UNIVERSITY OF
Southampton

UNIVERSITY OF Southampton

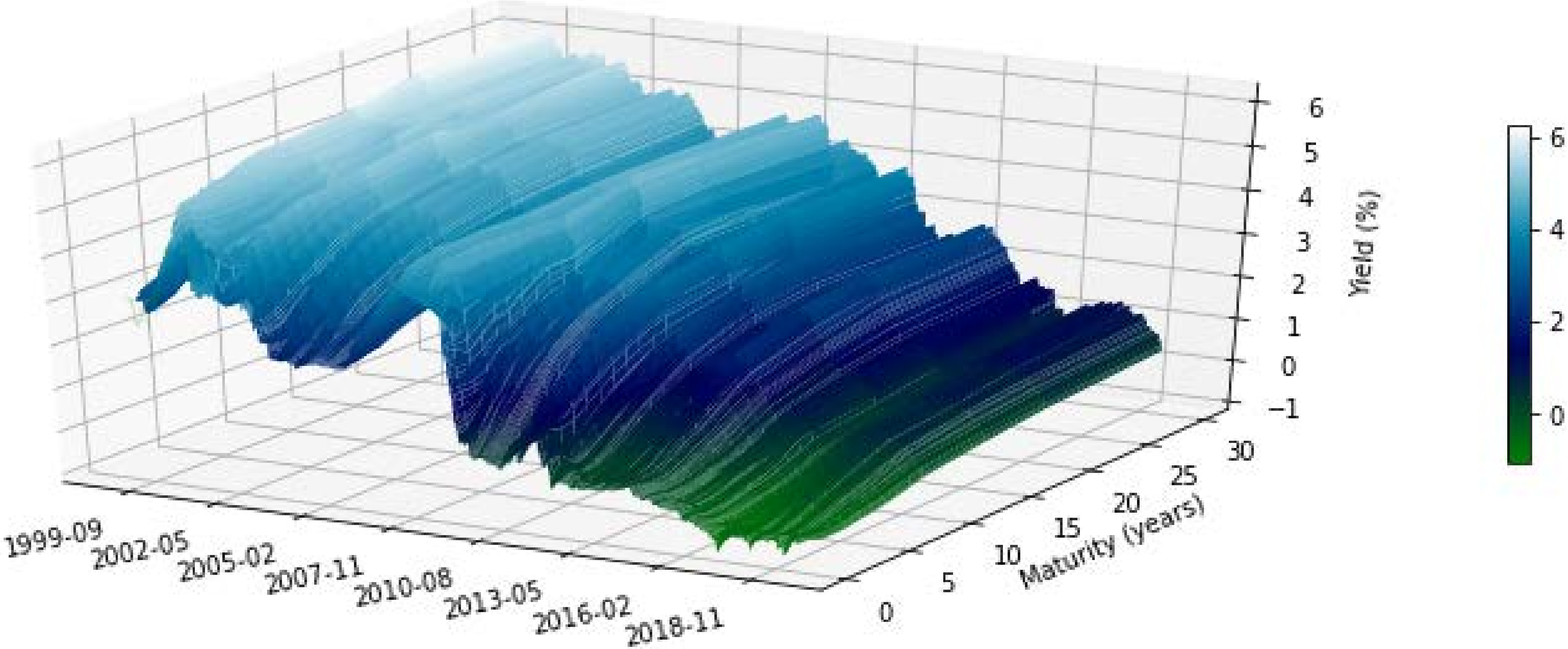
LSTM-LagLasso for bond yield forecasting: Peeping into the long short-term memory networks' black box

Manuel Nunes, Enrico Gerding, Frank McGroarty, Mahesan Niranjan

17 January 2020

EURO GOVERNMENT BOND YIELD CURVE

Target: 10Y yield



OBJECTIVES AND CONTRIBUTIONS

OBJECTIVES

1. To assess the potential of LSTMs for bond yield forecasting (memory advantage)
2. To demystify the preconceived notion of black box.

OBJECTIVES AND CONTRIBUTIONS

OBJECTIVES

1. To assess the potential of LSTMs for bond yield forecasting (memory advantage)
2. To demystify the preconceived notion of black box.

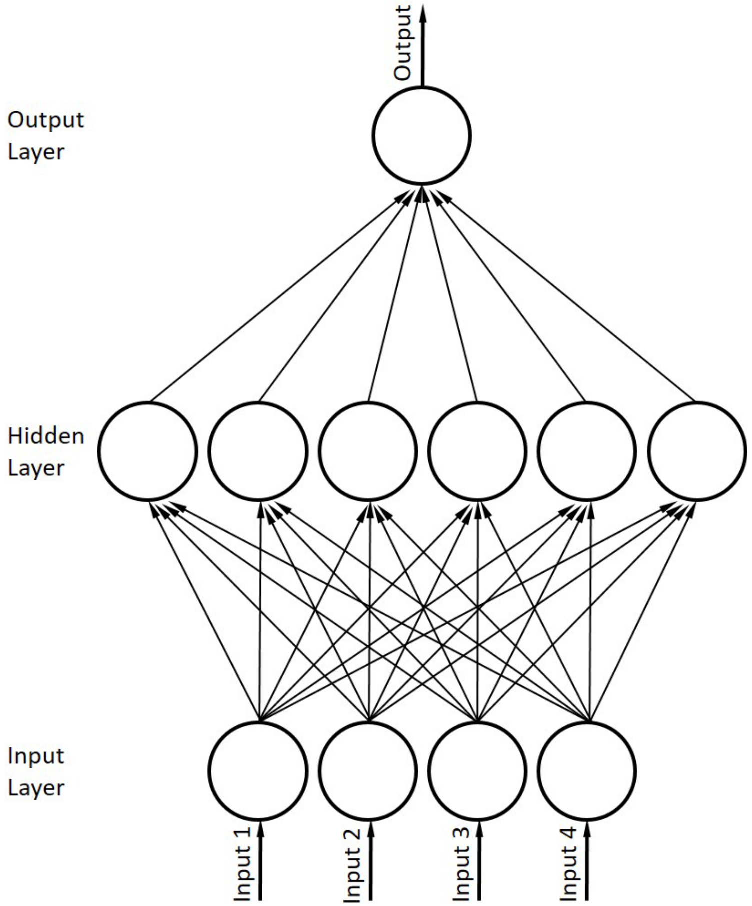
CONTRIBUTIONS

- [Obj. 1] First application of LSTMs to bond yield forecasting.
- [Obj. 2] First internal study of signals inside the LSTM memory cell (states & gates).
- [Obj. 2] Explain signals (states) with exogenous macroeconomic and market variables.
  - We develop a methodology here identified as LSTM-LagLasso.

Machine Learning Models Used

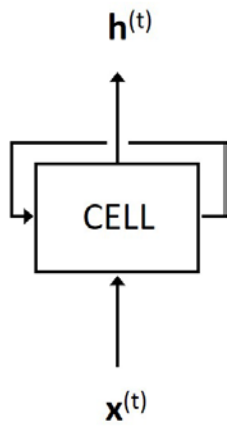
MULTILAYER PERCEPTRON (MLP)

- It is the standard feedforward neural network.
- This type of model has been used widely for forecasting, across several asset classes (incl. bonds).



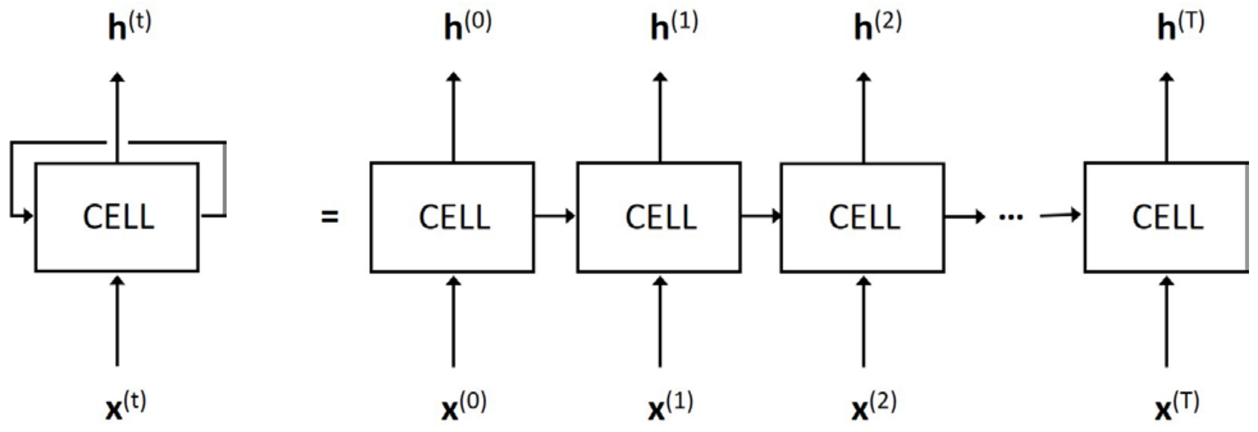
LONG SHORT-TERM MEMORY NETWORK (LSTM)

- Representation of Recurrent Neural Networks (RNN) unrolled in time.
  - x inputs; h outputs.



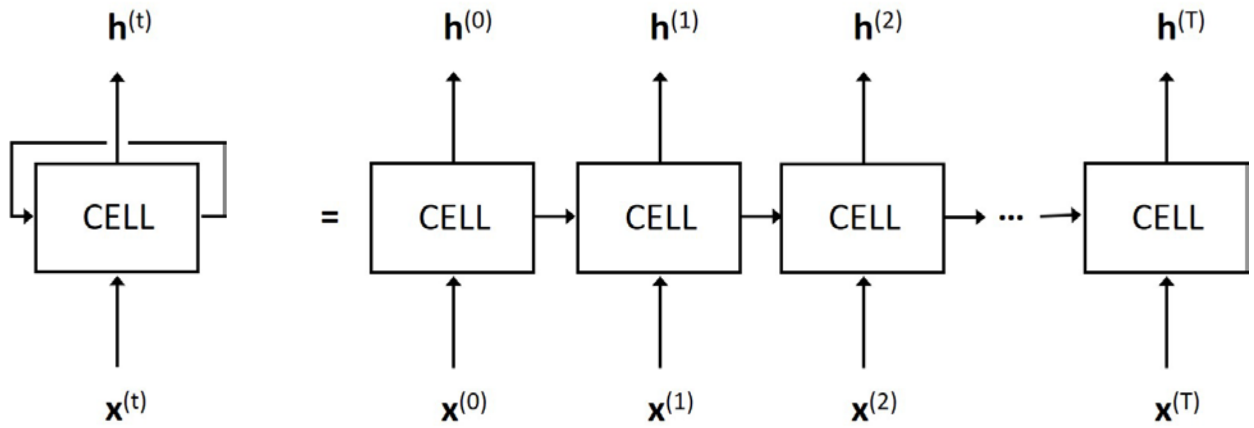
LONG SHORT-TERM MEMORY NETWORK (LSTM)

- Representation of Recurrent Neural Networks (RNN) unrolled in time.
  - x inputs; h outputs.

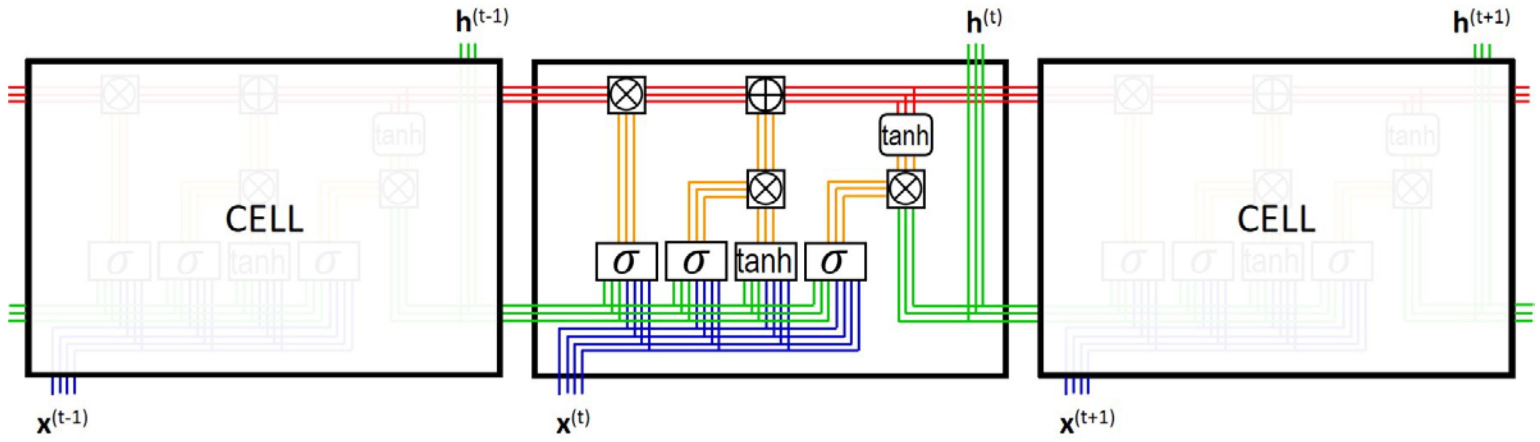


LONG SHORT-TERM MEMORY NETWORK (LSTM)

- Representation of Recurrent Neural Networks (RNN) unrolled in time.
  - x inputs; h outputs.



- The LSTM architecture was first introduced by Hochreiter and Schmidhuber (1997).



Data and Methodology

DATA

Target, features, dataset

- Target:
  - 10-year Euro government bond yield (January 1999 to April 2017).

DATA

Target, features, dataset

- Target:
  - 10-year Euro government bond yield (January 1999 to April 2017).
- Features (for MLPs):
  - Government bonds, corporate bonds, equities, currencies, commodities and volatility.
  - A vast range of economic indicators, from different geographic locations.
  - Calculated features: bond spreads, slope of the yield curve and simple technical analysis indicators.
  - Total number of features is 159 initial features.

DATA
Target, features, dataset

- Target:
  - 10-year Euro government bond yield (January 1999 to April 2017).
- Features (for MLPs):
  - Government bonds, corporate bonds, equities, currencies, commodities and volatility.
  - A vast range of economic indicators, from different geographic locations.
  - Calculated features: bond spreads, slope of the yield curve and simple technical analysis indicators.
  - Total number of features is 159 initial features.
- Feature selection (for MLPs):
  - selection of the most relevant features was carried out using Lasso regression.

THE MEMORY ADVANTAGE OF LSTM NETWORKS

Methodology

- We apply a univariate LSTM.
  - To assess LSTMs in the simplest form.

THE MEMORY ADVANTAGE OF LSTM NETWORKS

Methodology

- We apply a univariate LSTM.
  - To assess LSTMs in the simplest form.
- <u>**Direct comparison MLP vs. LSTM**</u>
  - Models (input sequence: 6 days).
  - Forecasting horizons: next day, +5, +10, +15, +20 days.

Model	Short name	Description
Direct comparison MLP vs. LSTM		
MLP	NN TgtOnly	MLP with target data only
LSTM	LSTM06	LSTM using input sequence of 6 time steps
LSTMs with different input sequences		
MLP	NN RelFeat	MLP with relevant features
	NN TgtOnly	MLP with target data only
LSTM	LSTM06	LSTM using input sequence of 6 time steps
	LSTM21	LSTM using input sequence of 21 time steps
	LSTM61	LSTM using input sequence of 61 time steps

Input sequence: number of past days passed to the model for forecasting.

THE MEMORY ADVANTAGE OF LSTM NETWORKS

Methodology

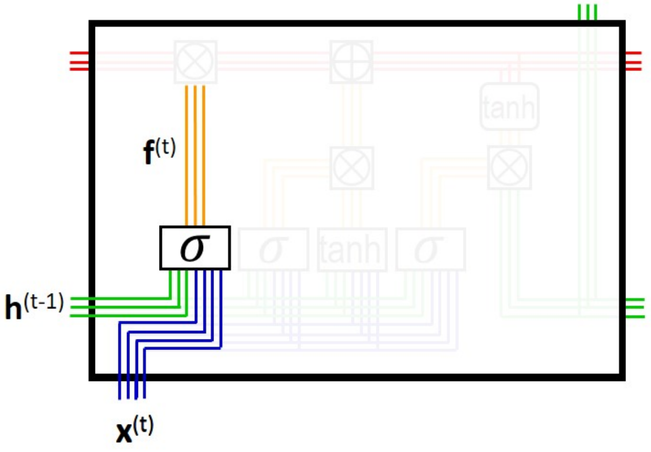
- We apply a univariate LSTM.
  - To assess LSTMs in the simplest form.
- **Direct comparison MLP vs. LSTM**
  - Models (input sequence: 6 days).
  - Forecasting horizons: next day, +5, +10, +15, +20 days.
- **LSTMs with different input sequences**
  - Models (input sequences: 6, 21, 61 days).
  - Forecasting horizons: next day, +5, +10, +15, +20 days.

Model	Short name	Description
Direct comparison MLP vs. LSTM		
MLP	NN TgtOnly	MLP with target data only
LSTM	LSTM06	LSTM using input sequence of 6 time steps
LSTMs with different input sequences		
MLP	NN RelFeat	MLP with relevant features
	NN TgtOnly	MLP with target data only
LSTM	LSTM06	LSTM using input sequence of 6 time steps
	LSTM21	LSTM using input sequence of 21 time steps
	LSTM61	LSTM using input sequence of 61 time steps

Input sequence: number of past days passed to the model for forecasting.

SIGNALS ANALYSIS

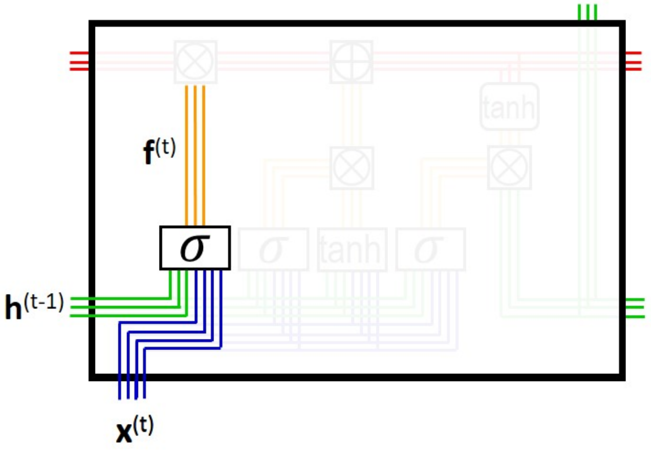
Gates and States - signals extracted at each location, time step, and hidden unit



$$f^{(t)} = \sigma\left(\boldsymbol{W}^{fx}\,\boldsymbol{x}^{(t)} + \boldsymbol{W}^{fh}\,\boldsymbol{h}^{(t-1)} + b_f\right)$$

SIGNALS ANALYSIS

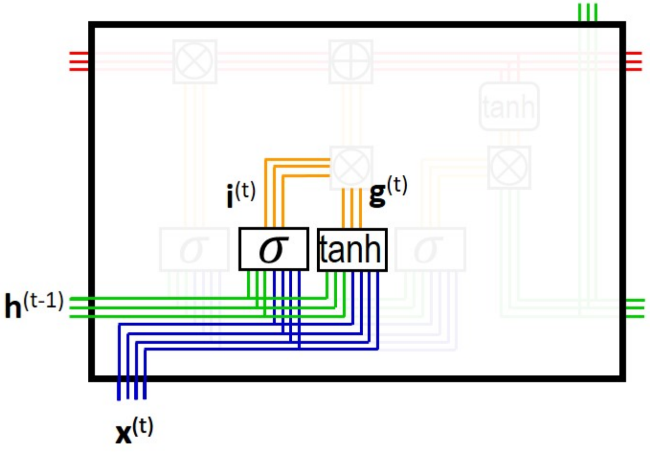
Gates and States - signals extracted at each location, time step, and hidden unit



$$f^{(t)} = \sigma\left(W^{fx}\,x^{(t)} + W^{fh}\,h^{(t-1)} + b_f\right)$$

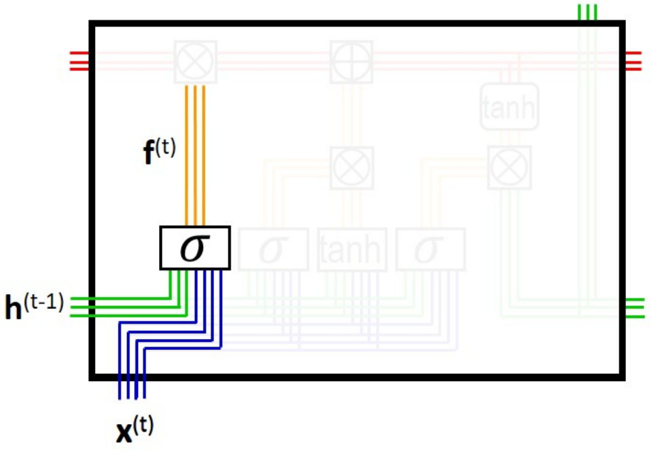
$$i^{(t)} = \sigma\left(W^{ix}\,x^{(t)} + W^{ih}\,h^{(t-1)} + b_i\right)$$

$$g^{(t)} = tanh\left(W^{gx}\,x^{(t)} + W^{gh}\,h^{(t-1)} + b_g\right)$$



SIGNALS ANALYSIS

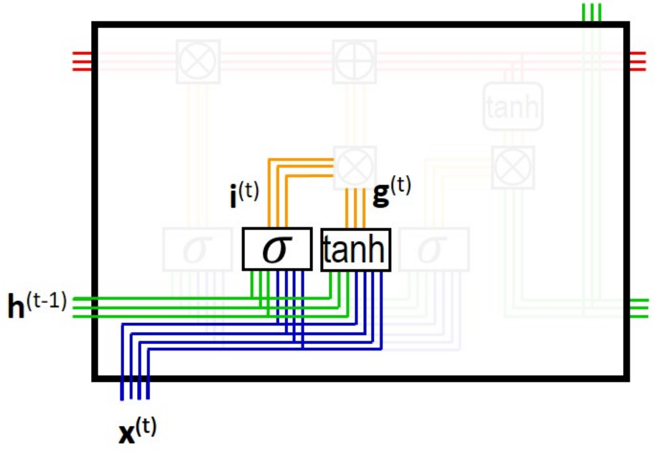
Gates and States - signals extracted at each location, time step, and hidden unit

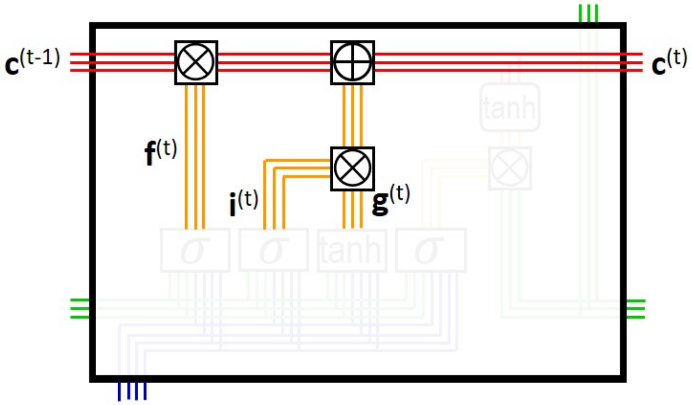


$$f^{(t)} = \sigma\left(W^{fx}\, x^{(t)} + W^{fh}\, h^{(t-1)} + b_f\right)$$

$$i^{(t)} = \sigma\left(W^{ix}\, x^{(t)} + W^{ih}\, h^{(t-1)} + b_i\right)$$

$$g^{(t)} = tanh\left(W^{gx}\, x^{(t)} + W^{gh}\, h^{(t-1)} + b_g\right)$$

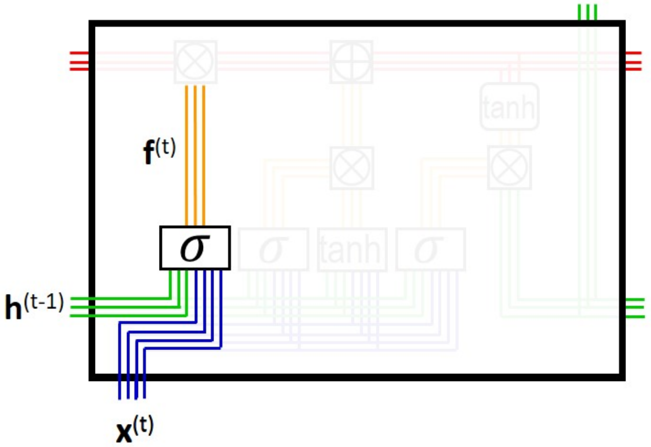




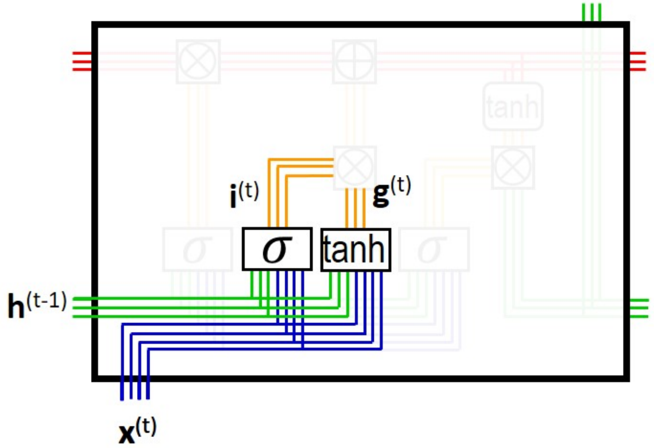
$$c^{(t)} = f^{(t)} \otimes c^{(t-1)} + i^{(t)} \otimes g^{(t)}$$

SIGNALS ANALYSIS

Gates and States - signals extracted at each location, time step, and hidden unit

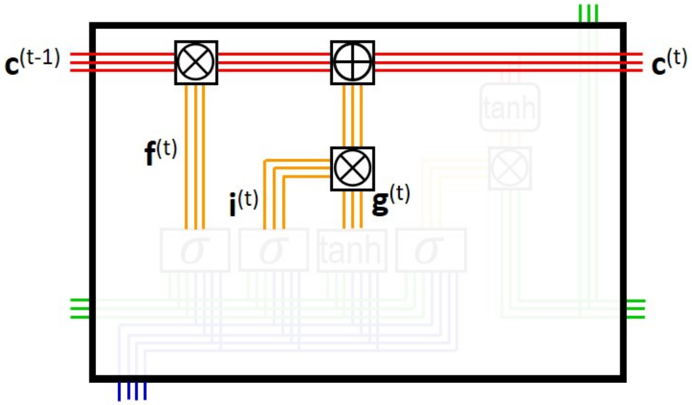


$$f^{(t)} = \sigma\left(\boldsymbol{W}^{fx}\,\boldsymbol{x}^{(t)} + \boldsymbol{W}^{fh}\,\boldsymbol{h}^{(t-1)} + b_f\right)$$

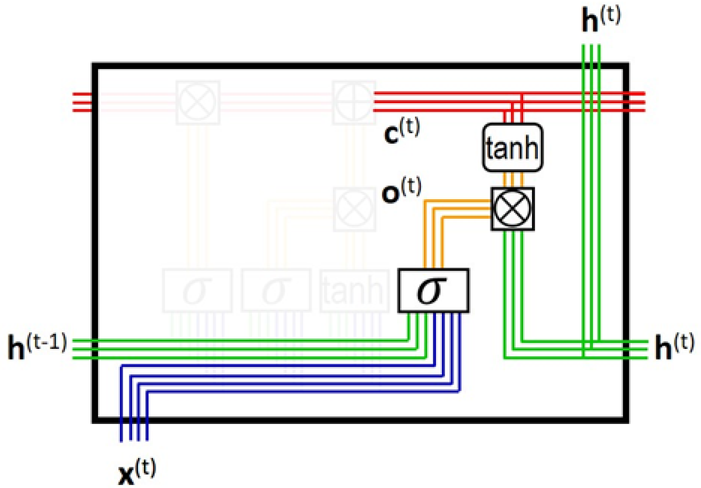


$$i^{(t)} = \sigma\left(\boldsymbol{W}^{ix}\,\boldsymbol{x}^{(t)} + \boldsymbol{W}^{ih}\,\boldsymbol{h}^{(t-1)} + b_i\right)$$

$$g^{(t)} = tanh\left(\boldsymbol{W}^{gx}\,\boldsymbol{x}^{(t)} + \boldsymbol{W}^{gh}\,\boldsymbol{h}^{(t-1)} + b_g\right)$$



$$c^{(t)} = f^{(t)} \otimes c^{(t-1)} + i^{(t)} \otimes g^{(t)}$$

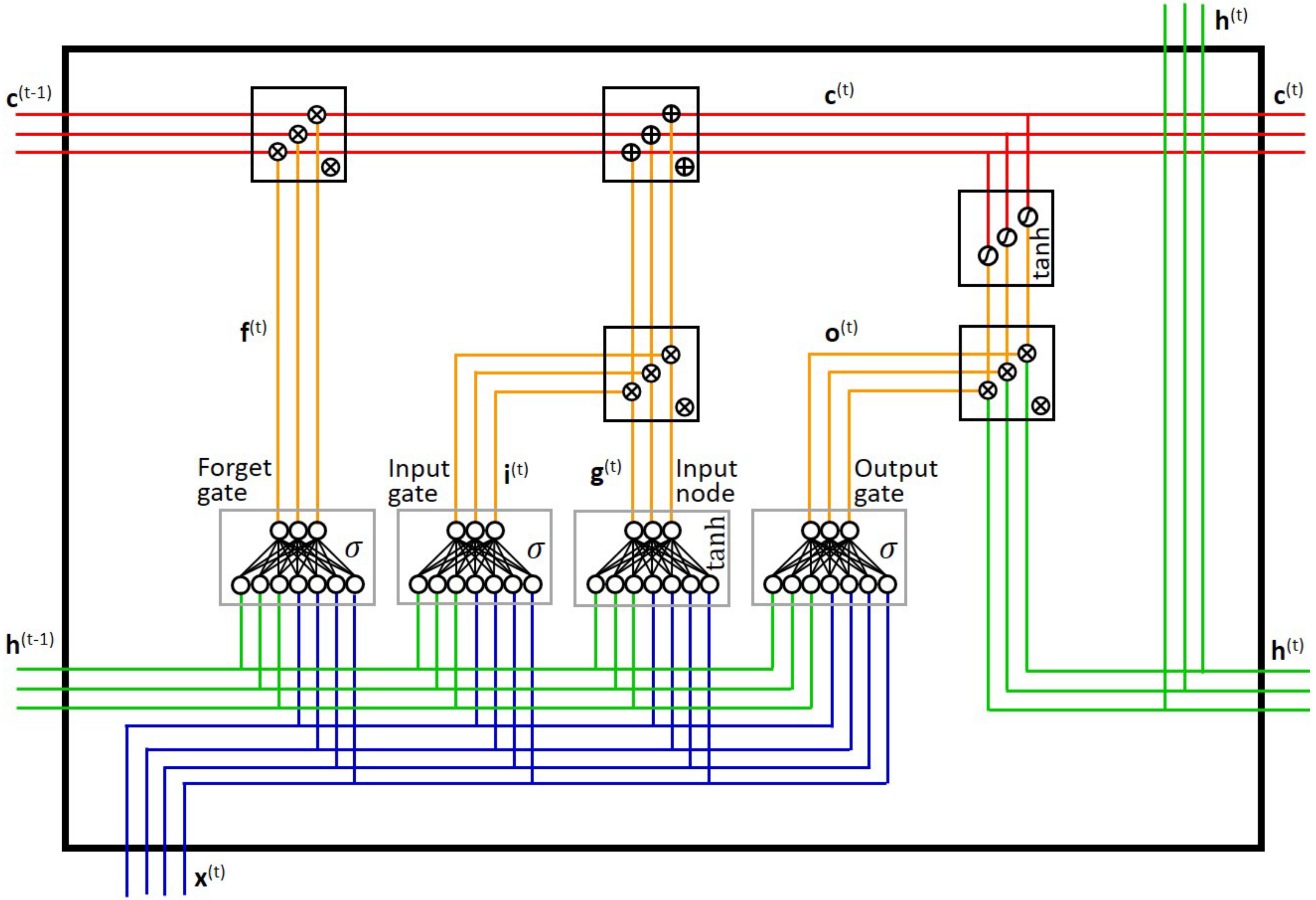


$$o^{(t)} = \sigma\left(\boldsymbol{W}^{ox}\,\boldsymbol{x}^{(t)} + \boldsymbol{W}^{oh}\,\boldsymbol{h}^{(t-1)} + b_o\right)$$

$$h^{(t)} = o^{(t)} \otimes tanh(c^{(t)})$$

SIGNALS ANALYSIS

Gates and States - signals extracted at each location, time step, and hidden unit



UNIVERSITY OF Southampton

SIGNALS ANALYSIS

Methodology – Extracting the signals

LSTM architectures used:

- Sequence-to-sequence
- Univariate and multivariate
  - 3 feature sets
- Hidden units: 3 (for interpretability)

LSTM architecture		
Features	set 1	10-year yield
	set 2	10-year, momentum indicator
	set 3	10-year, 5-year, 30-year yield
Target		10-year yield
Forecasting horizon		next day + 5 days
Moving window size		3000 days
Hidden units		3
Sequence in		6 days
Sequence out		6 days

LSTM-LAGLASSO

Methodology – Explaining the signals

- Based on Lasso [Tibshirani, 1996] and Kalman-LagLasso [Mahler, 2009].

LSTM-LAGLASSO

Methodology – Explaining the signals

- Based on Lasso [Tibshirani, 1996] and Kalman-LagLasso [Mahler, 2009].
- Modifications LSTM-LagLasso in relation to Kalman-LagLasso:
  - Model used.
  - Target variable.
  - Methodology to determine the relevant features & lags
    - <u>no denoising vs denoising</u> of variables and
    - all relevant lags considered vs. one per feature.

LSTM-LAGLASSO

Methodology – Explaining the signals

- Based on Lasso [Tibshirani, 1996] and Kalman-LagLasso [Mahler, 2009].
- Modifications LSTM-LagLasso in relation to Kalman-LagLasso:
  - Model used.
  - Target variable.
  - Methodology to determine the relevant features & lags
    - <u>no denoising vs denoising</u> of variables and
    - all relevant lags considered vs. one per feature.
- Feature set same as for the MLP study.

LSTM-LAGLASSO

Methodology – Explaining the signals

- Based on Lasso [Tibshirani, 1996] and Kalman-LagLasso [Mahler, 2009].
- Modifications LSTM-LagLasso in relation to Kalman-LagLasso:
  - Model used.
  - Target variable.
  - Methodology to determine the relevant features & lags
    - no denoising vs denoising of variables and
    - all relevant lags considered vs. one per feature.
- Feature set same as for the MLP study.
- Excluded 10-year yield from feature set.

Results and Discussion

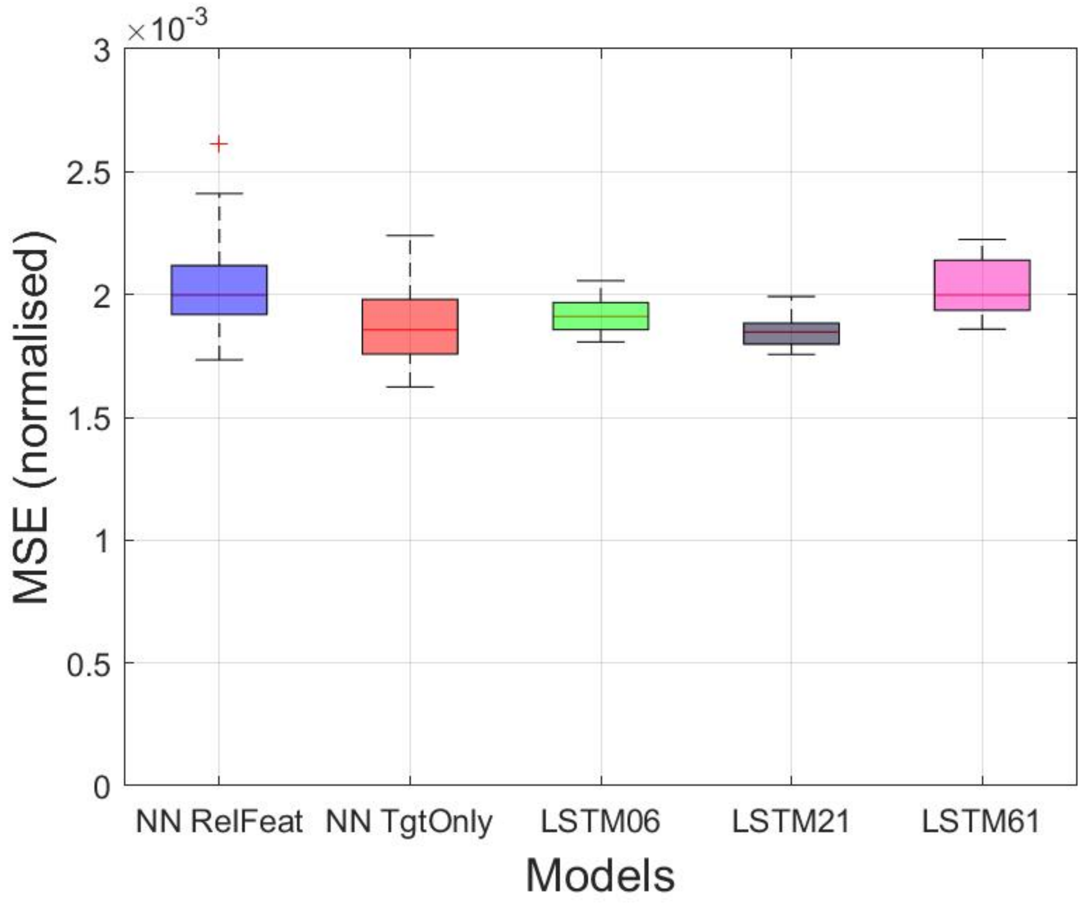
LSTMs vs. MLPs

THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (1/3)

Forecasting horizon: next day

- Similar results across models.
- Lower standard deviation for LSTMs.

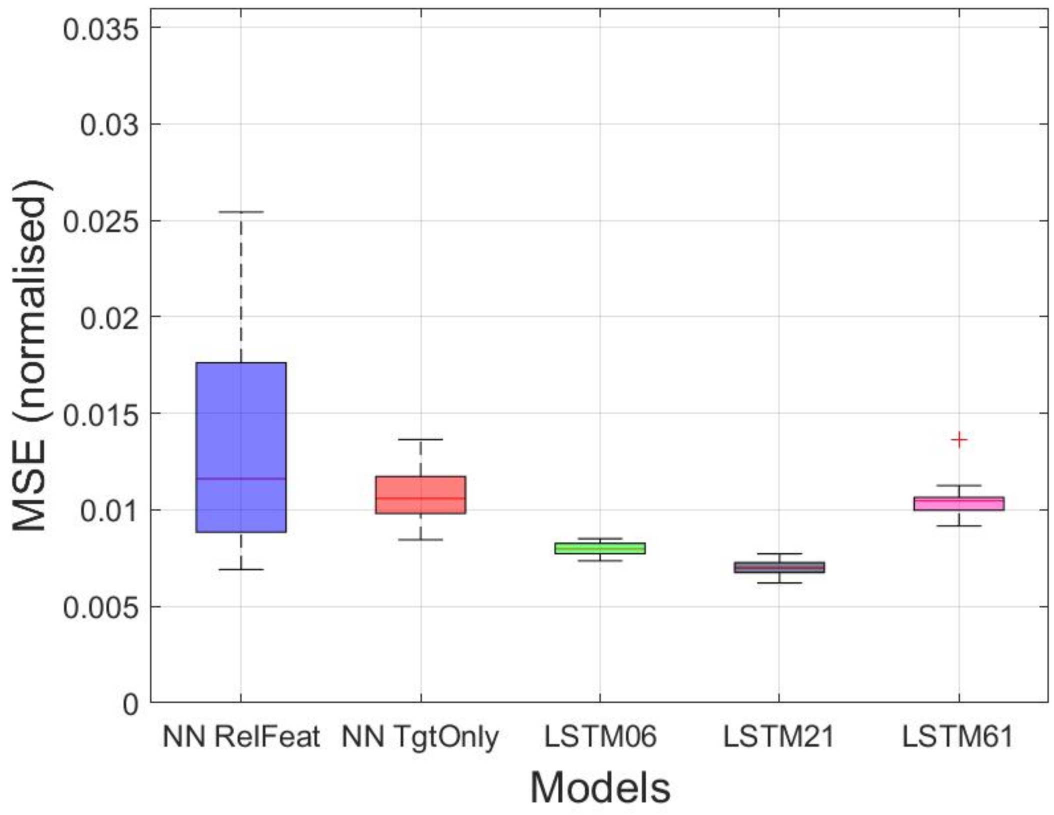


THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (2/3)

Forecasting horizon: next day + 5

- The benefits of LSTMs become more evident.
- For 6 and 21 time steps LSTM significantly outperform.

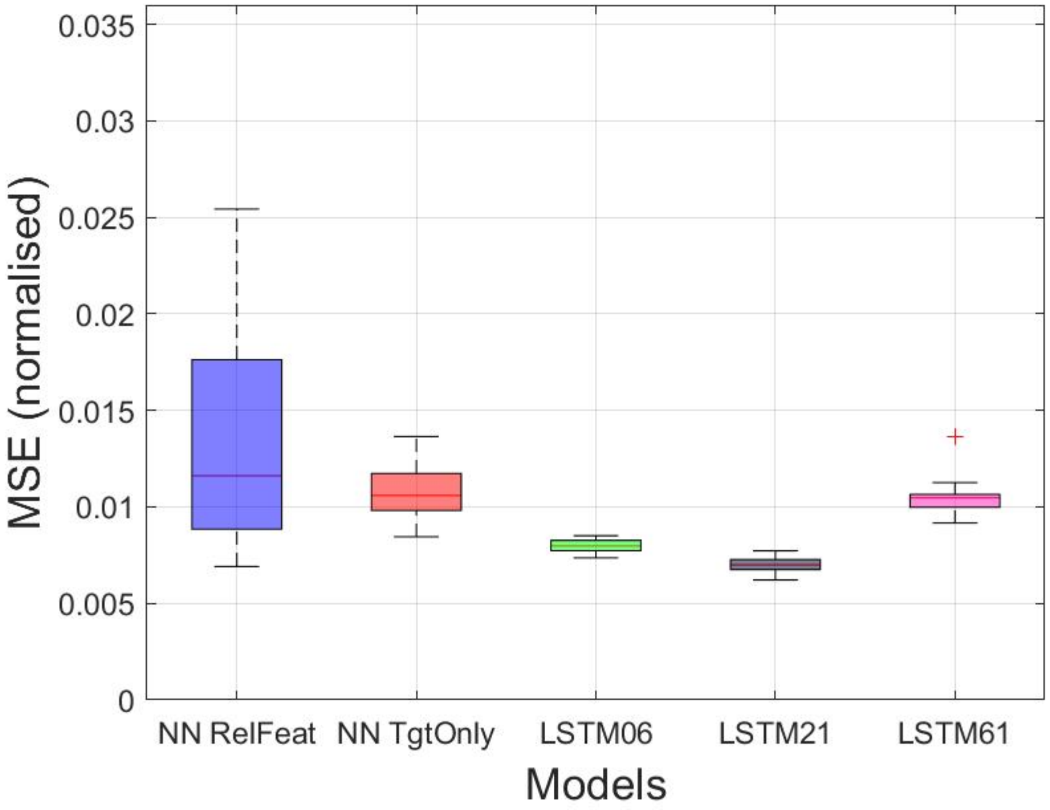


THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (2/3)

Forecasting horizon: next day + 5

- The benefits of LSTMs become more evident.
- For 6 and 21 time steps LSTM significantly outperform.
- For 61 time steps, it appears that such a long input sequence is not yet "necessary".



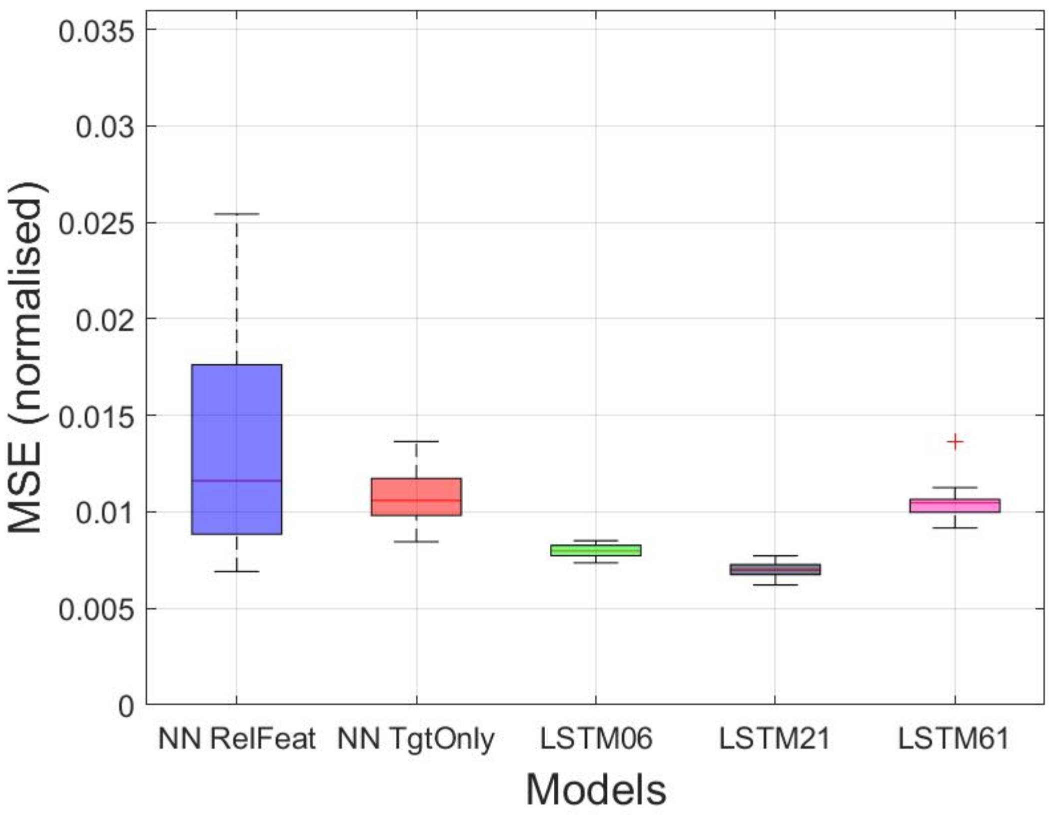
THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (2/3)

Forecasting horizon: next day + 5

- The benefits of LSTMs become more evident.
- For 6 and 21 time steps LSTM significantly outperform.
- For 61 time steps, it appears that such a long input sequence is not yet “necessary”.

Horizon selected for further study

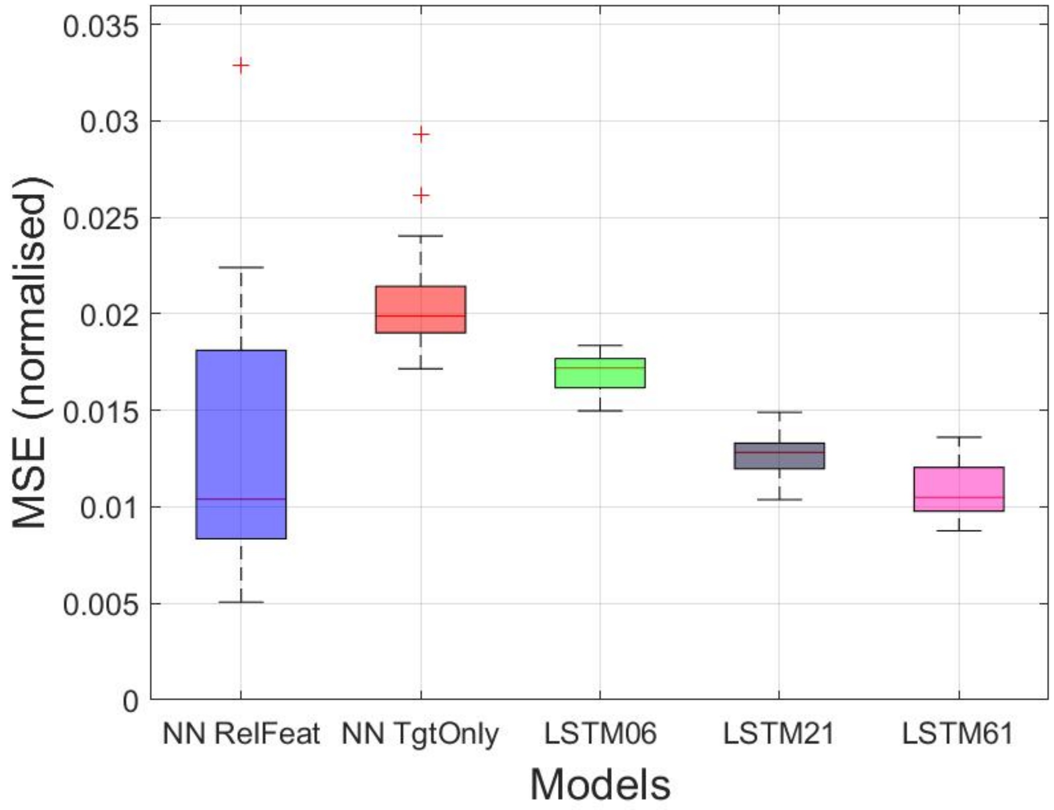


THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (3/3)

Forecasting horizon: next day + 10 & 15

- LSTMs tend to perform better.
- LSTMs produce lower standard deviations.

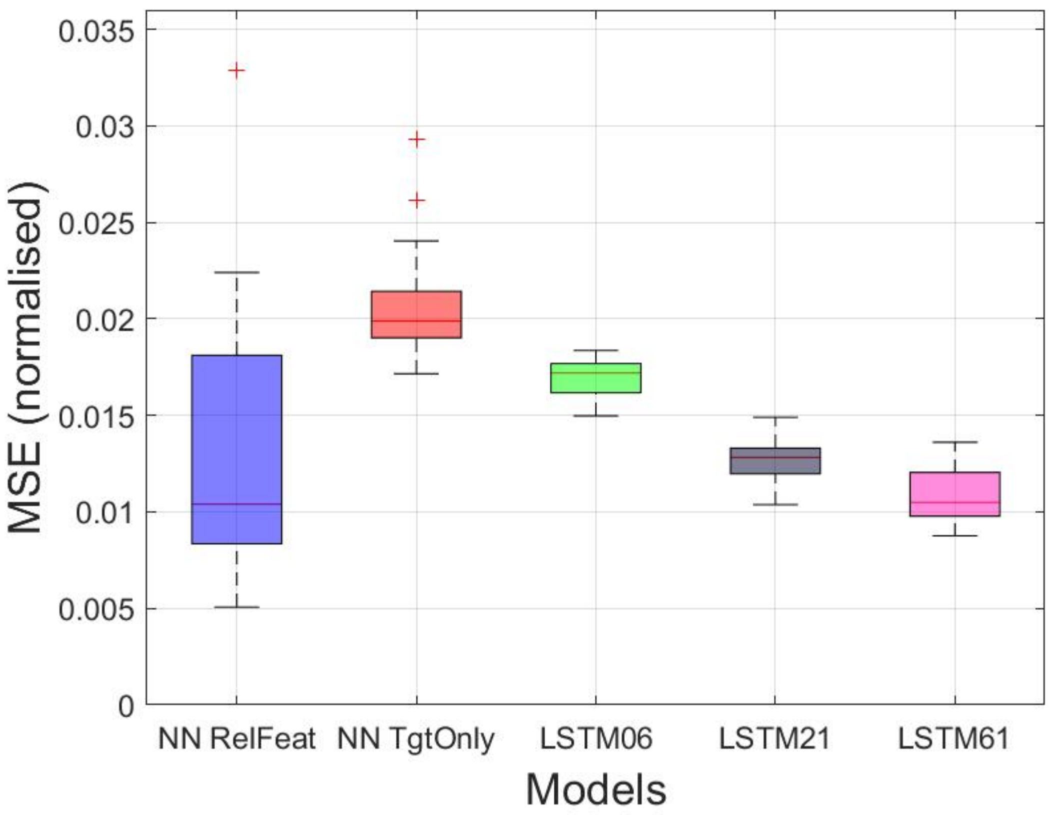


THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (3/3)

Forecasting horizon: next day + 10 & 15

- LSTMs tend to perform better.
- LSTMs produce lower standard deviations.
- LSTM61 reach similar median error levels to the MLP with the most relevant features.

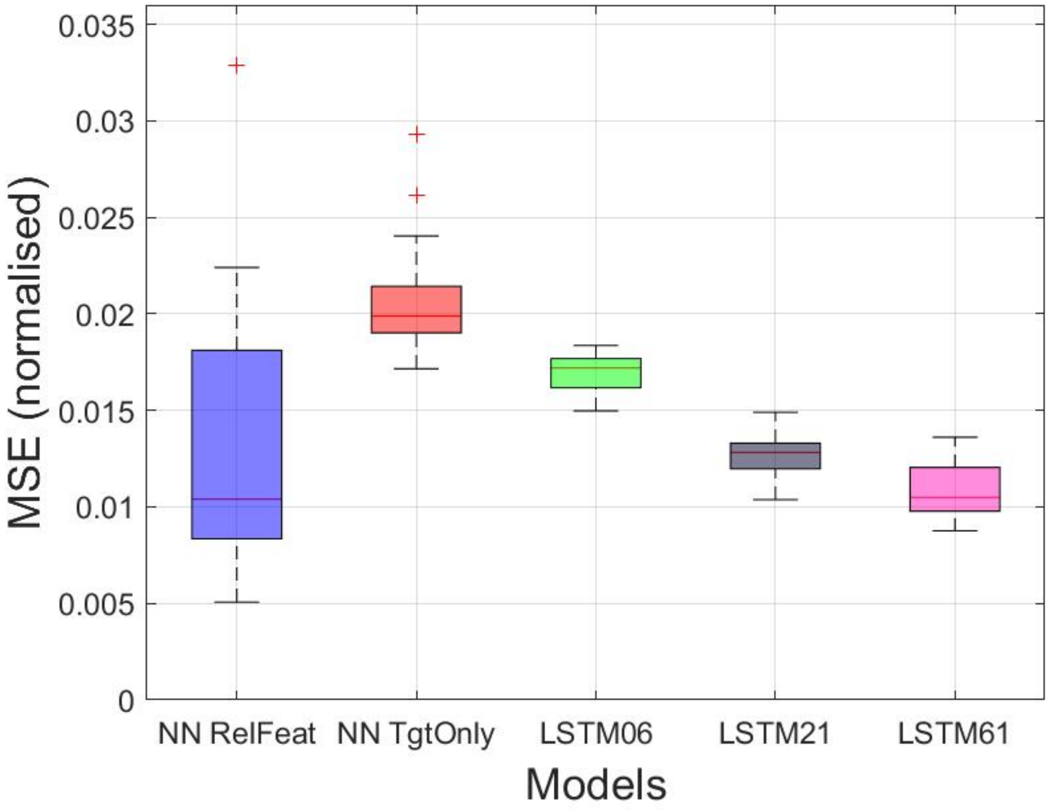


THE MEMORY ADVANTAGE OF LSTM NETWORKS

Results - LSTMs with different input sequences (3/3)

Forecasting horizon: next day + 10 & 15

- LSTMs tend to perform better.
- LSTMs produce lower standard deviations.
- LSTM61 reach similar median error levels to the MLP with the most relevant features.
- LSTMs appear capable of compensating the lack of additional information from markets with additional memory via longer input sequences.



Results and Discussion

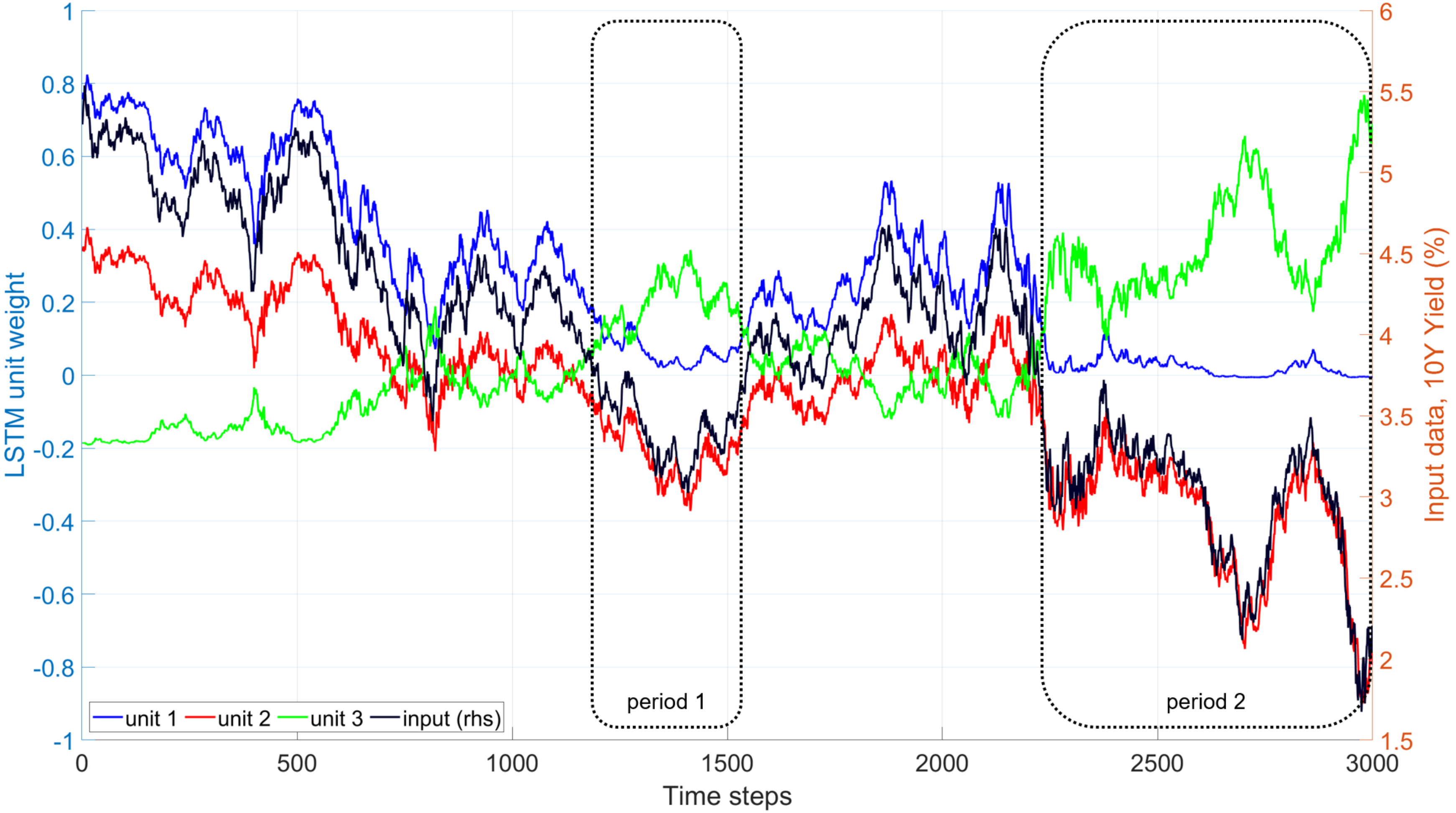
Signals Analysis

SIGNALS ANALYSIS

Results – Feature set 1 (10Y yield), Hidden State

During period 1 & 2:

- Unit 1 ≈ inactive.
- Units 2 & 3 take over.



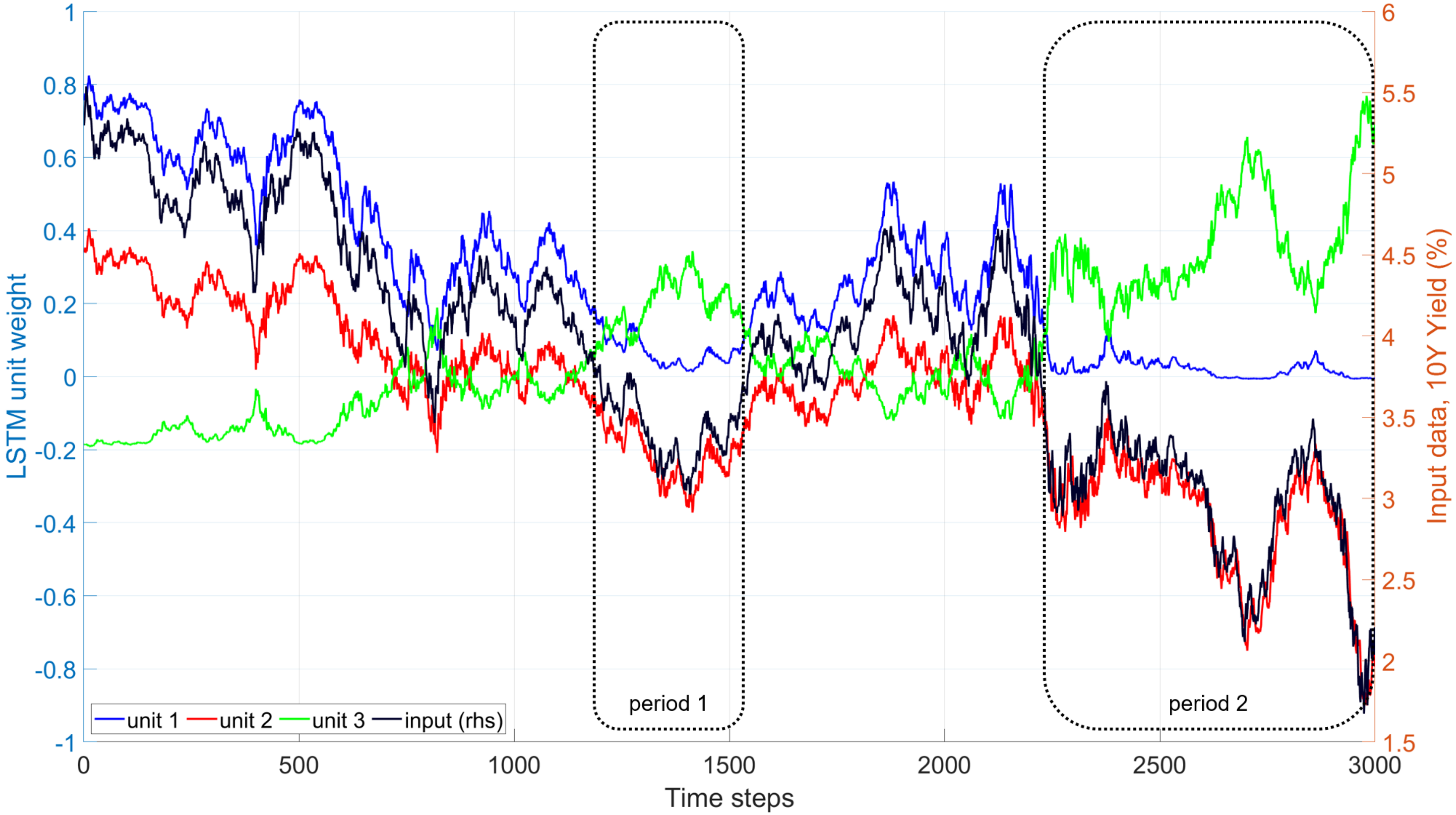
“units”=LSTM hidden units

SIGNALS ANALYSIS

Results – Feature set 1 (10Y yield), Hidden State

During period 1 & 2:

- Unit 1 ≈ inactive.
- Units 2 & 3 take over.
- 10Y yield assumes downward extreme values
  - below 3.6-3.7%.



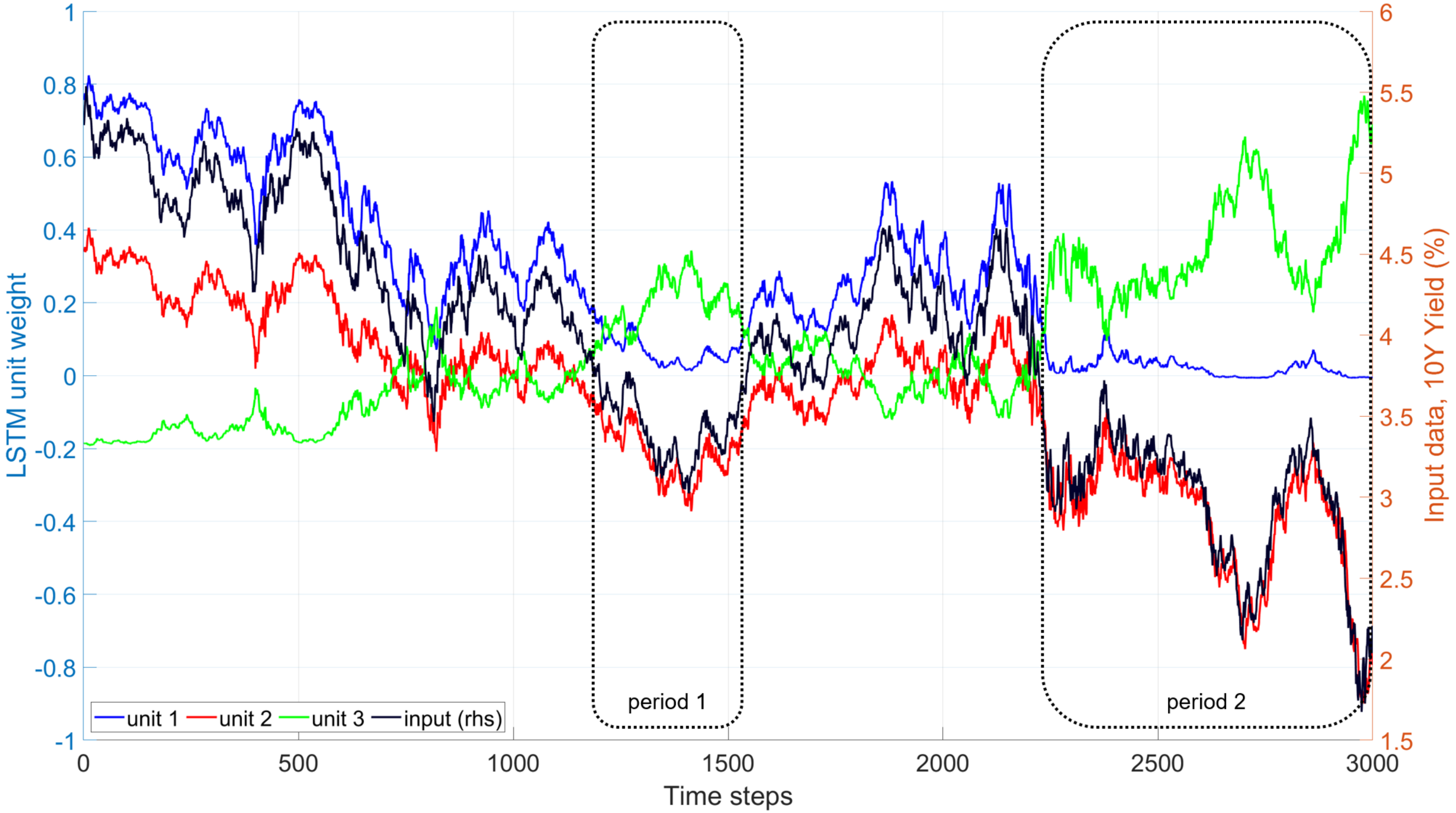
“units”=LSTM hidden units

SIGNALS ANALYSIS

Results – Feature set 1 (10Y yield), Hidden State

During period 1 & 2:

- Unit 1 ≈ inactive.
- Units 2 & 3 take over.
- 10Y yield assumes downward extreme values
  - below 3.6-3.7%.
- “Specialisation” of hidden units.



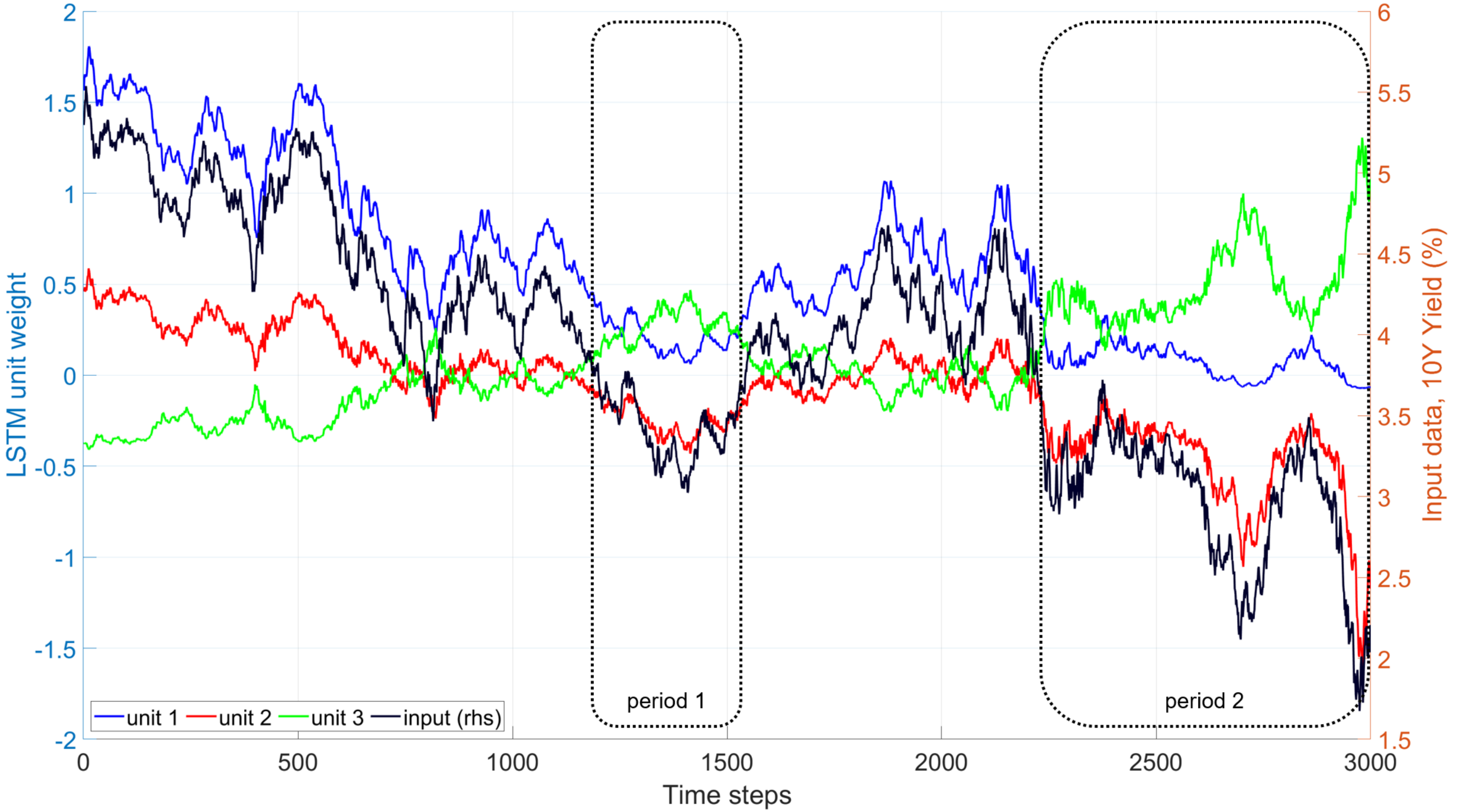
“units”=LSTM hidden units

SIGNALS ANALYSIS

Results – Feature set 1 (10Y yield), Cell State

During period 1 & 2:

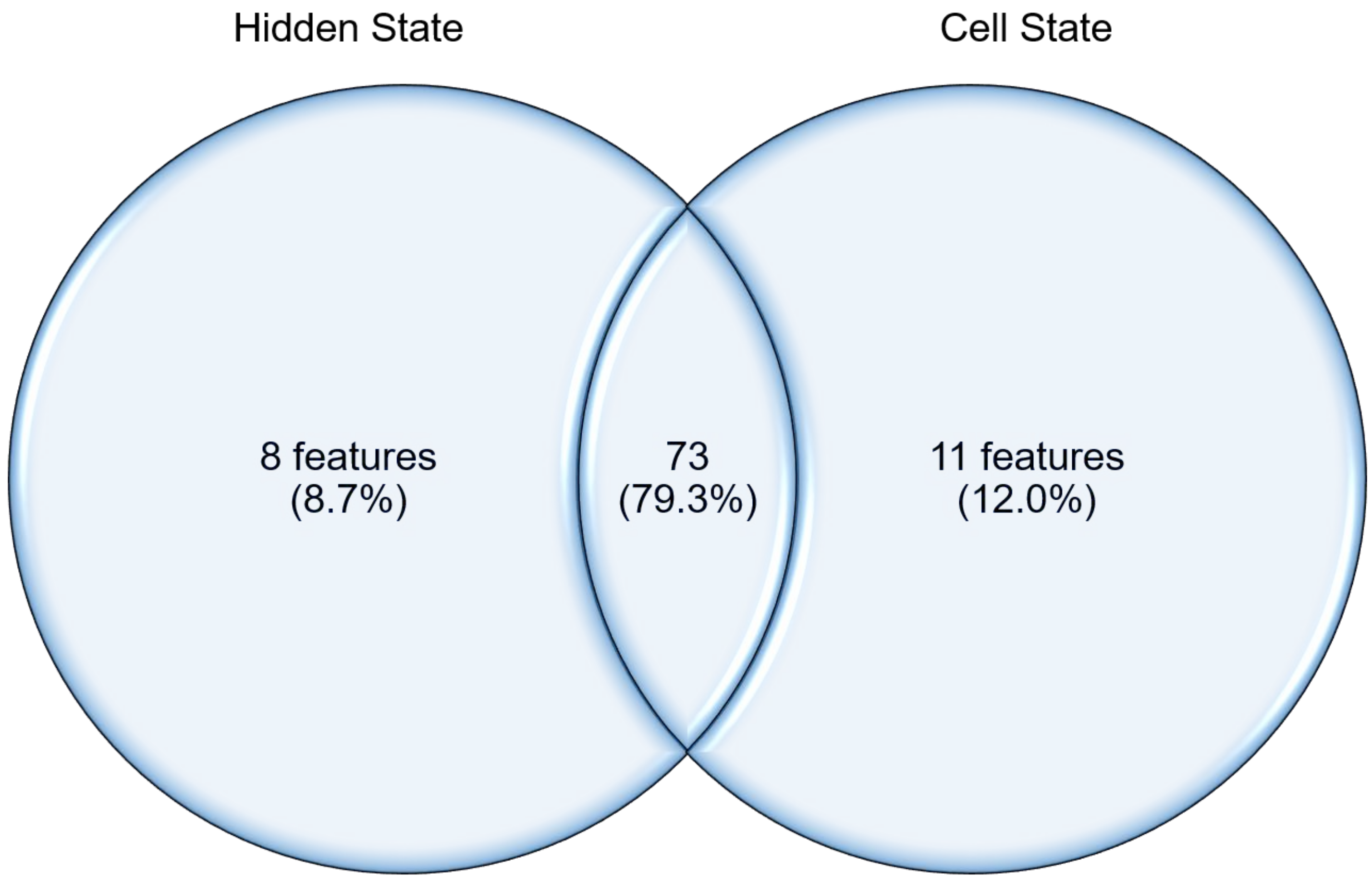
- The Cell State presents a similar overall pattern.



“units”=LSTM hidden units

LSTM-LAGLASSO METHOD

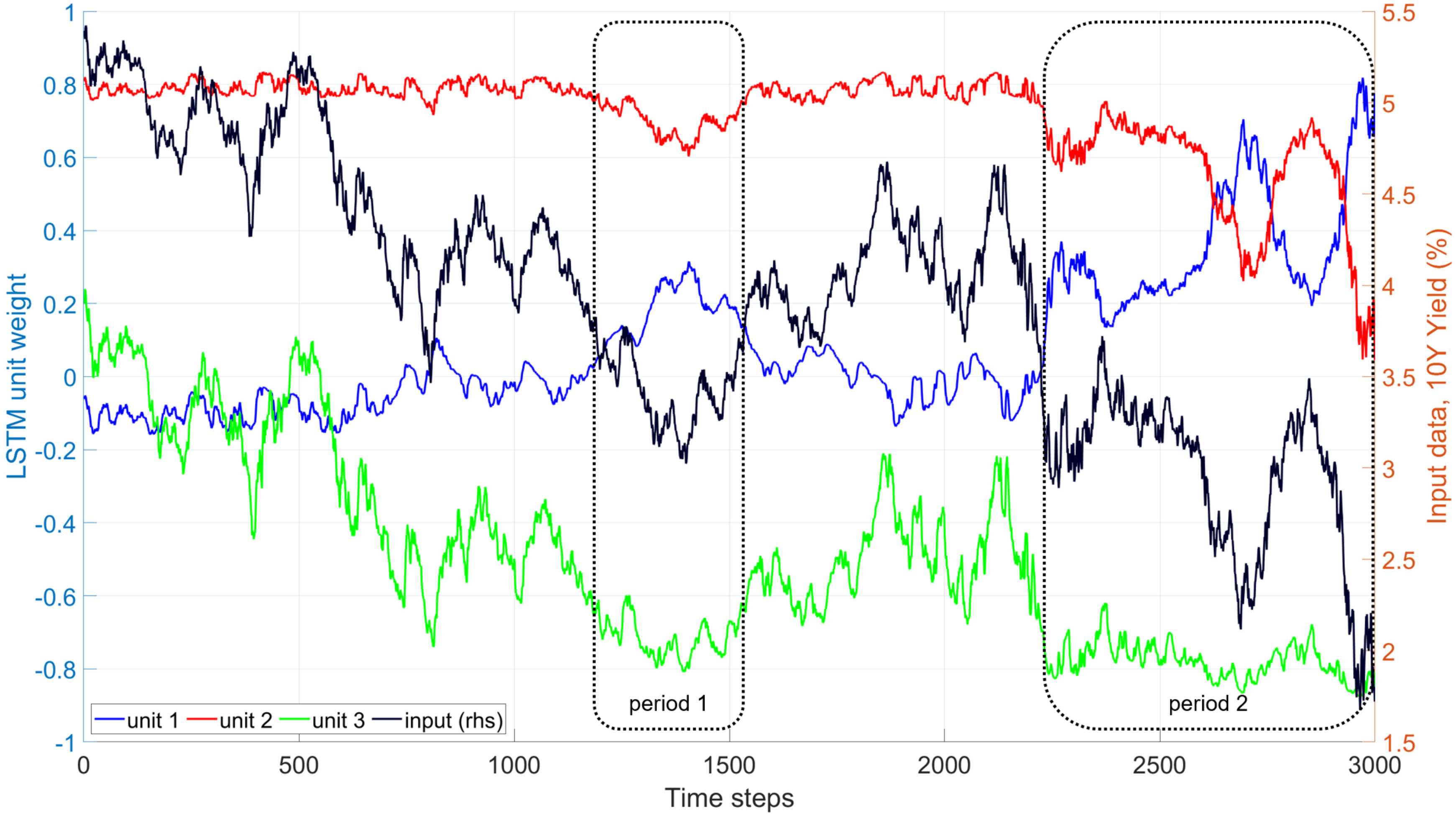
Results – Relevant features for hidden and cell states



SIGNALS ANALYSIS

Results – Feature set 2 (10Y yield + momentum indicator), Hidden State

- For feature set 2 the same periods can be identified.

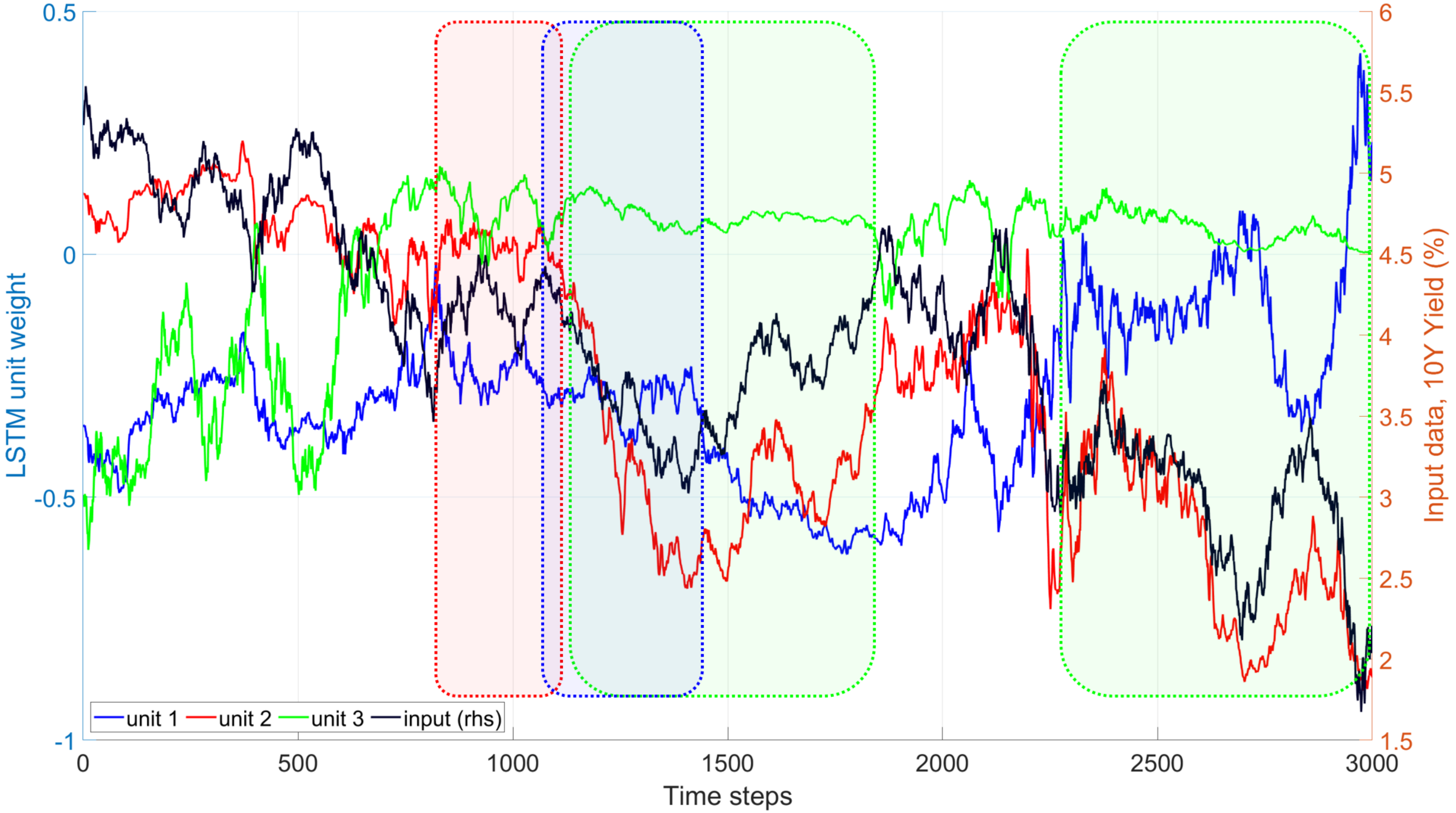


"units"=LSTM hidden units

SIGNALS ANALYSIS

Results – Feature set 3 (10Y yield + 5Y, 30Y yield), Hidden State

- For feature set 3 the activation / deactivation of units through time is still visible.
- Some examples of those periods are highlighted.



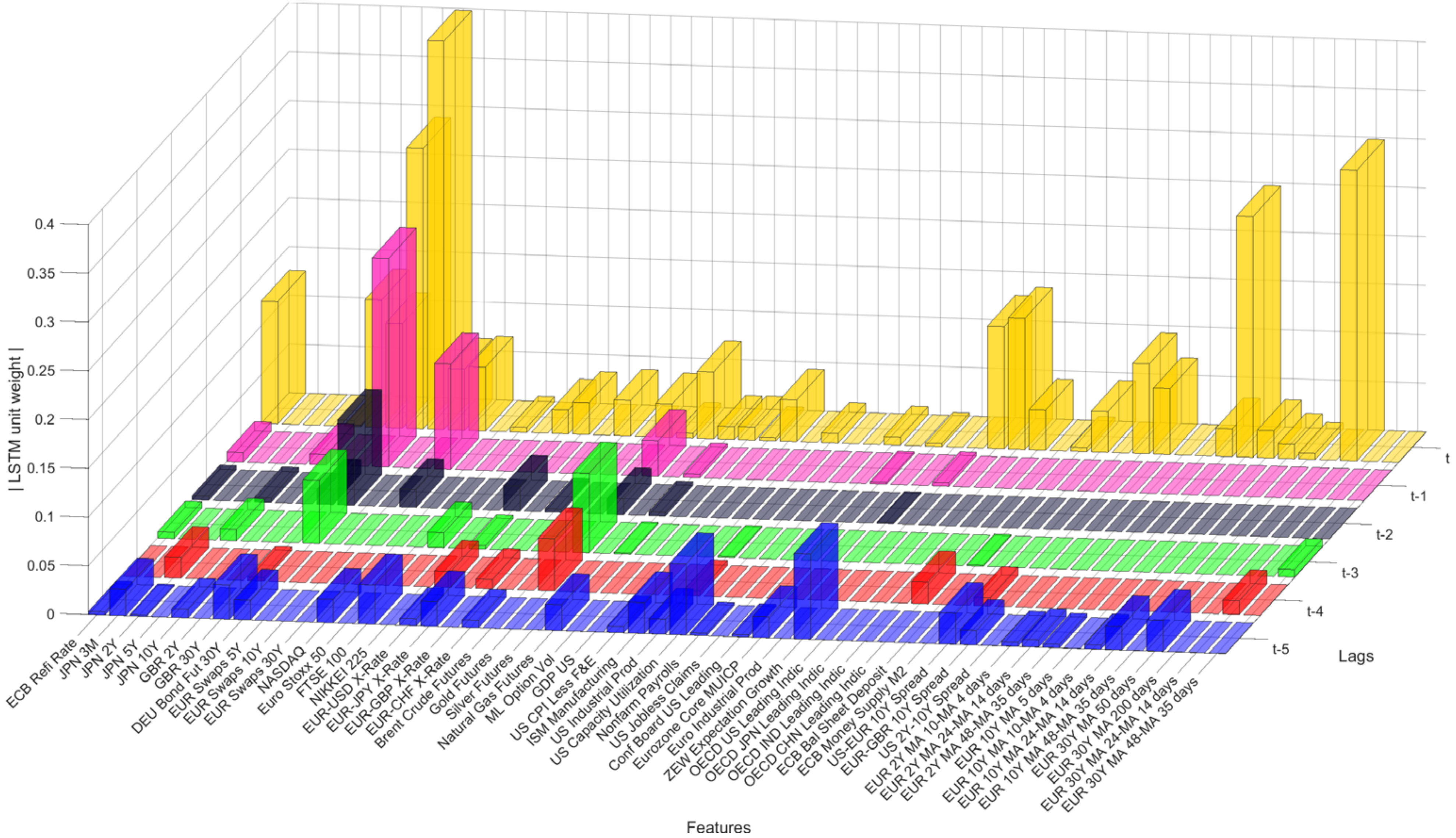
“units”=LSTM hidden units

Results and Discussion

LSTM-LagLasso Method

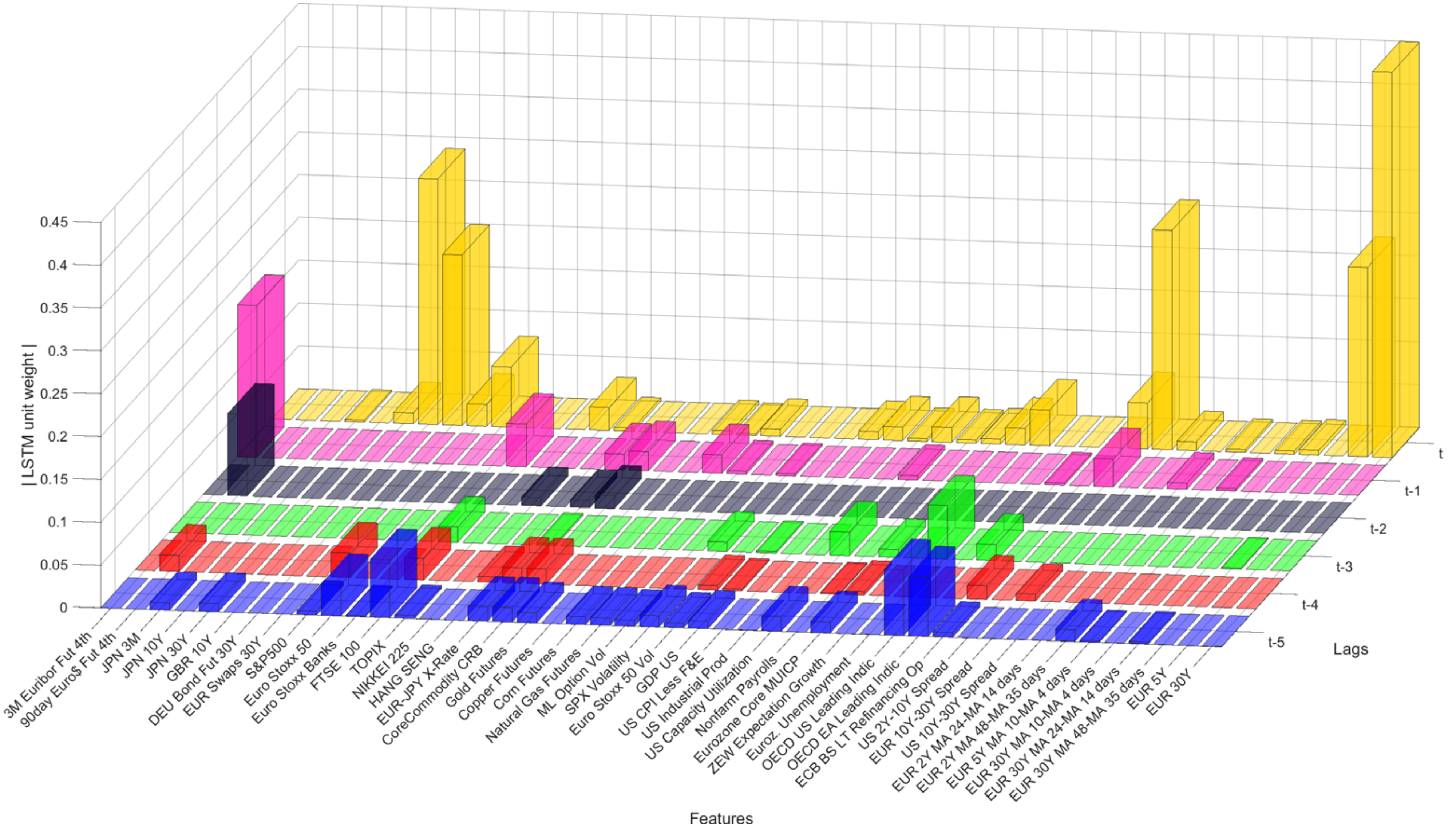
LSTM-LAGLASSO METHOD

Results – Relevant features for the hidden state, unit 1



LSTM-LAGLASSO METHOD

Results – Relevant features for the hidden state, unit 2

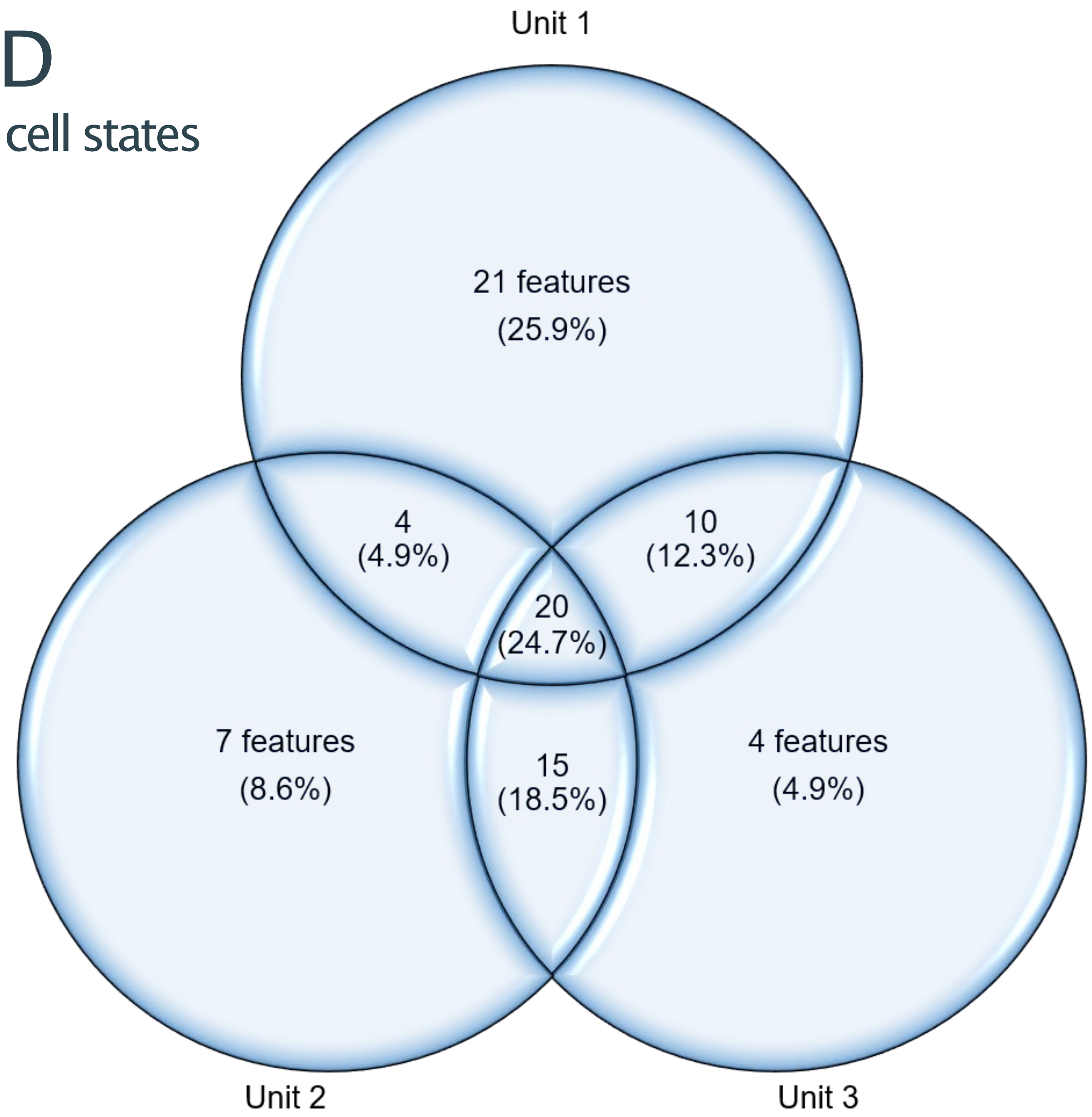


LSTM-LAGLASSO METHOD

Results – Relevant features for hidden and cell states

More diversity among units:

- Not all relevant features are common to all units, additionally pointing to unit differentiation.
  - Approx. 25% are common.



LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

UNIVERSITY OF Southampton

LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

University of Southampton

LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

UNIVERSITY OF Southampton

LSTM-LAGLASSO METHOD

Results – Relevant features with | Weight | > 0.15

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Hidden state			
ECB Refi Rate	0.150		
3M Euribor Fut 4th		0.177	0.248
GBR 30Y	0.161		0.053
DEU Bond Fut 30Y	0.674	0.287	0.410
EUR Swaps 5Y	0.398		
EUR Swaps 30Y	0.192	0.199	0.315
FTSE 100	0.056	0.160	0.107
Gold Futures	0.193	0.012	0.004
US 2Y-10Y Spread	0.068	0.087	0.198
EUR 10Y-30Y Spread		0.264	0.214
EUR 10Y MA 5 days	0.247		0.082
EUR 30Y MA 200 days	0.298		
EUR 5Y		0.221	0.202
EUR 30Y		0.450	0.168

Feature name	Weight		
	Unit 1	Unit 2	Unit 3
Cell state			
3M Euribor Fut 4th			0.212
90day Euro$ Fut 5th		0.203	
GBR 30Y	0.151	0.038	
DEU Bond Fut 30Y	0.723	0.921	0.533
EUR Swaps 5Y	0.258		
EUR Swaps 30Y	0.220		0.186
Gold Futures	0.168	0.109	0.051
EUR 10Y-30Y Spread		0.271	0.270
EUR 10Y MA 5 days	0.525		
EUR 30Y MA 200 days	0.273		
EUR 5Y		0.003	0.172
EUR 30Y			0.240

The listed features have an absolute weight greater than or equal to 0.15 in at least one of the LSTM units. The cells in the table for which the weight is equal to zero are left empty.

LSTM-LAGLASSO METHOD

Results – Relevant features

Conventional features:

- Central Bank reference rates
- Inflation
- Economic activity
  - ISM Manufacturing, Industrial Production, Capacity Utilisation, ZEW Expectation Growth
- Labour markets

LSTM-LAGLASSO METHOD

Results – Relevant features

Conventional features:

- Central Bank reference rates
- Inflation
- Economic activity
  - ISM Manufacturing, Industrial Production, Capacity Utilisation, ZEW Expectation Growth
- Labour markets

Non-conventional features:

- Futures German 30Y bond
- Futures 3M Euribor & 90d Euro$
- Adjacent maturities
- Intra and inter-curve spreads
- OECD Leading Indicators
- ECB Balance Sheet indicators

Southampton (University of)

LSTM-LAGLASSO METHOD

Results – Relevant features

Conventional features:

- Central Bank reference rates
- Inflation
- Economic activity
  - ISM Manufacturing, Industrial Production, Capacity Utilisation, ZEW Expectation Growth
- Labour markets

Non-conventional features:

- Futures German 30Y bond
- Futures 3M Euribor & 90d Euro$
- Adjacent maturities
- Intra and inter-curve spreads
- OECD Leading Indicators
- ECB Balance Sheet indicators

LSTM-LagLasso for Feature Selection.

Conclusions and Future Work

UNIVERSITY OF Southampton

MAIN CONCLUSIONS

The memory advantage of LSTM networks & Signals analysis

LSTMs with different input sequences vs. MLPs

- LSTMs with longer input sequences achieve similar levels of forecasting accuracy to the MLP with the most relevant features, with lower standard deviation.
  - Additional memory in the LSTM "equivalent" to additional information in the MLP.

MAIN CONCLUSIONS

The memory advantage of LSTM networks & Signals analysis

LSTMs with different input sequences vs. MLPs

- LSTMs with longer input sequences achieve similar levels of forecasting accuracy to the MLP with the most relevant features, with lower standard deviation.
  - Additional memory in the LSTM "equivalent" to additional information in the MLP.
- Shorter forecasting horizons require smaller input sequences and vice-versa.
  - Need to adjust the LSTM architecture to the forecasting horizon and in general terms to the conditions of the problem.

UNIVERSITY OF Southampton

MAIN CONCLUSIONS

The memory advantage of LSTM networks & Signals analysis

LSTMs with different input sequences vs. MLPs

- LSTMs with longer input sequences achieve similar levels of forecasting accuracy to the MLP with the most relevant features, with lower standard deviation.
  - Additional memory in the LSTM "equivalent" to additional information in the MLP.
- Shorter forecasting horizons require smaller input sequences and vice-versa.
  - Need to adjust the LSTM architecture to the forecasting horizon and in general terms to the conditions of the problem.

Signals analysis

- The most remarkable property found consistently in the LSTM signals, is the activation / deactivation of units through time.
- LSTM units tend to specialise in different yield ranges or features.

UNIVERSITY OF Southampton

MAIN CONCLUSIONS

LSTM-LagLasso - Explaining the signals

LSTM-LagLasso

- The information contained in the LSTM states is complex, but may be explained by exogenous variables.
- It identifies lags as important, in particular t, t-1 and t-5.

MAIN CONCLUSIONS

LSTM-LagLasso - Explaining the signals

LSTM-LagLasso

- The information contained in the LSTM states is complex, but may be explained by exogenous variables.
- It identifies lags as important, in particular t, t-1 and t-5.
- LSTM-LagLasso may be used as an alternative feature selection method.
- On the relevant features, it selects conventional as well as non-conventional market/macro indicators, contributing to the prediction process, but which are not commonly used in forecasting models.

MAIN CONCLUSIONS

LSTM-LagLasso - Explaining the signals

LSTM-LagLasso

- The information contained in the LSTM states is complex, but may be explained by exogenous variables.
- It identifies lags as important, in particular t, t-1 and t-5.
- LSTM-LagLasso may be used as an alternative feature selection method.
- On the relevant features, it selects conventional as well as non-conventional market/macro indicators, contributing to the prediction process, but which are not commonly used in forecasting models.
- Finally, the LSTMs can capture this information and maintain it in the states through time.

FUTURE WORK

- Financial asset forecasting is one type of problem. Portfolio Management and Trading is a different one.
- Future work will focus on the implementation of this type of model in **autonomous systems**.

Thank You!