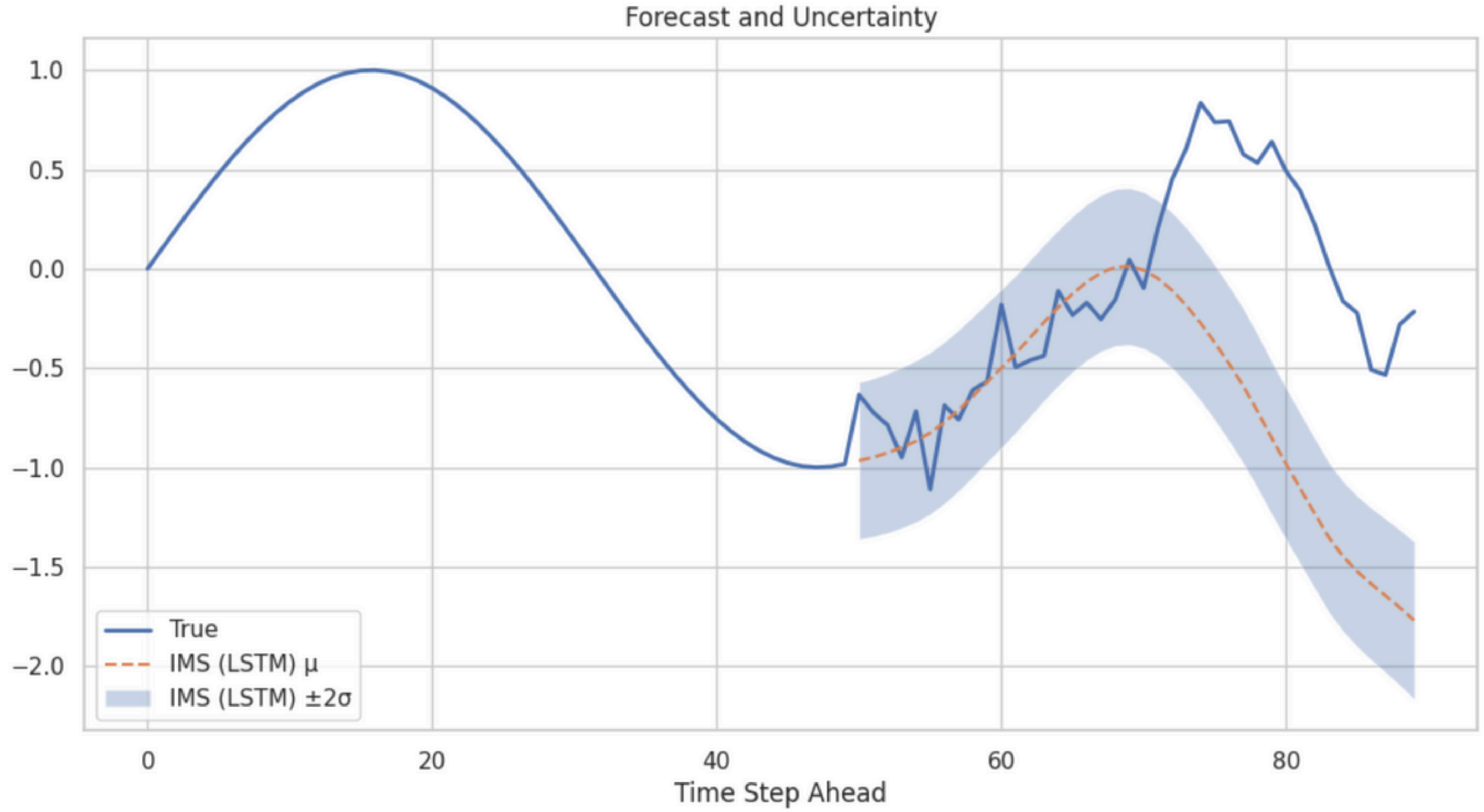
- aktueller Stand
    - Single/Multi-world
    - HPO
- mögliche nächste Schritte

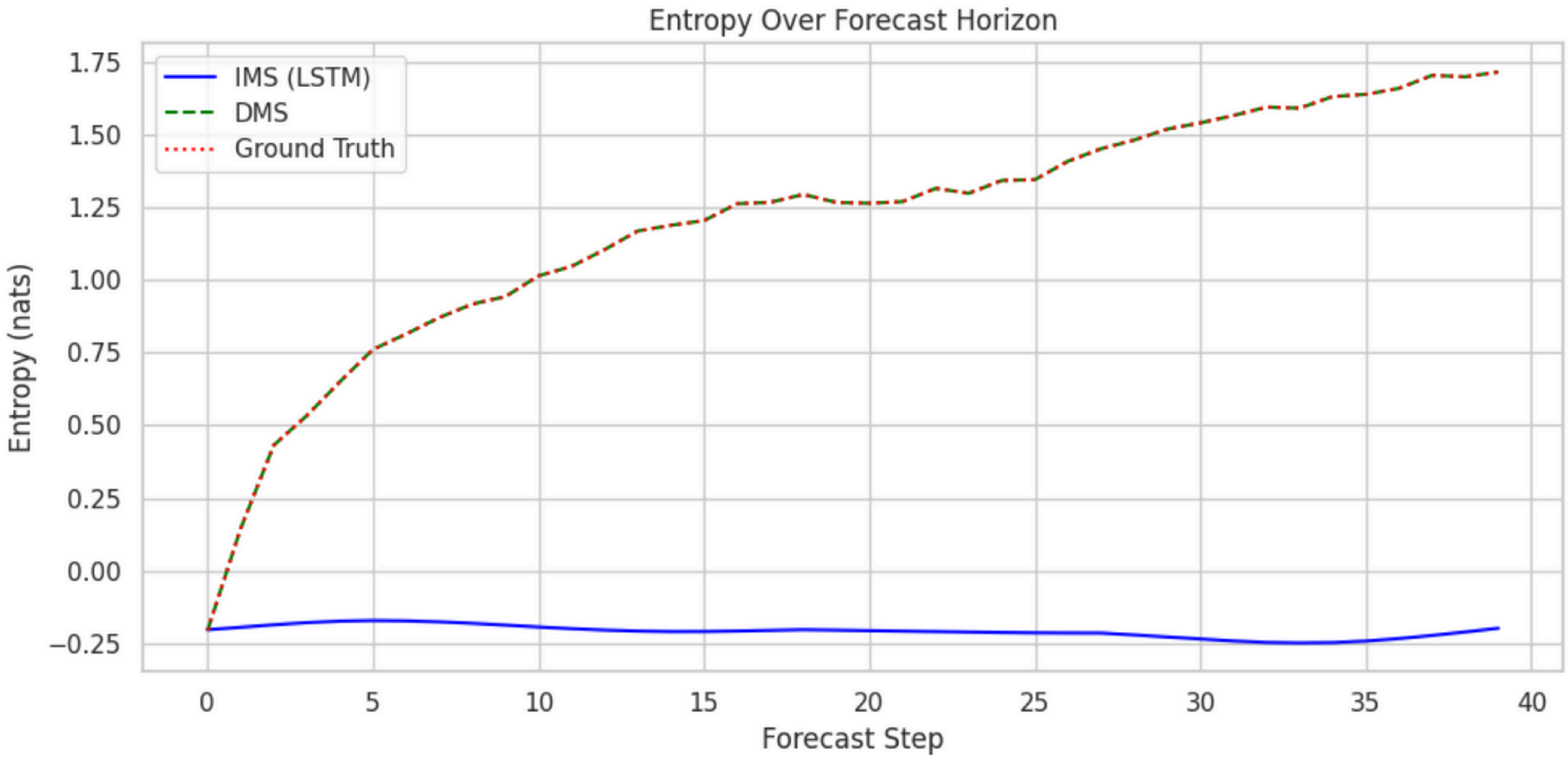
Single World



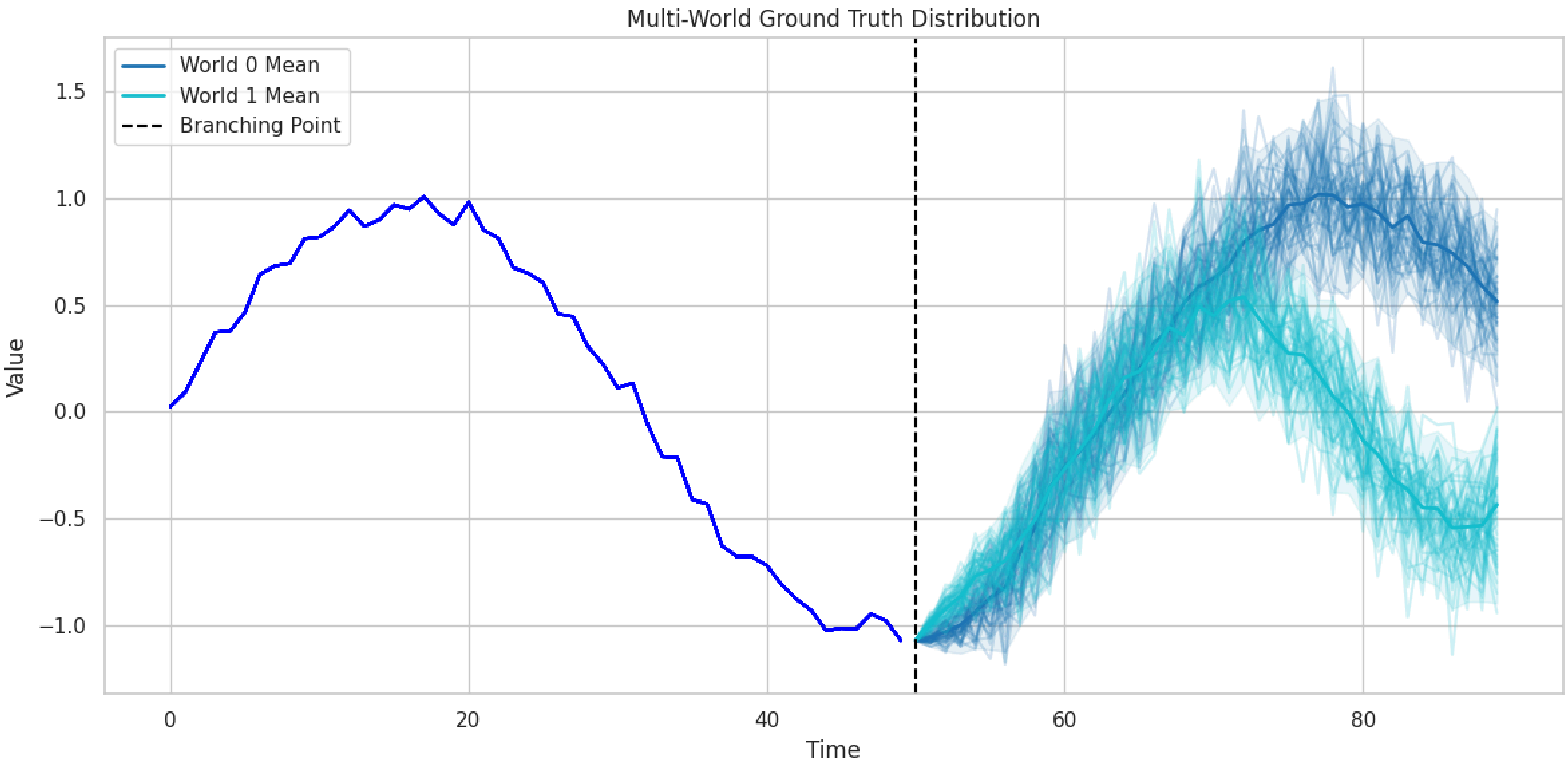



Forecast and Uncertainty
True
DMS μ
DMS ±2σ
3
2
1
0
−1
−2
0
20
40
60
80
Time Step Ahead

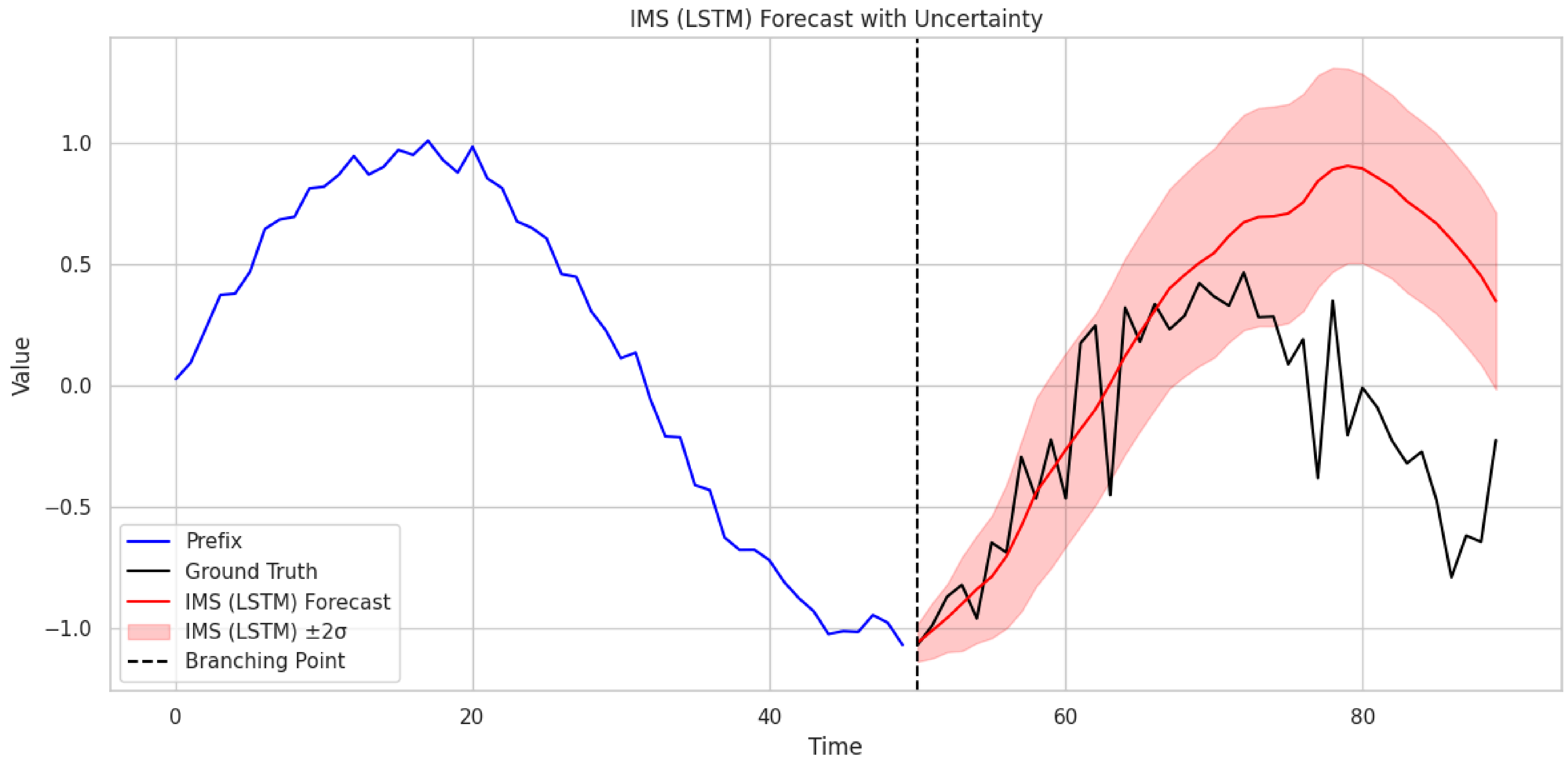

Forecast and Uncertainty
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
True
IMS (LSTM) μ
IMS (LSTM) ±2σ
0
20
40
60
80
Time Step Ahead

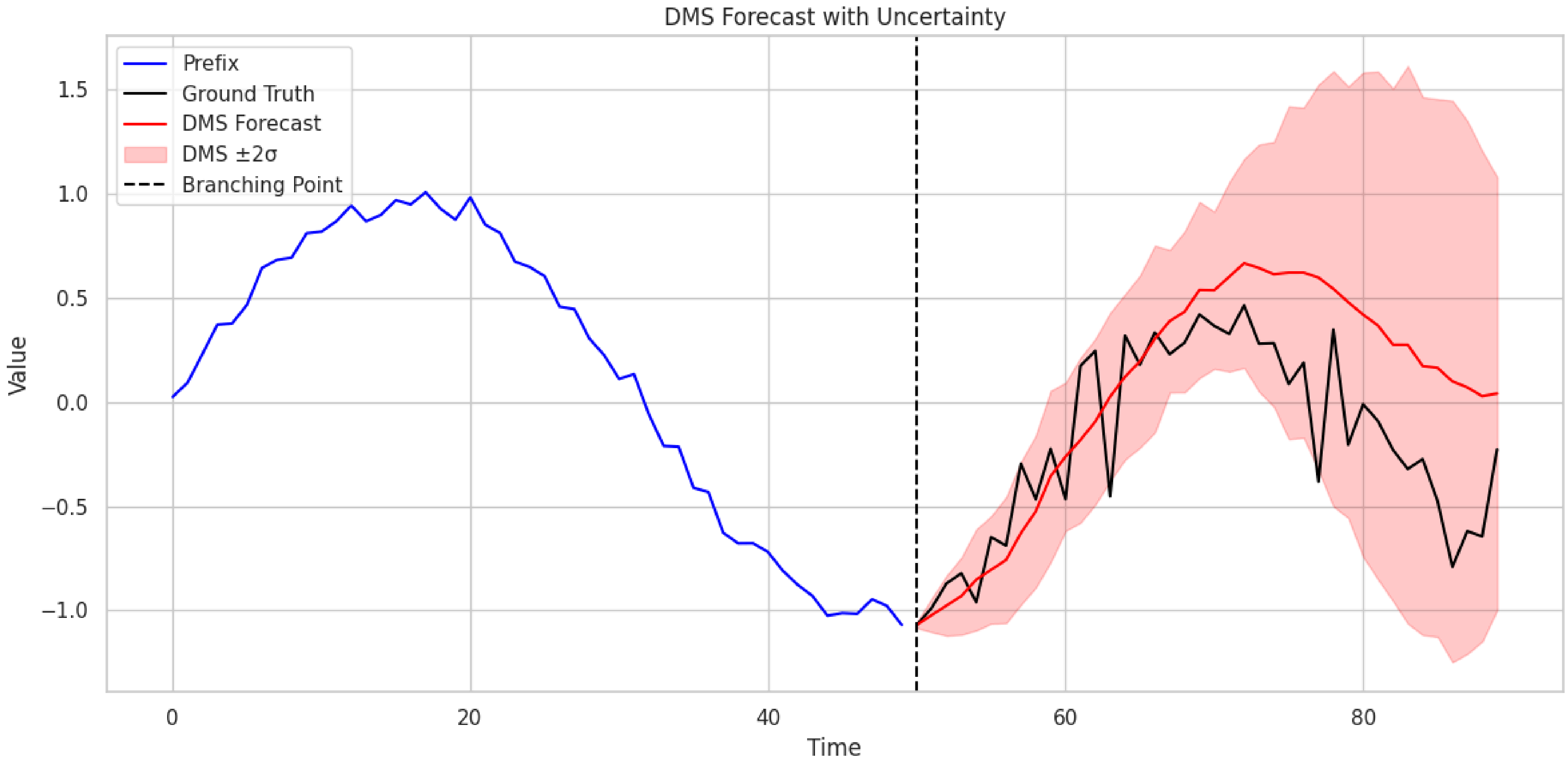

Entropy Over Forecast Horizon
IMS (LSTM)
DMS
Ground Truth
Entropy (nats)
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
−0.25
0
5
10
15
20
25
30
35
40
Forecast Step

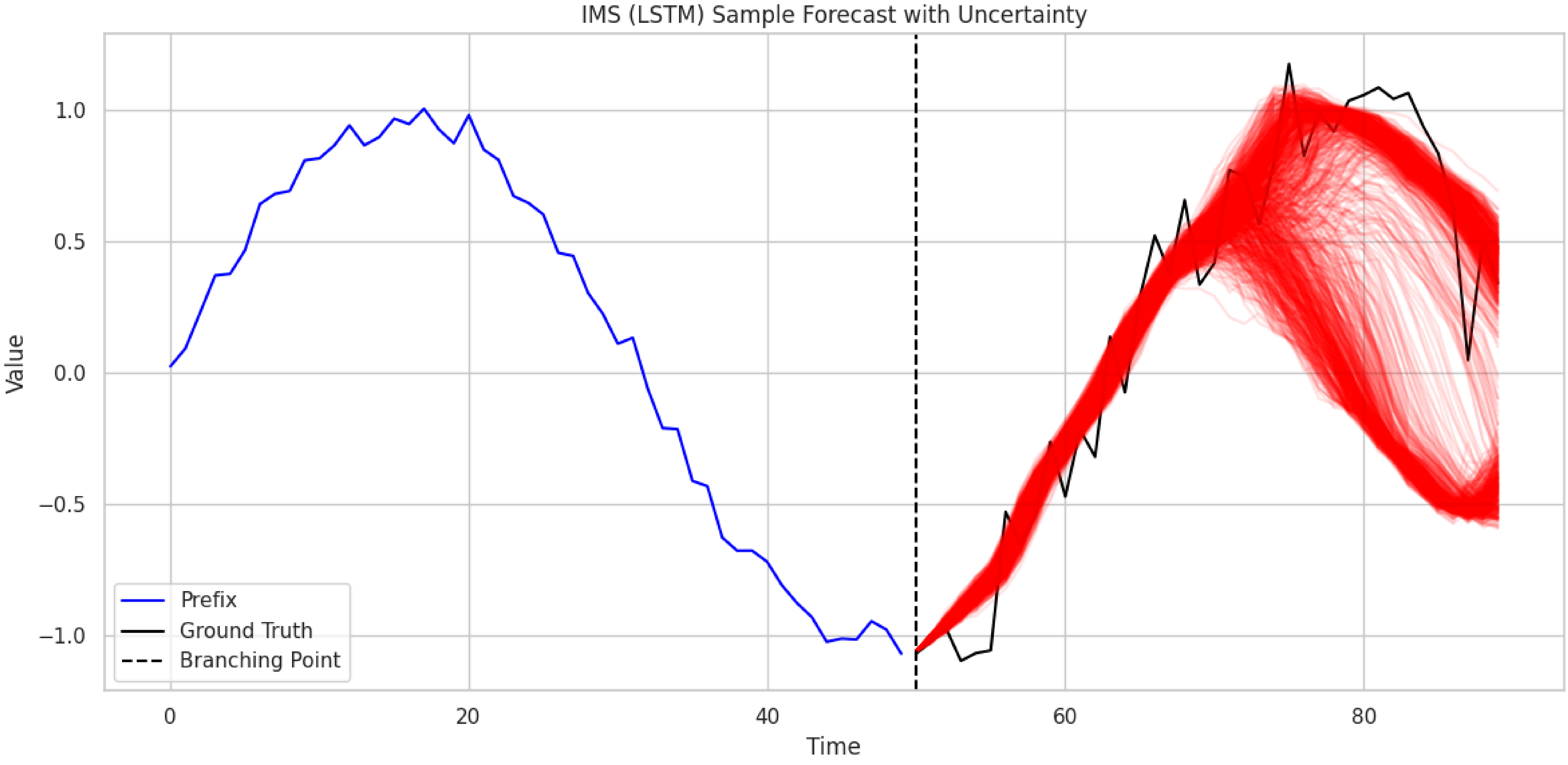
Multi world



Multi-World Ground Truth Distribution
World 0 Mean
World 1 Mean
Branching Point
1.5
1.0
0.5
0.0
−0.5
−1.0
Value
0
20
40
60
80
Time


IMS (LSTM) Forecast with Uncertainty
Value
Time
Prefix
Ground Truth
IMS (LSTM) Forecast
IMS (LSTM) ±2σ
Branching Point
1.0
0.5
0.0
−0.5
−1.0
0
20
40
60
80


DMS Forecast with Uncertainty
Value
Time
Prefix
Ground Truth
DMS Forecast
DMS ±2σ
Branching Point
1.5
1.0
0.5
0.0
−0.5
−1.0
0
20
40
60
80


IMS (LSTM) Sample Forecast with Uncertainty
Value
Time
1.0
0.5
0.0
−0.5
−1.0
0
20
40
60
80
Prefix
Ground Truth
Branching Point

Multi-World Entropy Formula

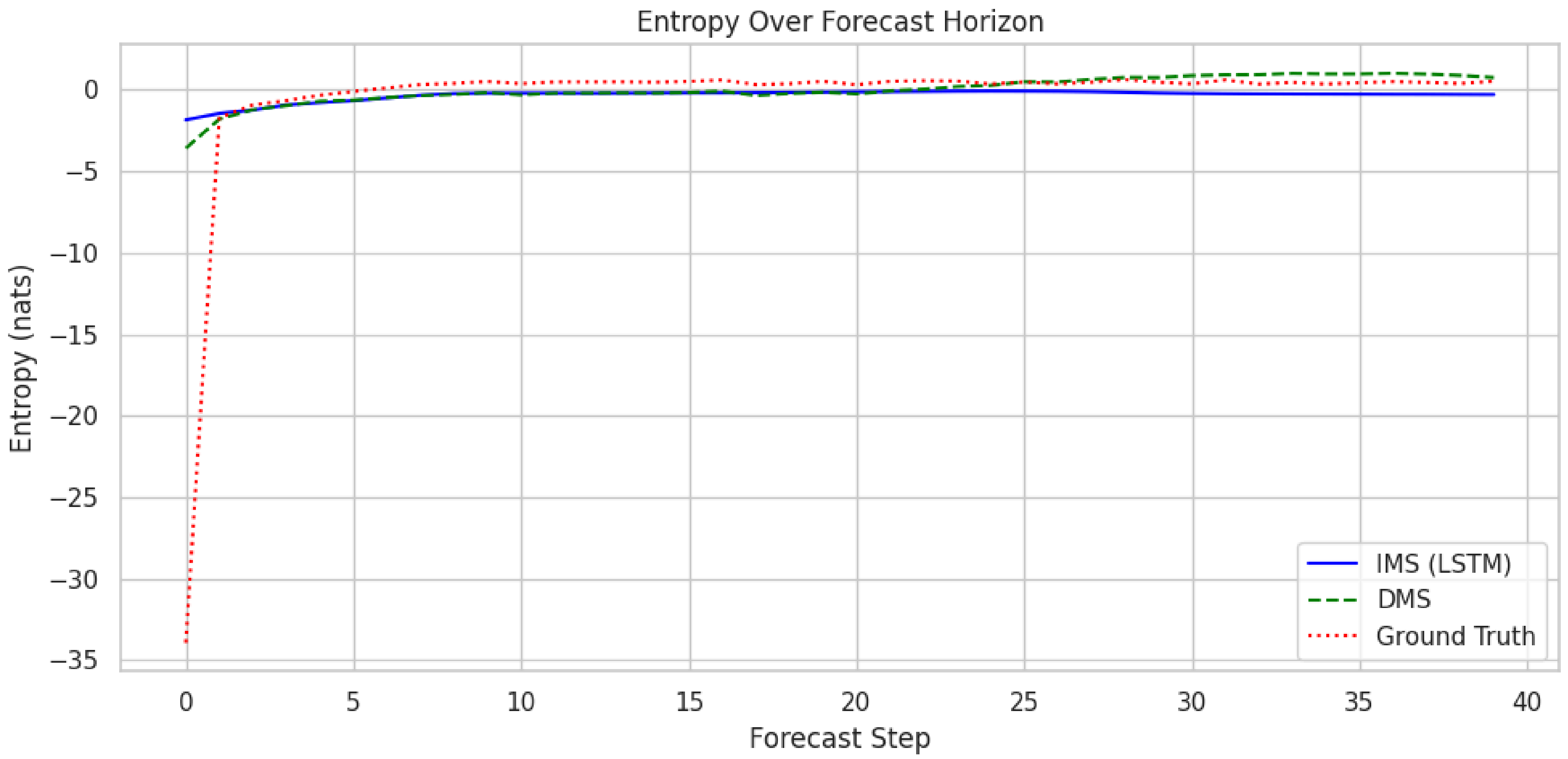
For a given time step t after the branching point:

$$H(X_t) = H_{between} + H_{within}$$

Where:

- $H_{between}$ is the entropy from choosing between worlds
- H_{within} is the expected entropy within each world

More specifically:



$$H_{between} = -\sum_{i=1}^{n_{worlds}} p(w_i) \log p(w_i)$$

Since all worlds have equal probability $p(w_i) = \frac{1}{n_{worlds}}$, this simplifies to:

$$H_{between} = -n_{worlds} \cdot \frac{1}{n_{worlds}} \log \frac{1}{n_{worlds}} = \log(n_{worlds})$$

For the within-world entropy, assuming Gaussian distributions within each world:

$$H_{within} = \sum_{i=1}^{n_{worlds}} p(w_i) \cdot H(X_t|w_i)$$

For a Gaussian with standard deviation σ_i in world i:

$$H(X_t|w_i) = \frac{1}{2}\log(2\pi e \sigma_i^2)$$

Therefore, the total entropy is:

$$H(X_t) = \log(n_{worlds}) + \frac{1}{n_{worlds}} \sum_{i=1}^{n_{worlds}} \frac{1}{2}\log(2\pi e \sigma_i^2)$$

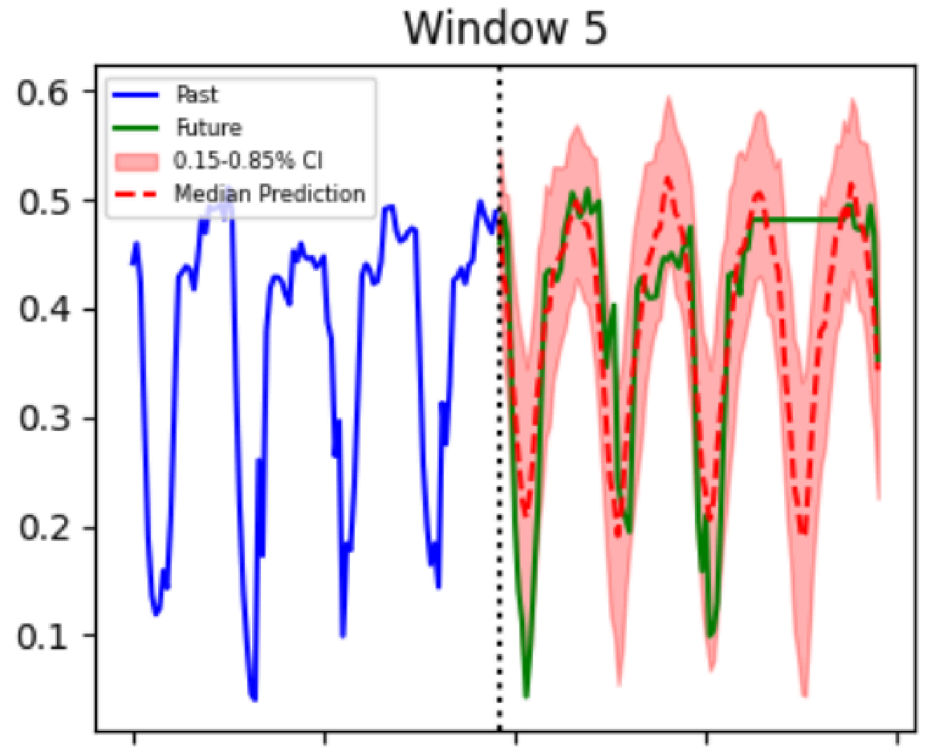
Ist das ein Forecasting szenario?

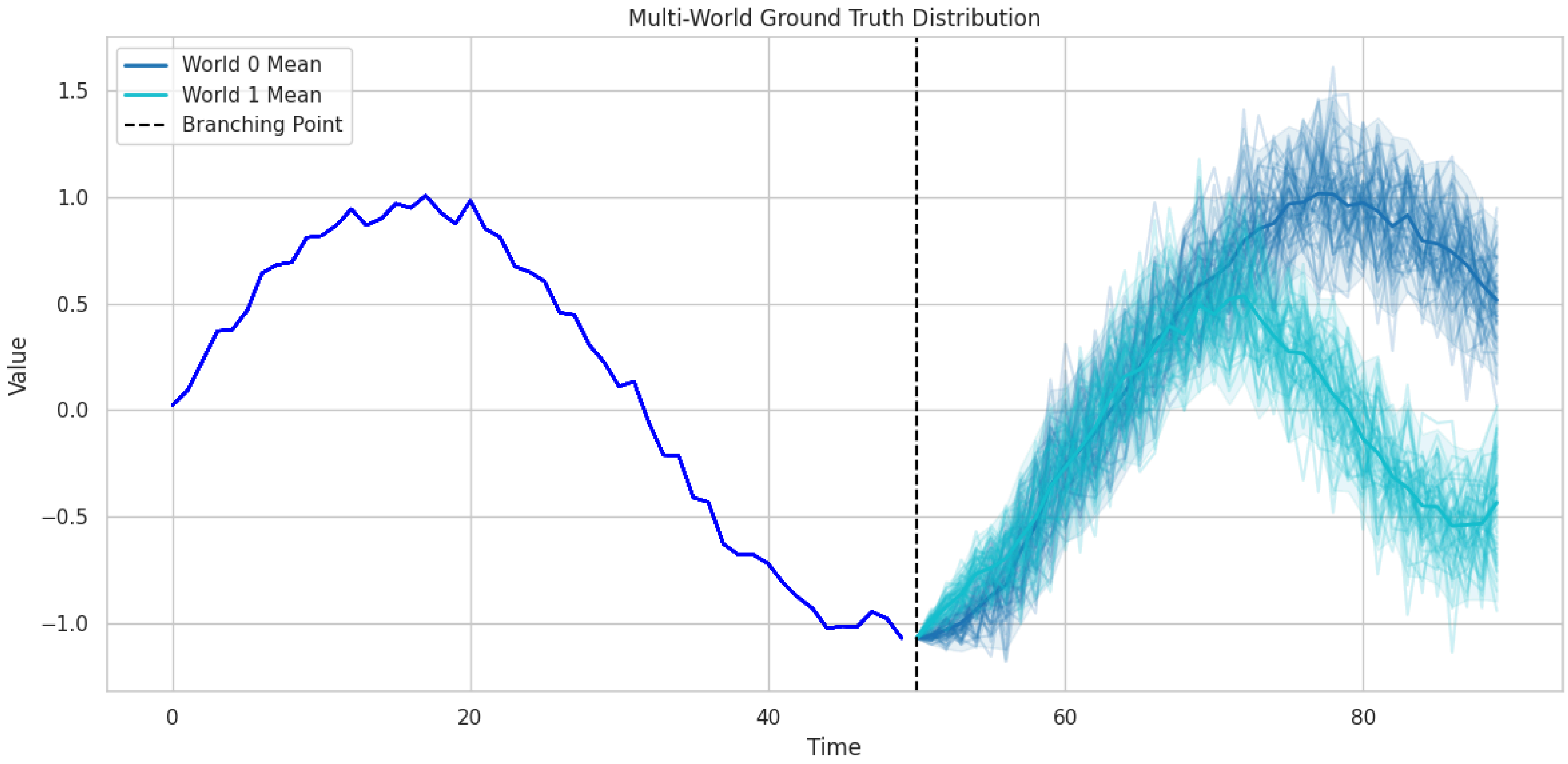
Multi-world:

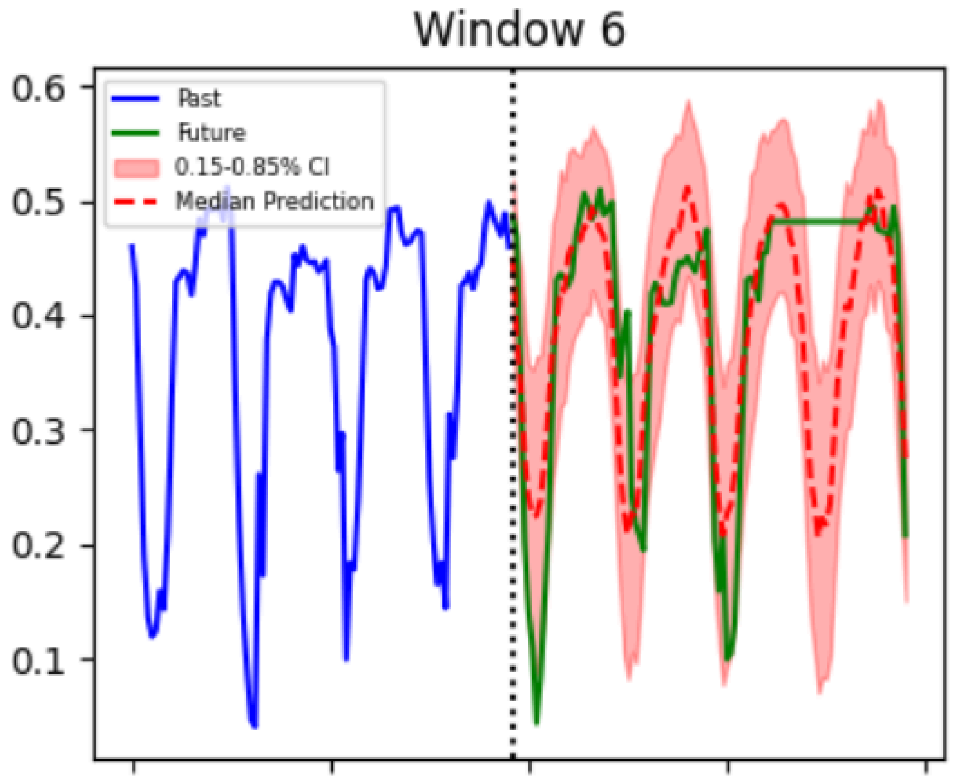
- Fokus auf identifizierung des richtigen Weges
- immer der gleiche prefix → unrealistisch
- Aber: man könnte es als zusammegefaltete Zeitreihe sehen (prefix müsste auch noise bekommen, ohne Informationen für world pfad zu verraten)

Forecasting:

- extrapolation in die Zukunft
- unterschiedlicher prefix basierend auf dem man vorhersagt







HPO

Data

+ ExchangeRate
+ ILI

Dataset	ETTh1 & ETTh2	ETTm1 & ETTm2	Electricity	Solar-Energy	Traffic	Weather
Timesteps	17,420	69,680	26,304	52,560	17,544	52,696
Channels	7	7	321	137	862	21
Frequency	1 hour	15 mins	1 hour	10 mins	1 hour	10 mins
Cyclic Patterns	Daily	Daily	Daily & Weekly	Daily	Daily & Weekly	Daily
Cycle Length	**24**	**96**	**168**	**144**	**168**	**144**

Name🌐	BeijingAirQuality	ExchangeRate	Illness
Domain📏	Beijing Air Quality	Exchange Rate	Ilness Data
Length📊	36000	7588	966
Time Series Count	7	8	7
🔄Graph	False	False	False
⏱️Freq. (m)	60	1440	10080
🎯Task	LTSF	LTSF	LTSF

HPO

Forecasting task

Typical forecasting horizons {96, 192, 336, 720} with 96 look back length

Models

1. PatchTST
2. DeepAR
   - only IMS model in benchmark
3. DLinear
   - MLP based
4. iTransformer
   - newer Tranformer modle that does not use patching

For ETTh1 and forecasting from 96 steps to 720

MODEL	HPO FERTIG/TODO
PatchTST	quantile, i_quantile, multivariate, univariate
DeepAR	quantile, univariate, i_quantile
iTransformer	quantile, i_quantile, multivariate, univariate
DLinear	quantile, i_quantile, multivariate, univariate

Next steps

1. HPO fertig für 2-3 weitere Datensätze
2. IMS-DMS
    a. Upscale Dataset
    b. Domain Bezug
    c. GP on top of DMS?
    d. multivariate DMS
3. (Copula and Flows)