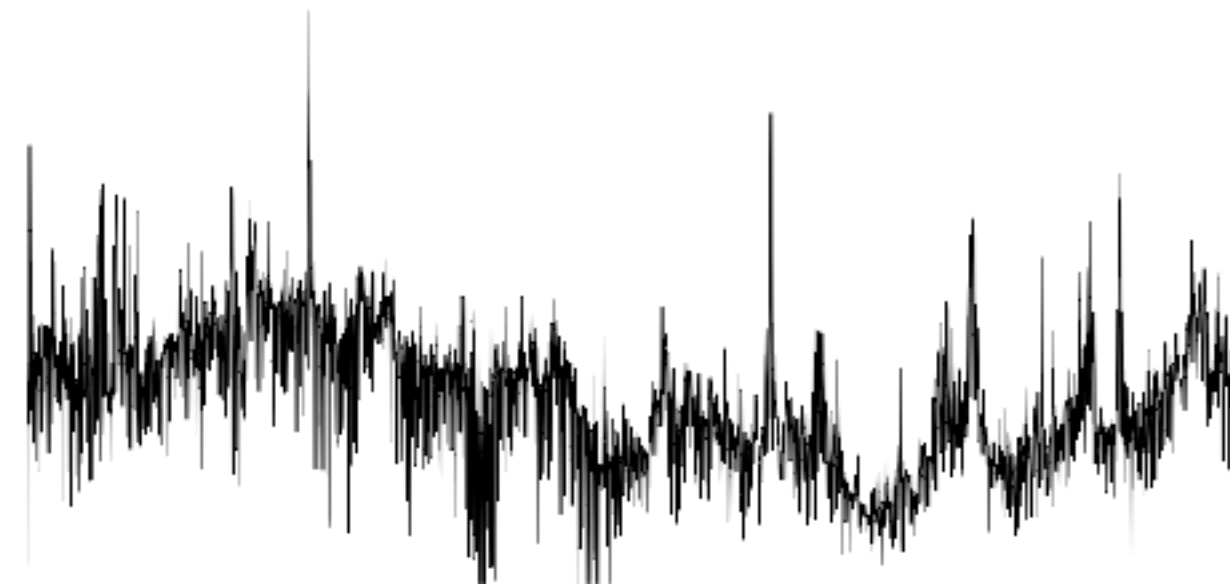


ELECTRICITY PRICE FORECASTING



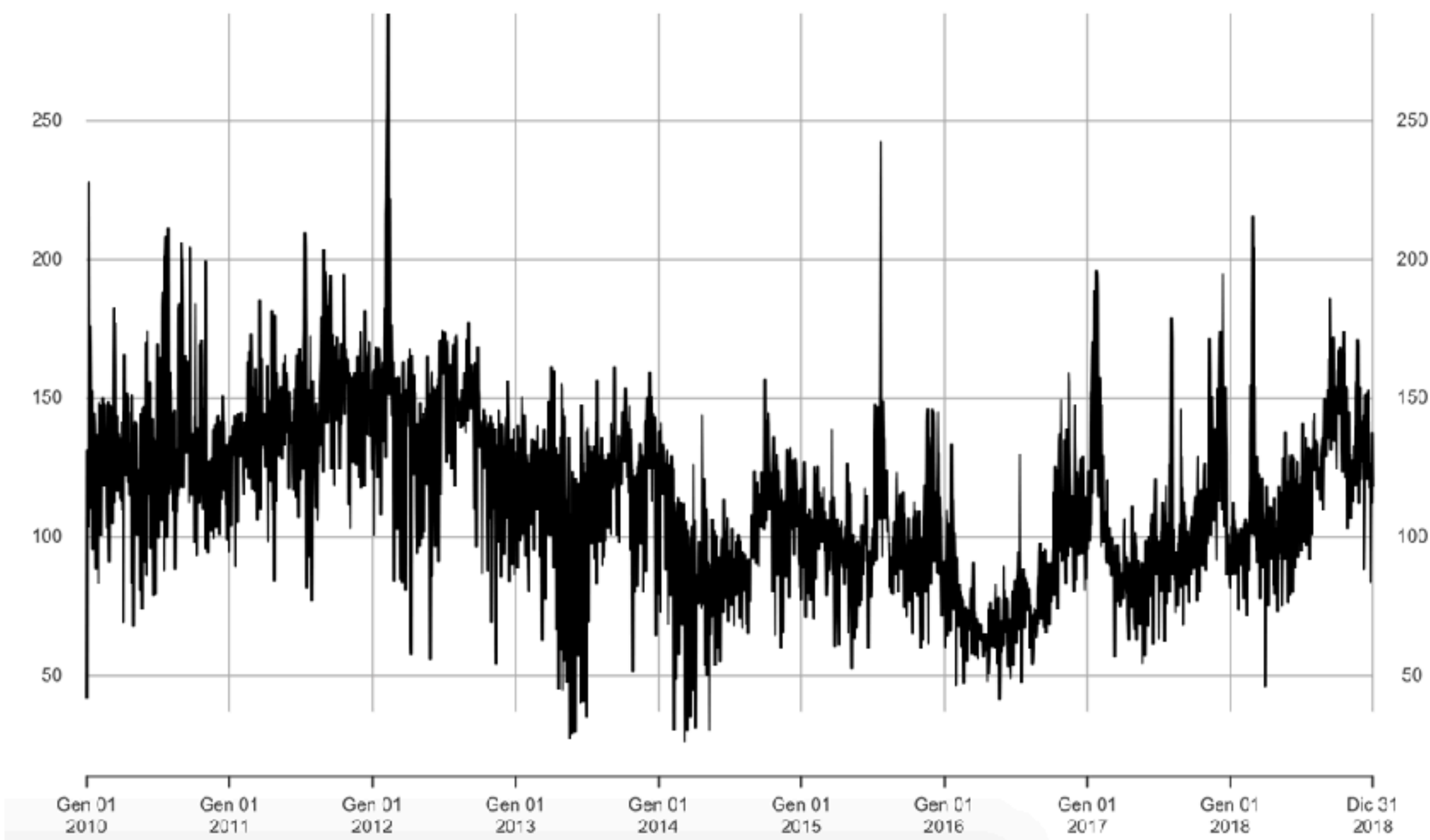
Streaming Data Management and
Time Series Analysis

Academic Year 2019/2020

PROJECT WORKFLOW

THE TASK

The goal is the daily prediction over a 11 months horizon using 9-years price data with ARIMA, UCM and Machine Learning models



PREPROCESSING

- Logarithmic Transformation
- Train-Val Split (90% - 10%)

METRICS

- MAE

MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON
AND
CONCLUSIONS

MODELS

ARIMA

UCM

KNN

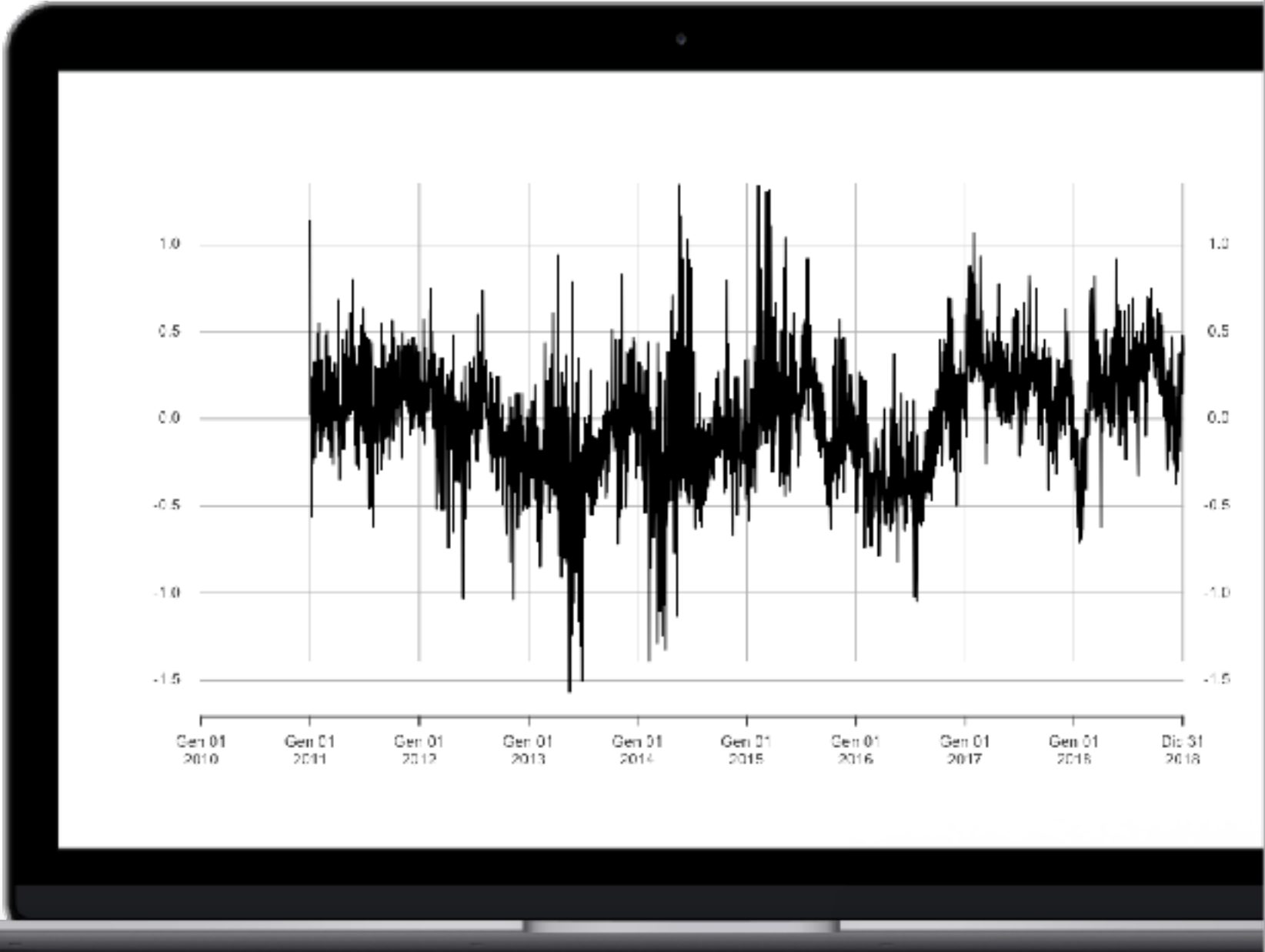
LSTM RNN

COMPARISON
AND
CONCLUSIONS

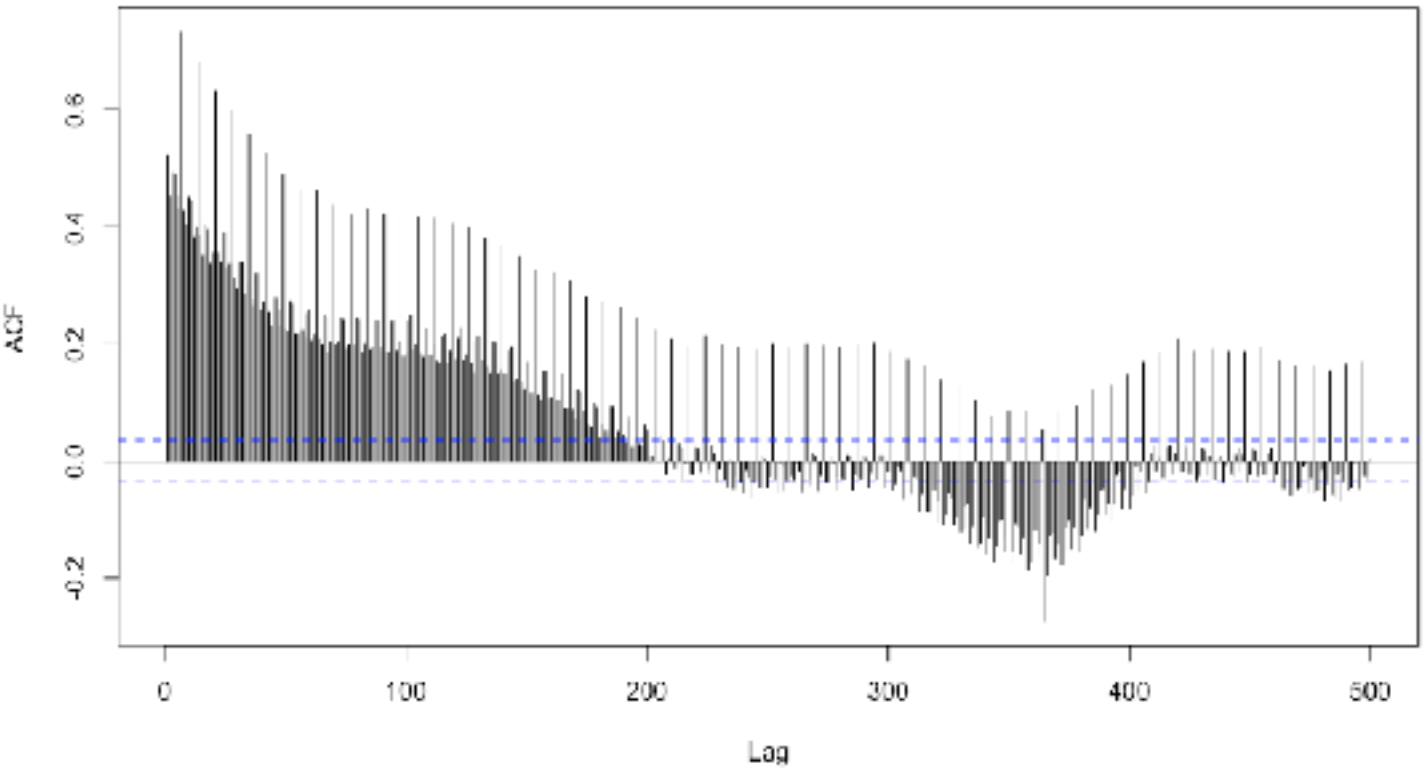
ARIMA

Identification

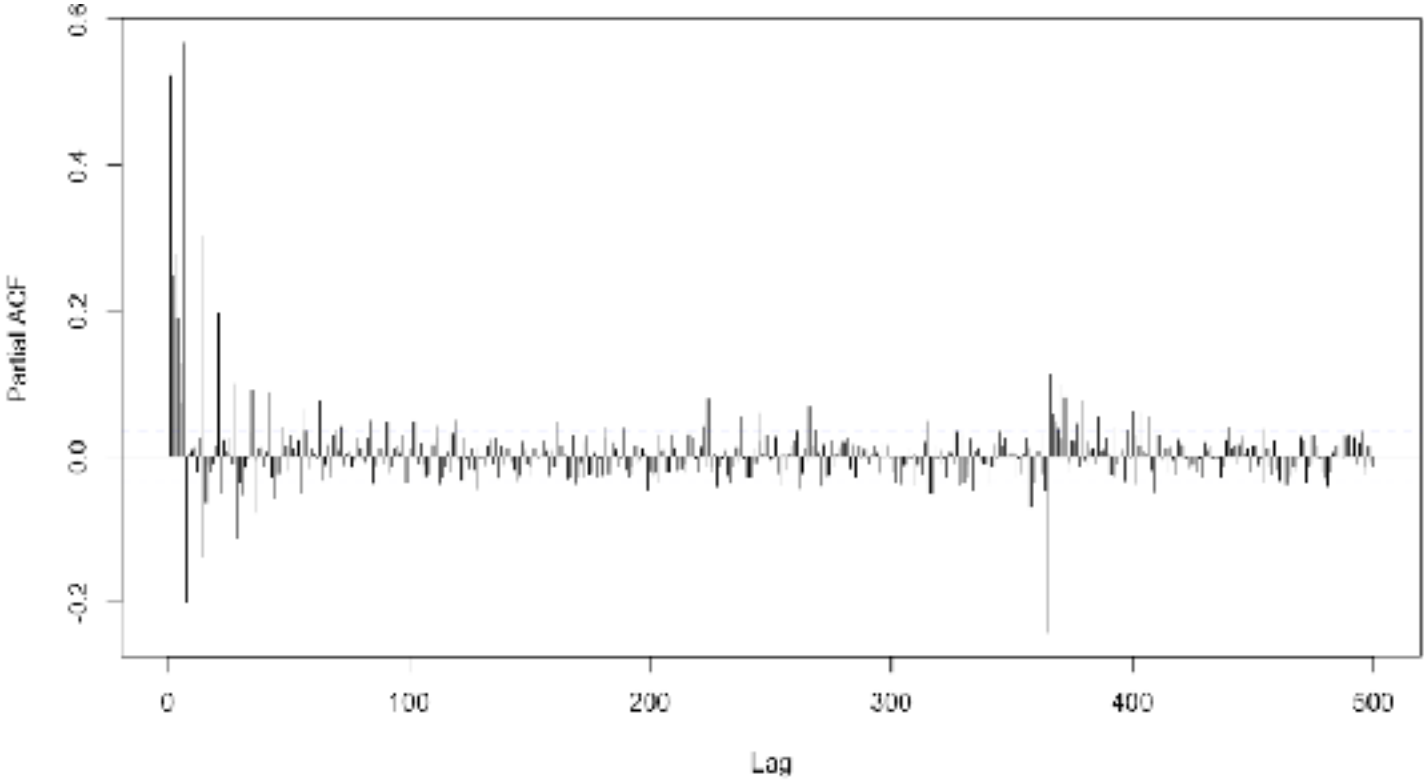
Ly with differences (order 365.25)



ACF



PACF



MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON AND CONCLUSIONS

ARIMA

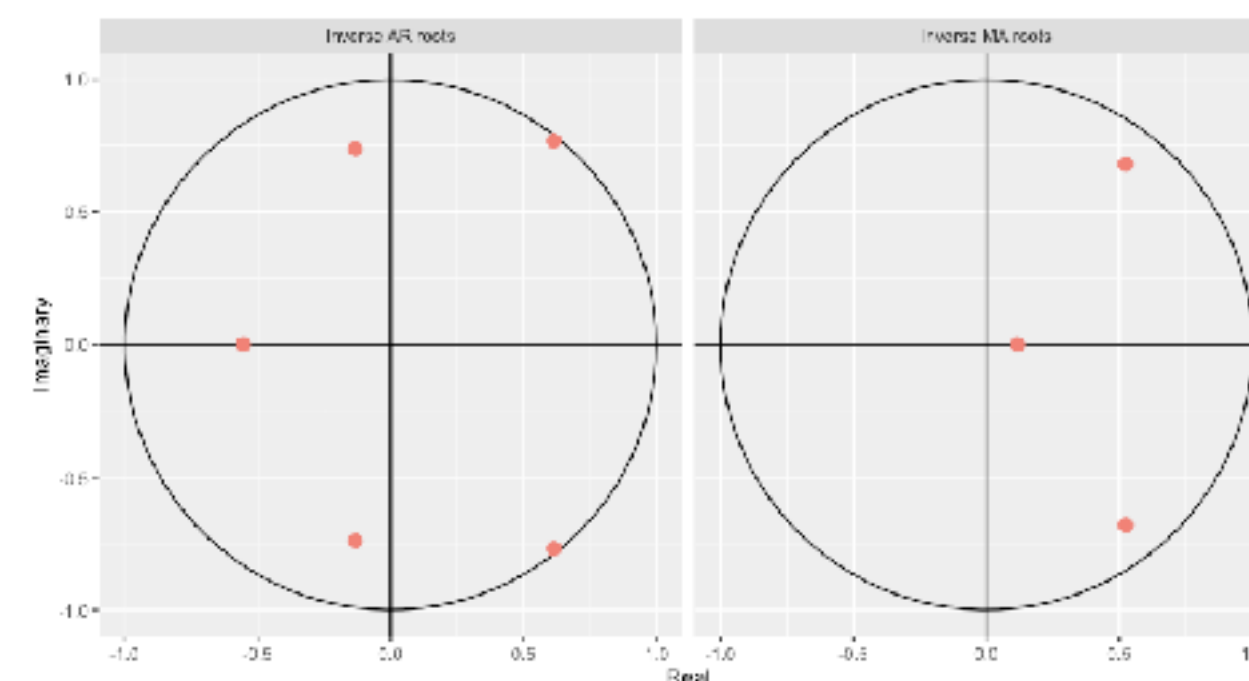
Estimation

ISSUES

- R can't handle differences of order greater than 350. SARIMA is not applicable for daily data

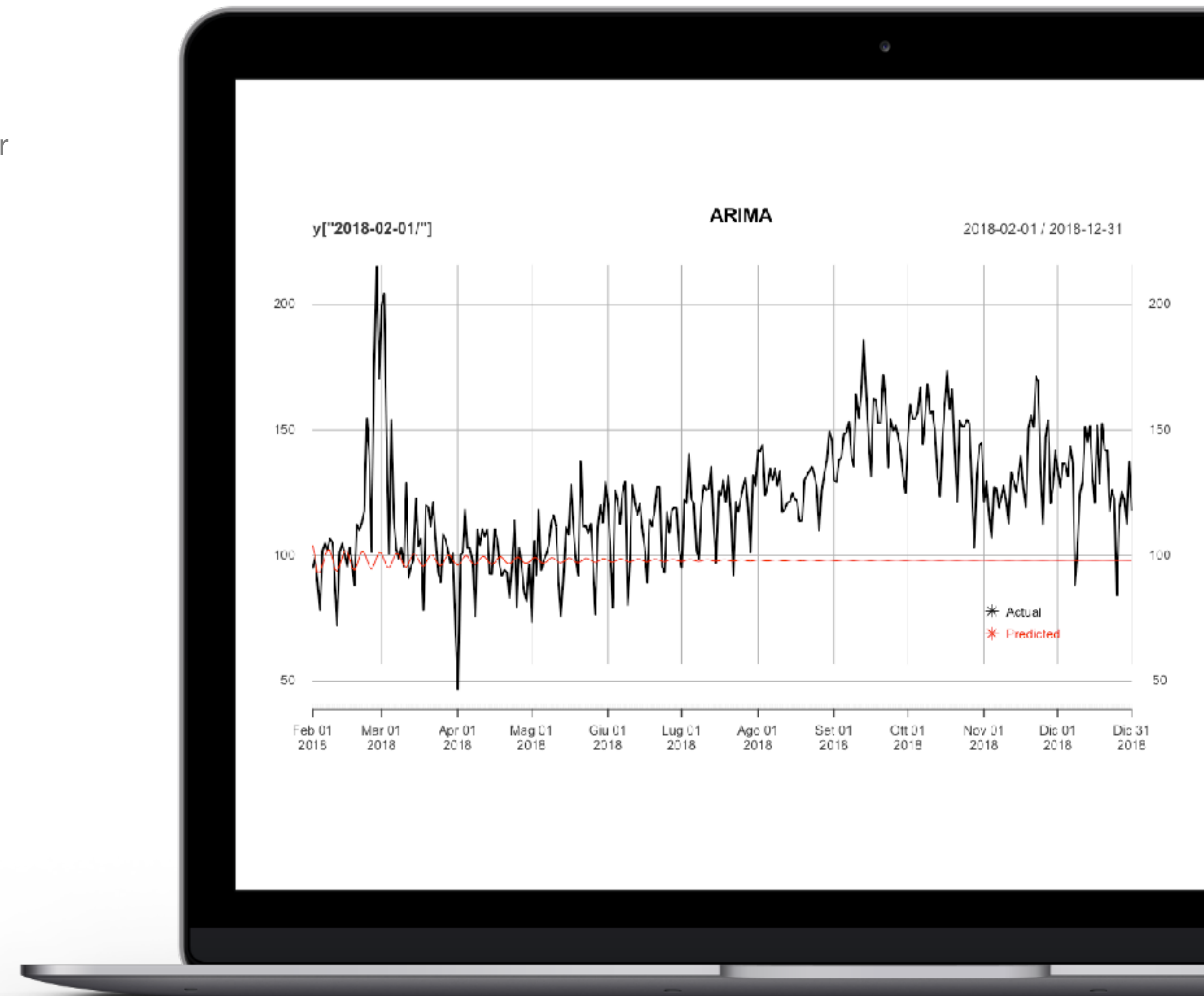
MODEL

- AutoARIMA is deployed for the identification
- ARIMA(5,1,3) is used for forecasting



EVALUATION

- MAE = 0.236165



MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON
AND
CONCLUSIONS

U C M

Overview

MODEL	COMPONENTS
1	LLT + DUMMY SEASONALITY (7)
2	IRW + CYCLE + DUMMY SEASONALITY (7)
3	IRW + CYCLE + TRIG DETERMINISTIC SEASONALITY (365)
4	LLT + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)
5	IRW + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)
6	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG DETERMINISTIC SEASONALITY (365)
7	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)
8	LLT + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)

MODELS

ARIMA

UCM

KNN

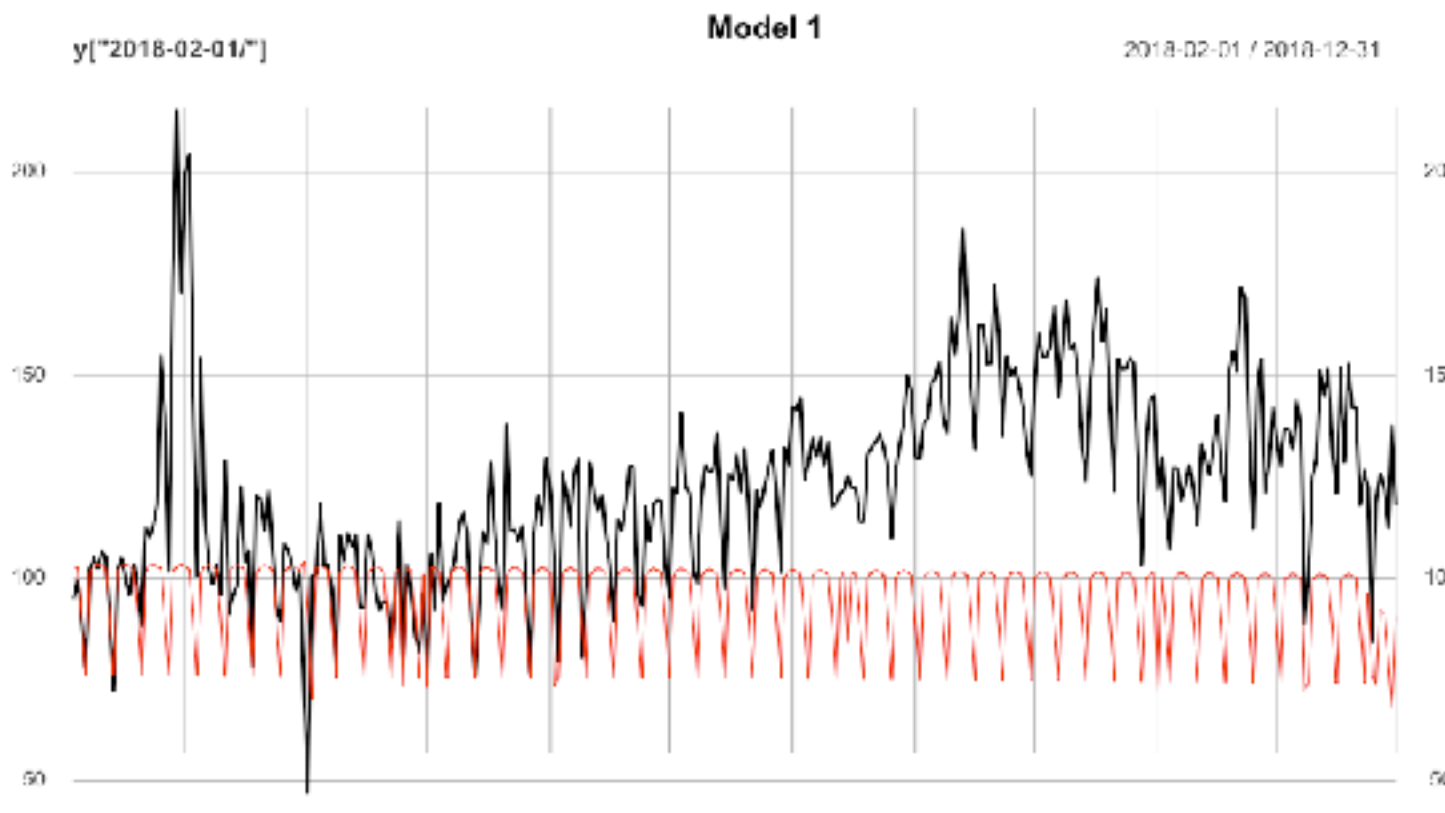
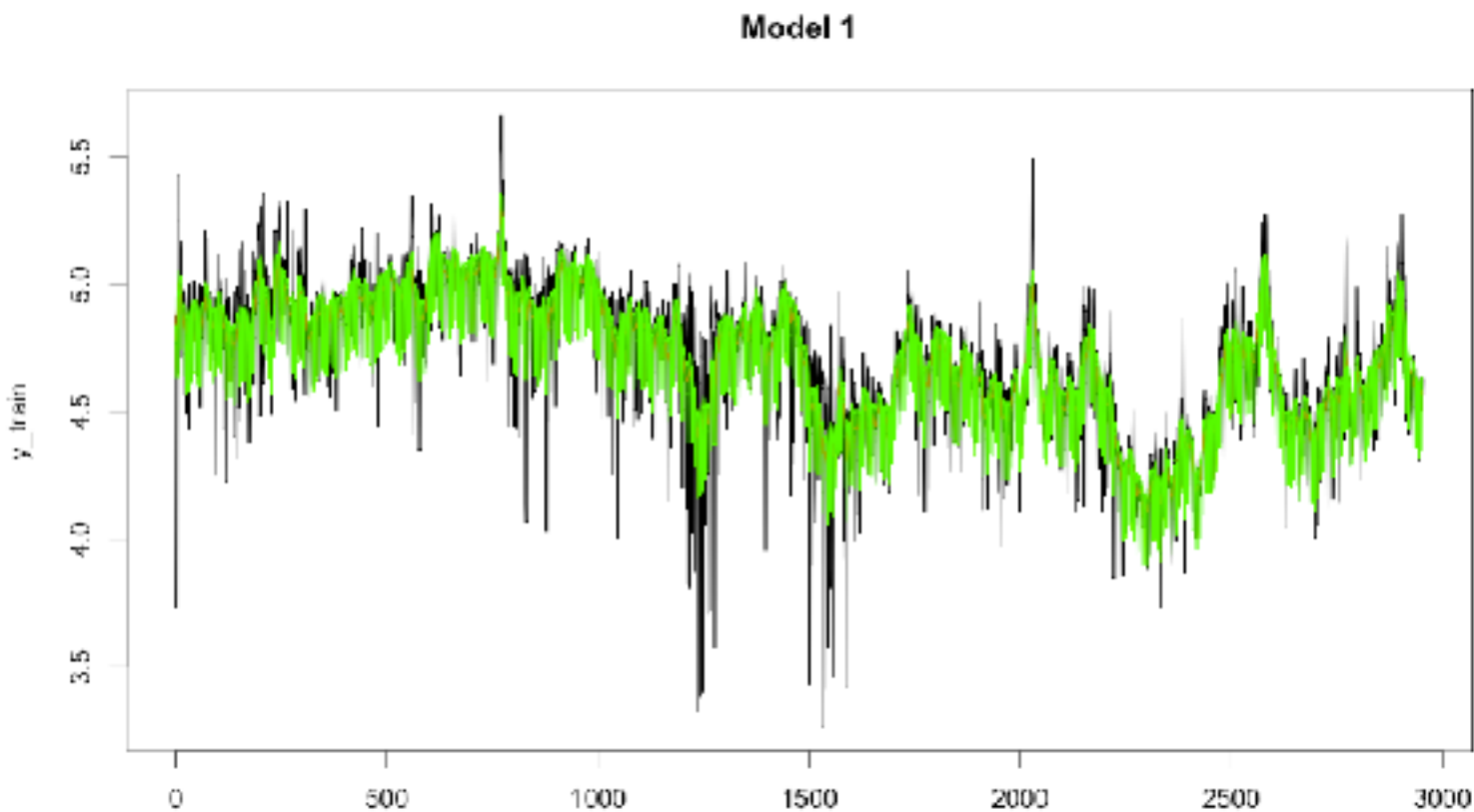
LSTM RNN

COMPARISON
AND
CONCLUSIONS

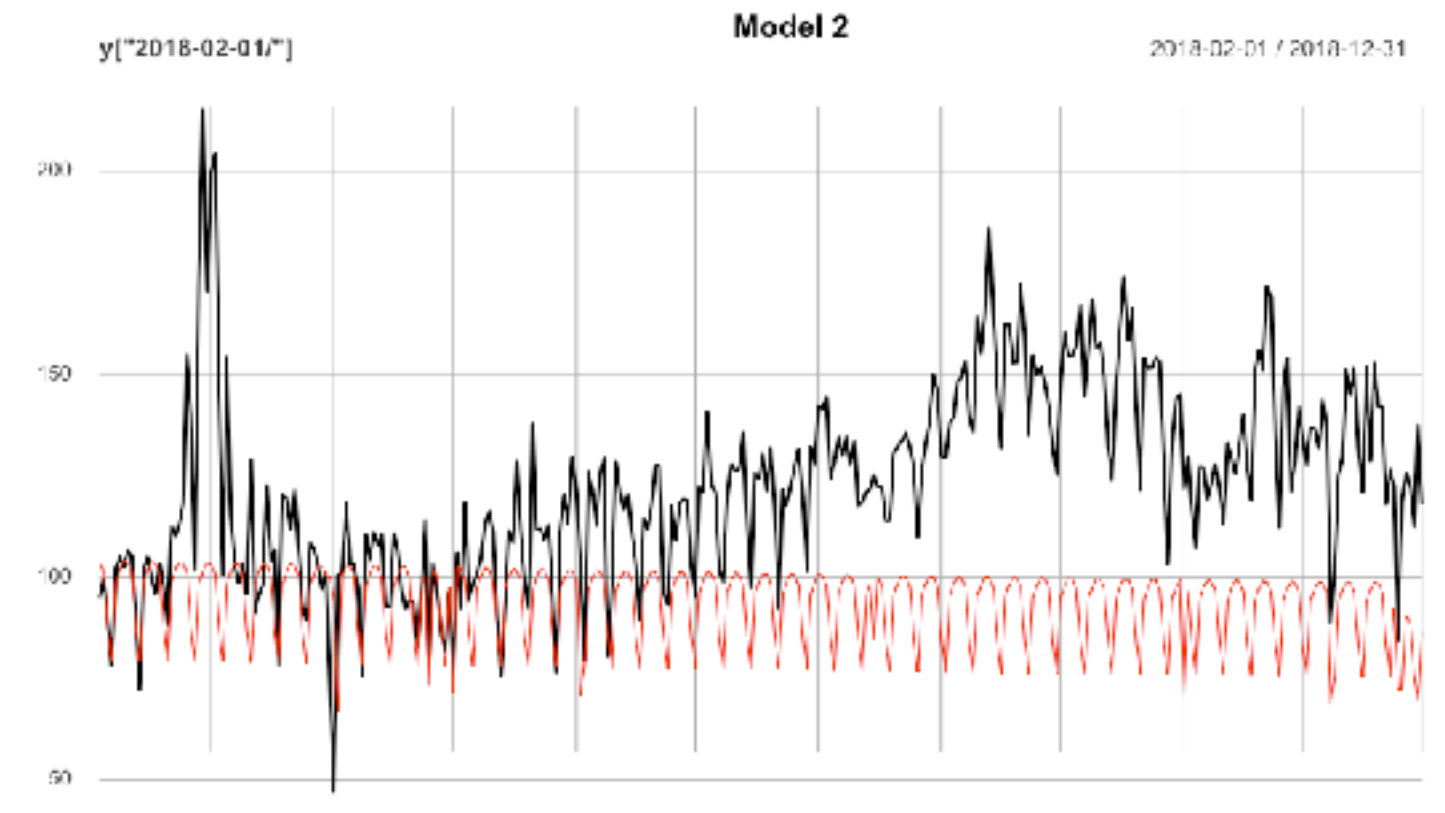
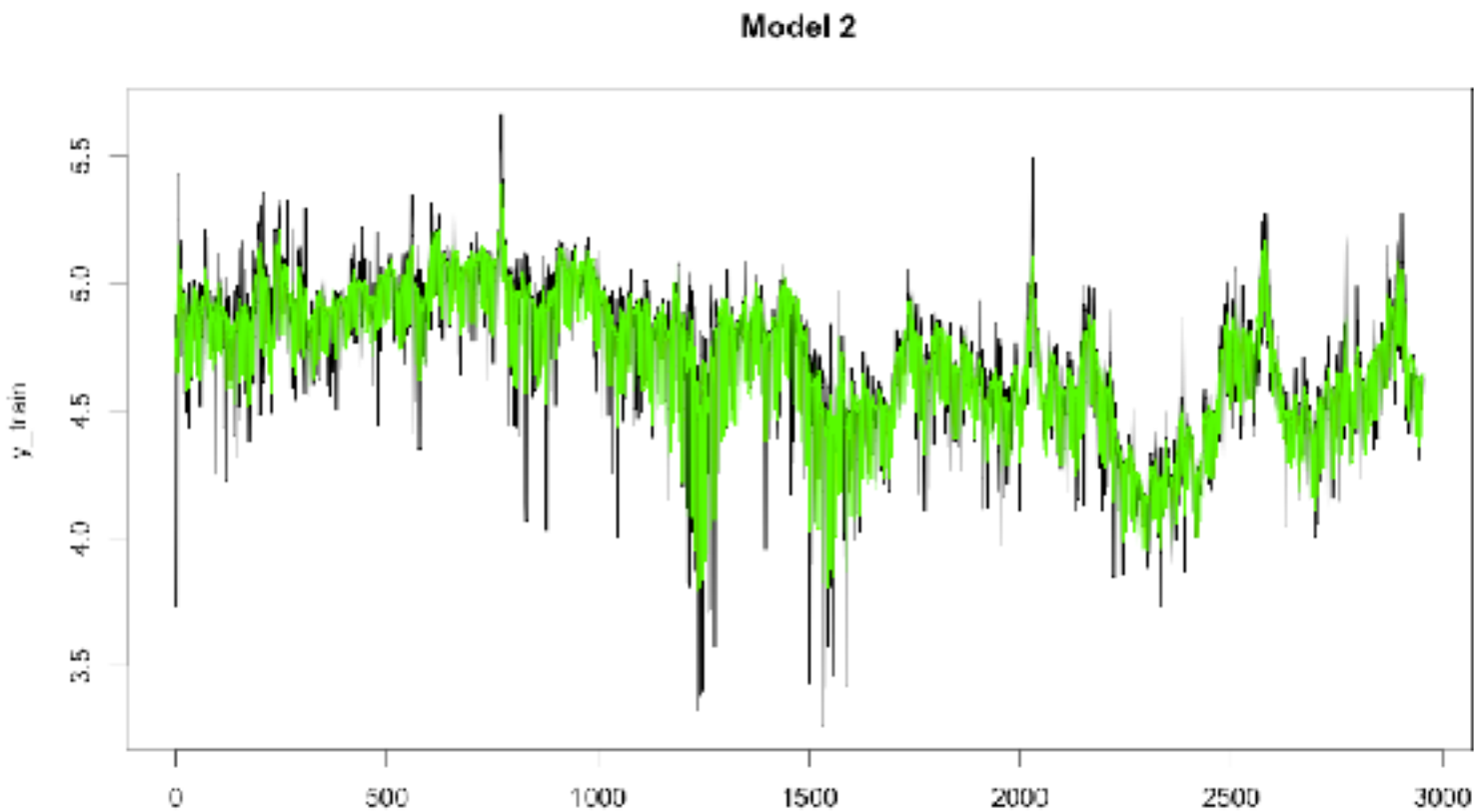
U C M

MODELS 1-2

LLT + DUMMY SEASONALITY (7)



IRW + CYCLE + DUMMY SEASONALITY (7)



MODELS

ARIMA

UCM

KNN

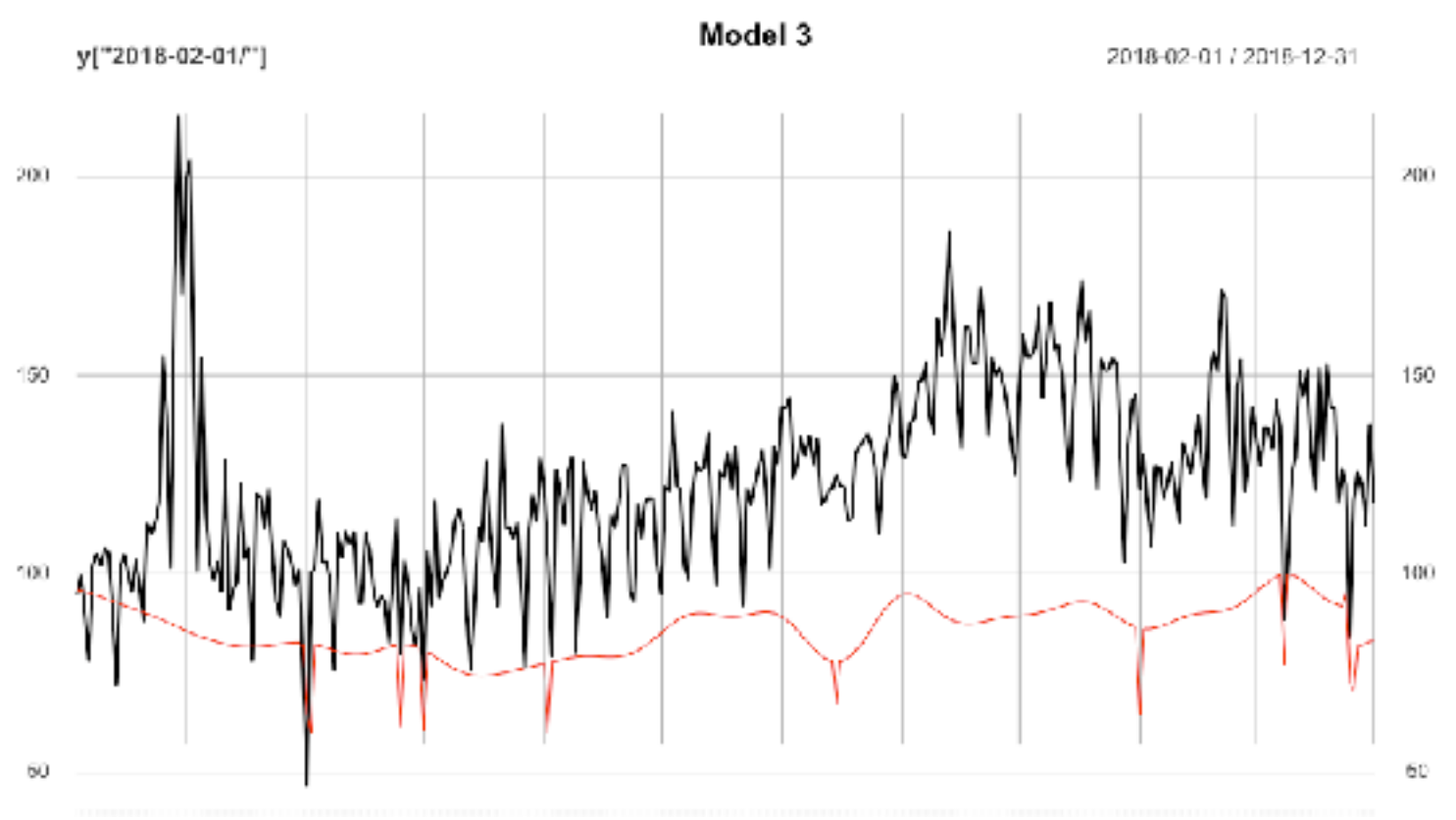
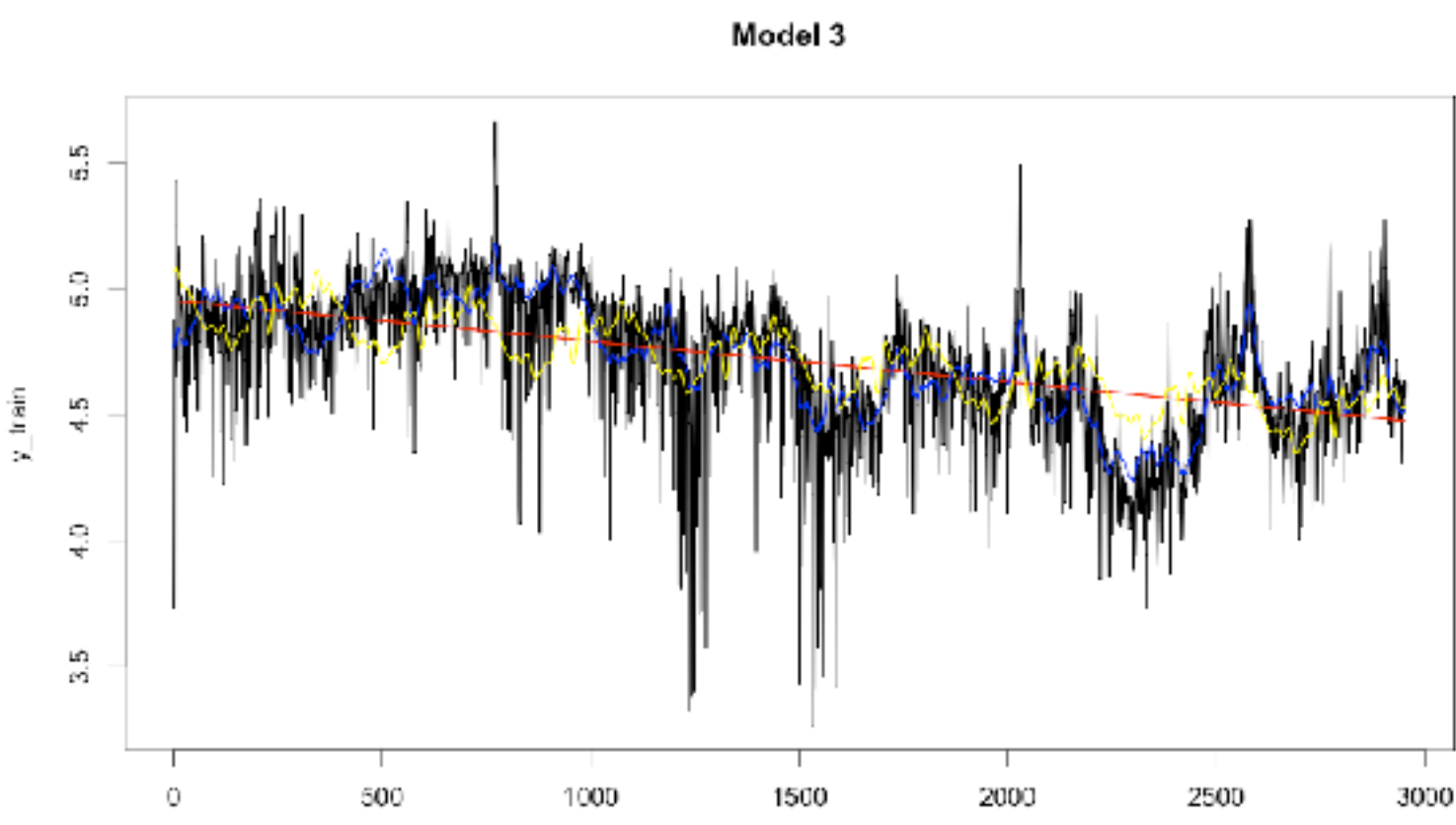
LSTM RNN

COMPARISON
AND
CONCLUSIONS

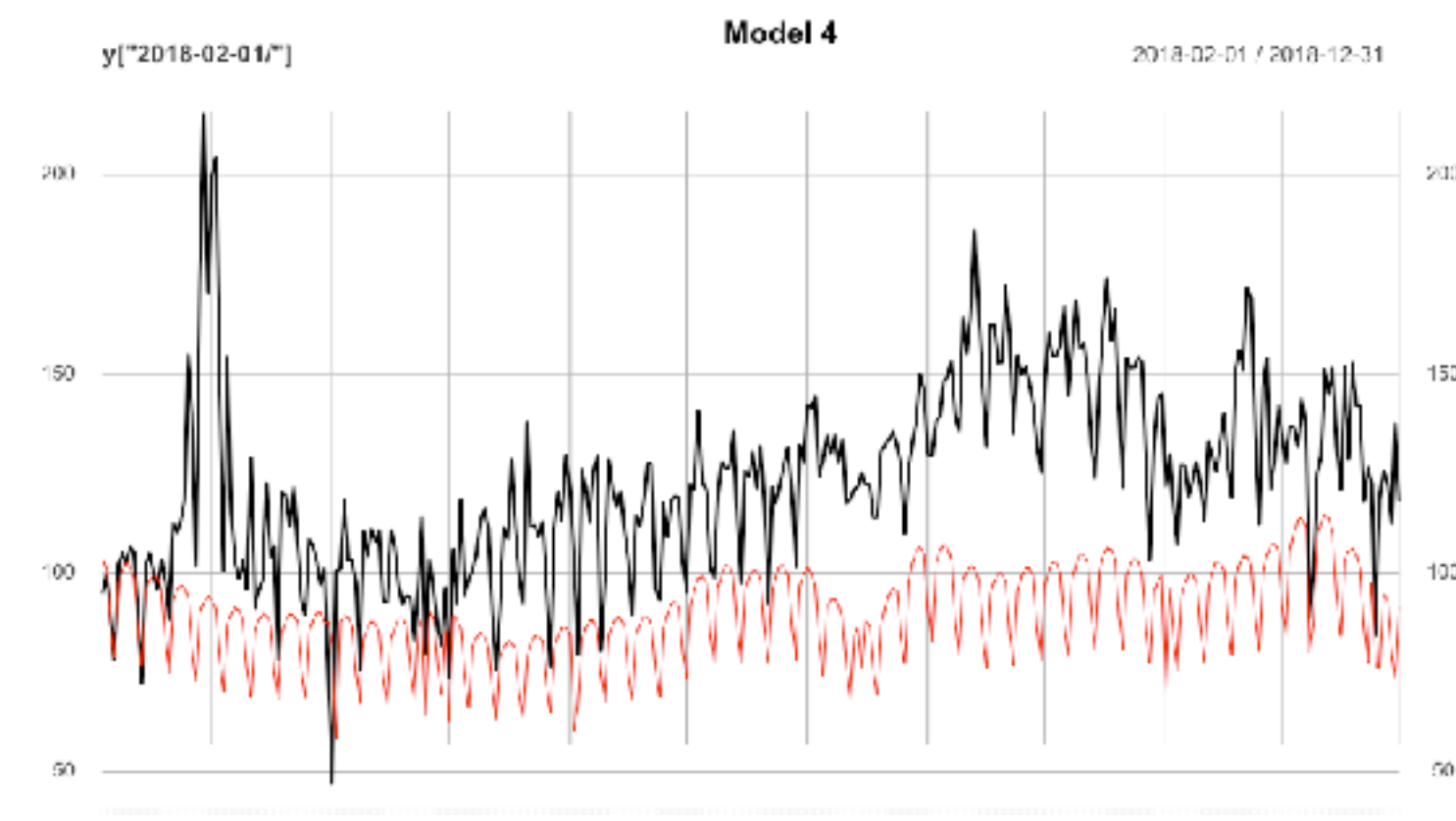
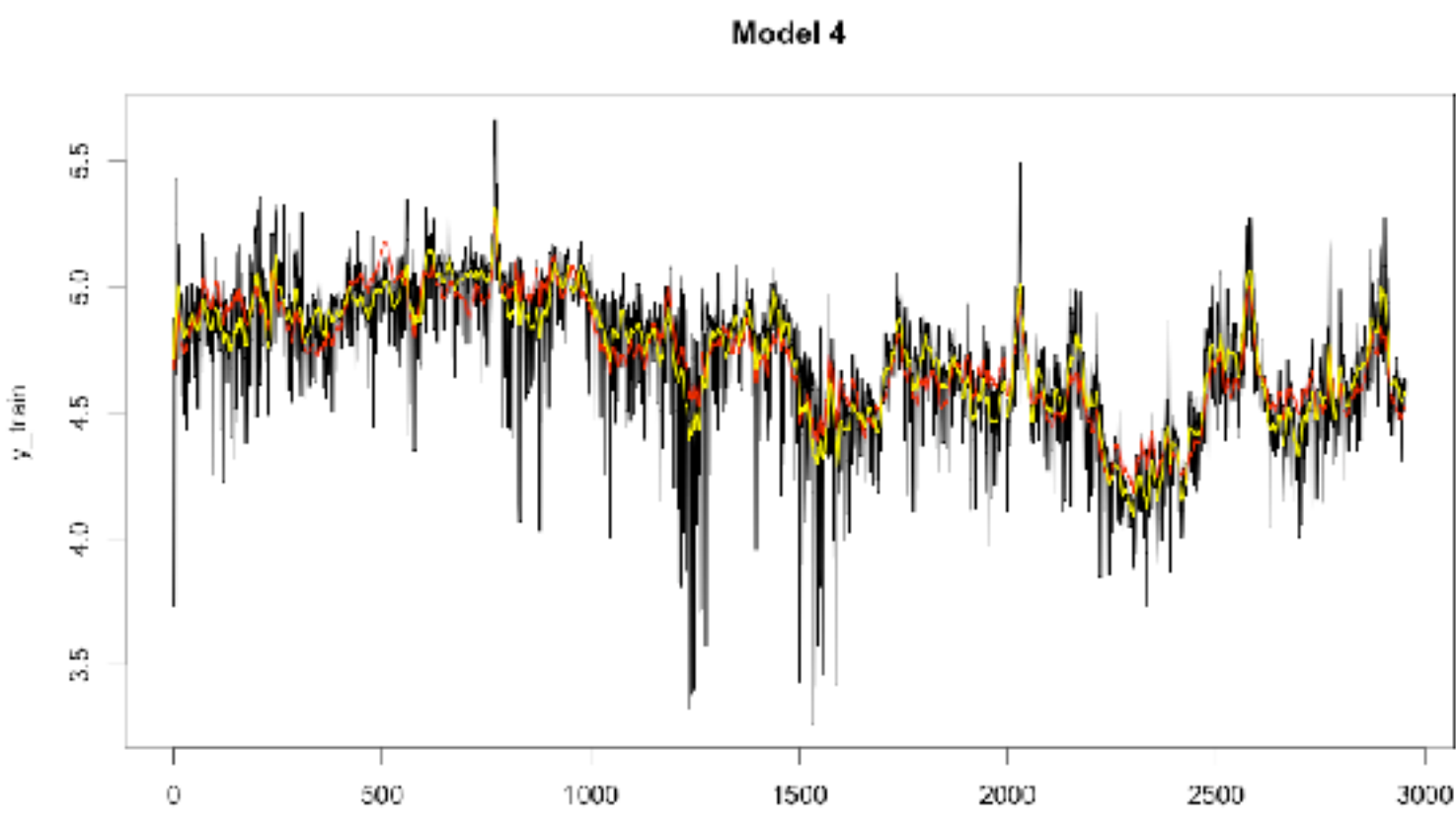
U C M

MODELS 3-4

IRW + CYCLE + TRIG DETERMINISTIC SEASONALITY (365)



LLT + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)



MODELS

ARIMA

UCM

KNN

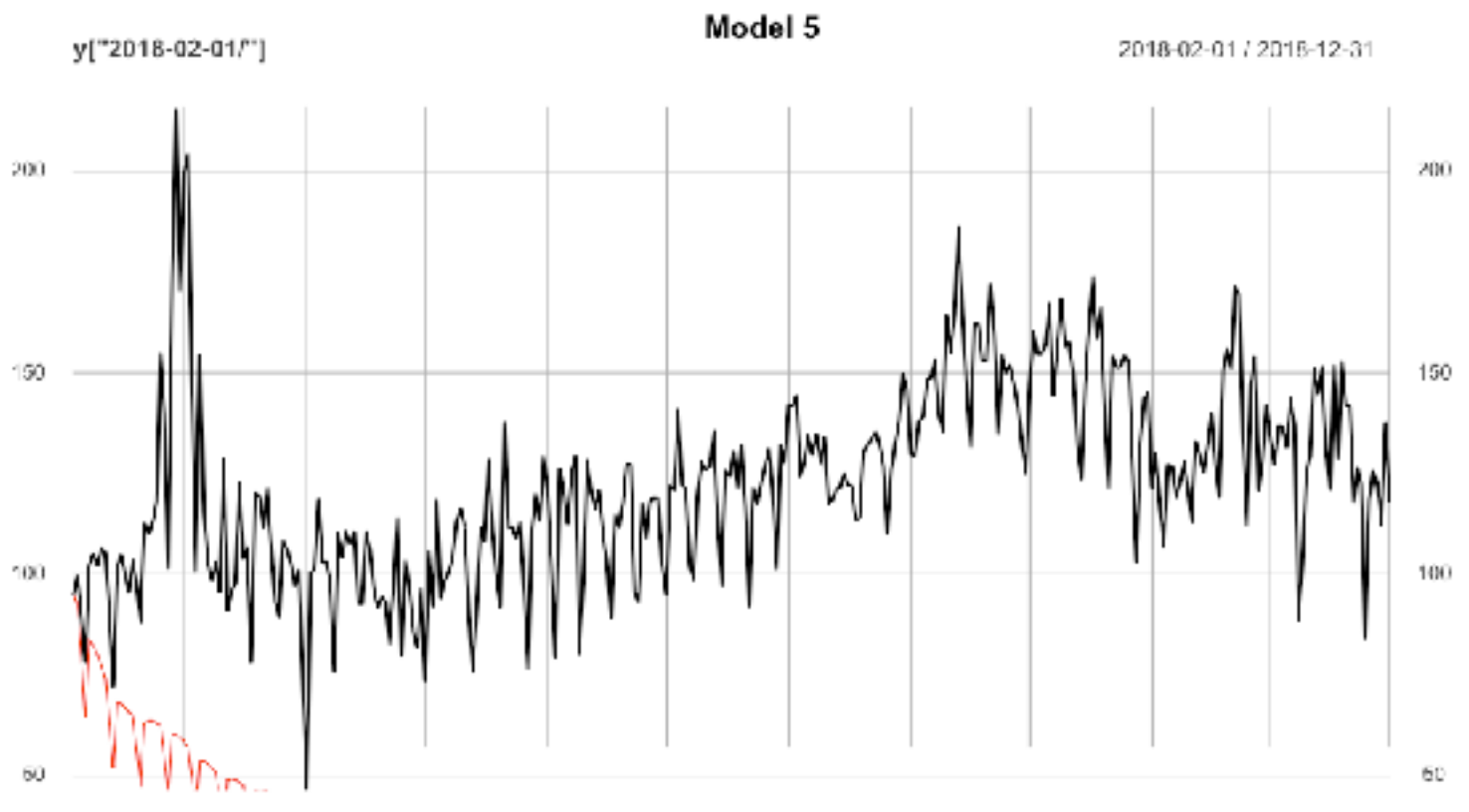
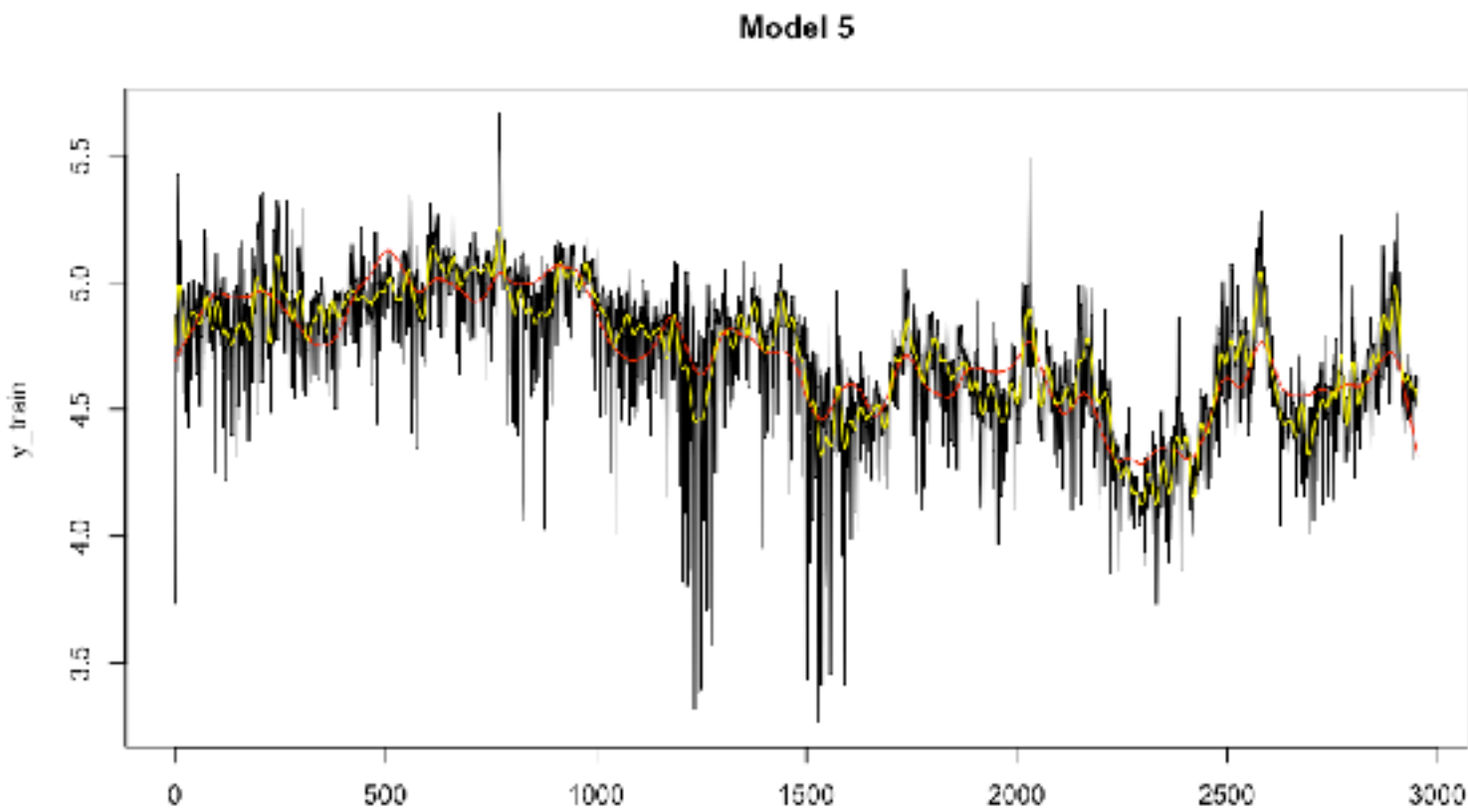
LSTM RNN

COMPARISON
AND
CONCLUSIONS

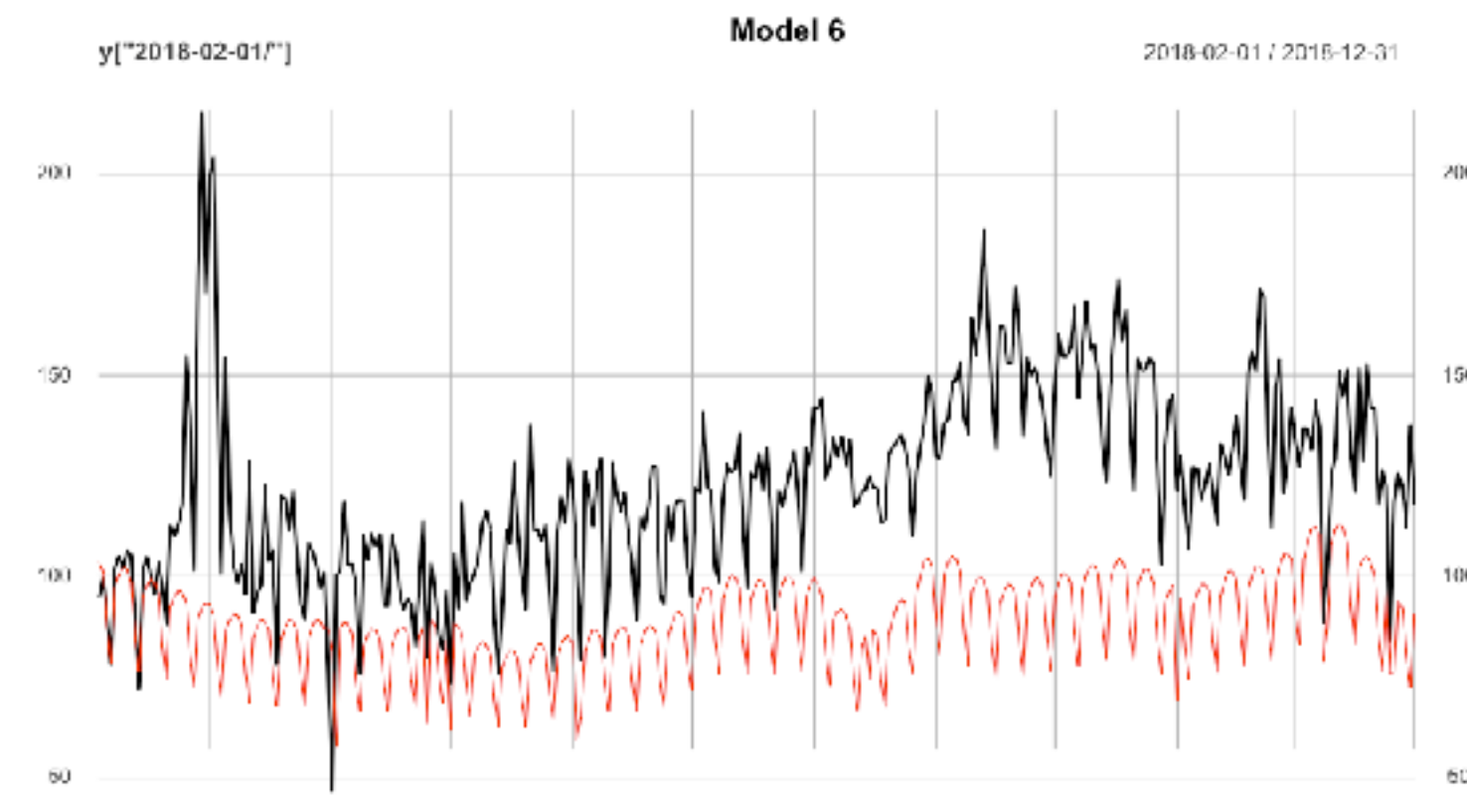
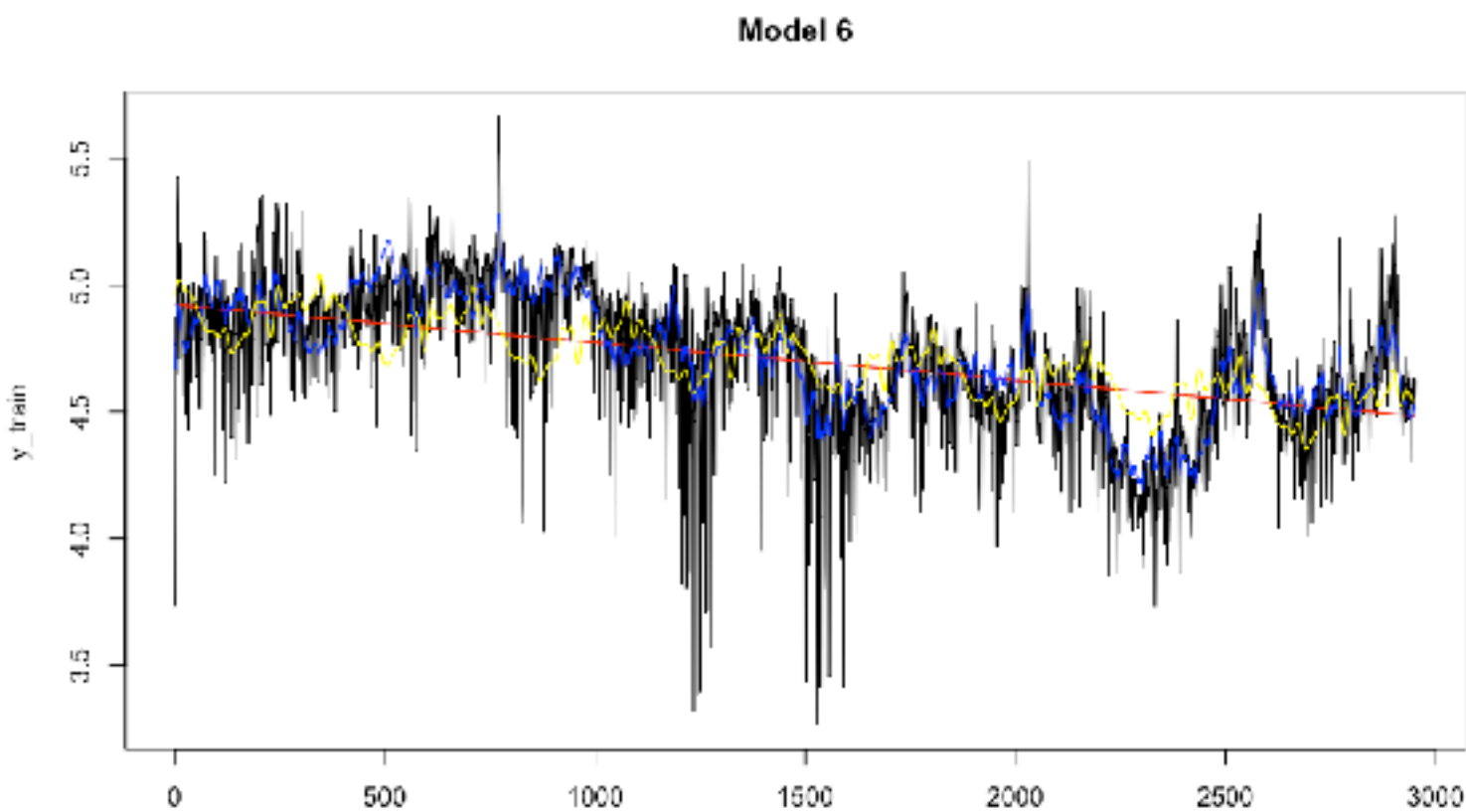
U C M

MODELS 5-6

IRW + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)



IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG DETERMINISTIC SEASONALITY (365)



MODELS

ARIMA

UCM

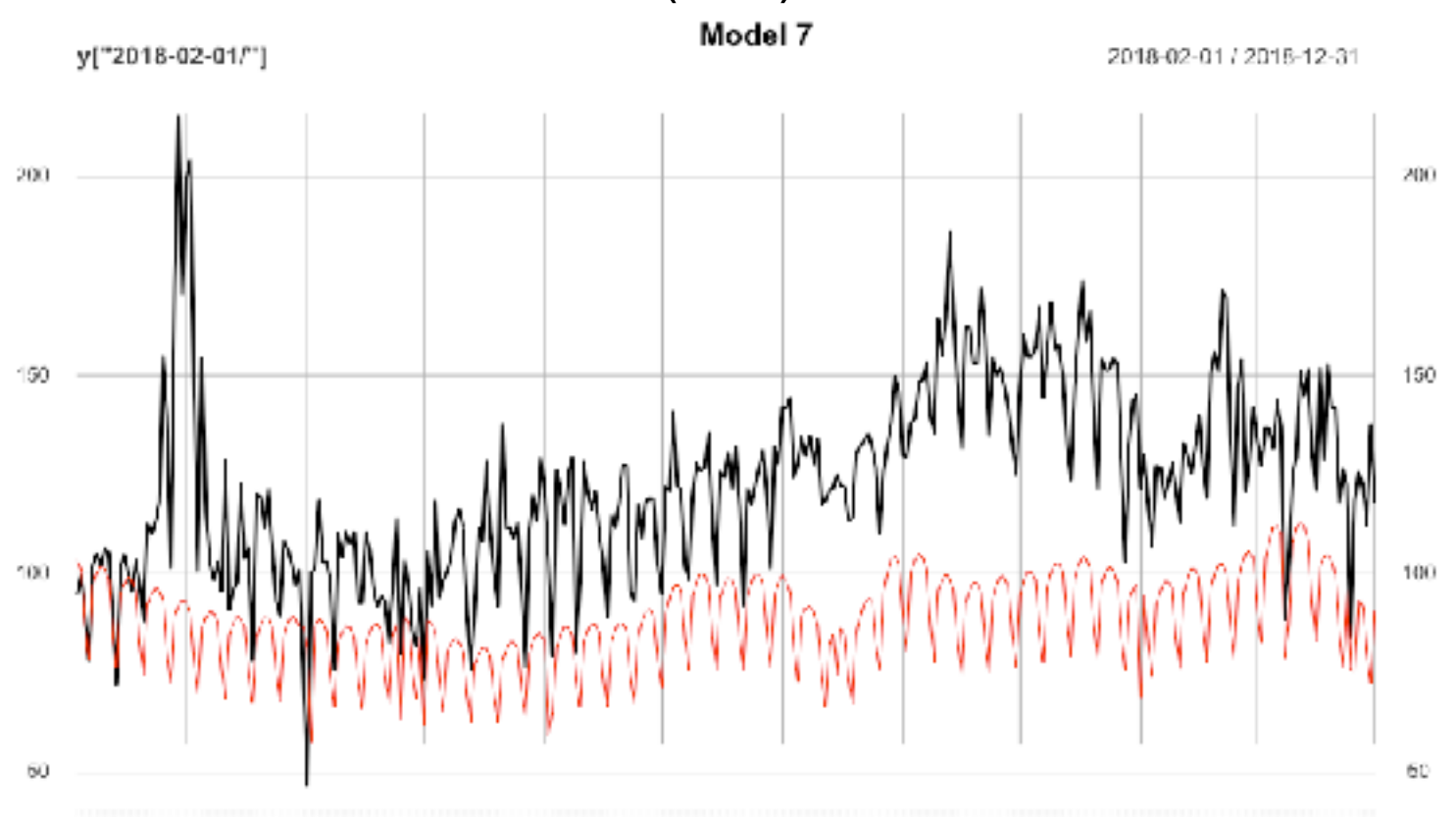
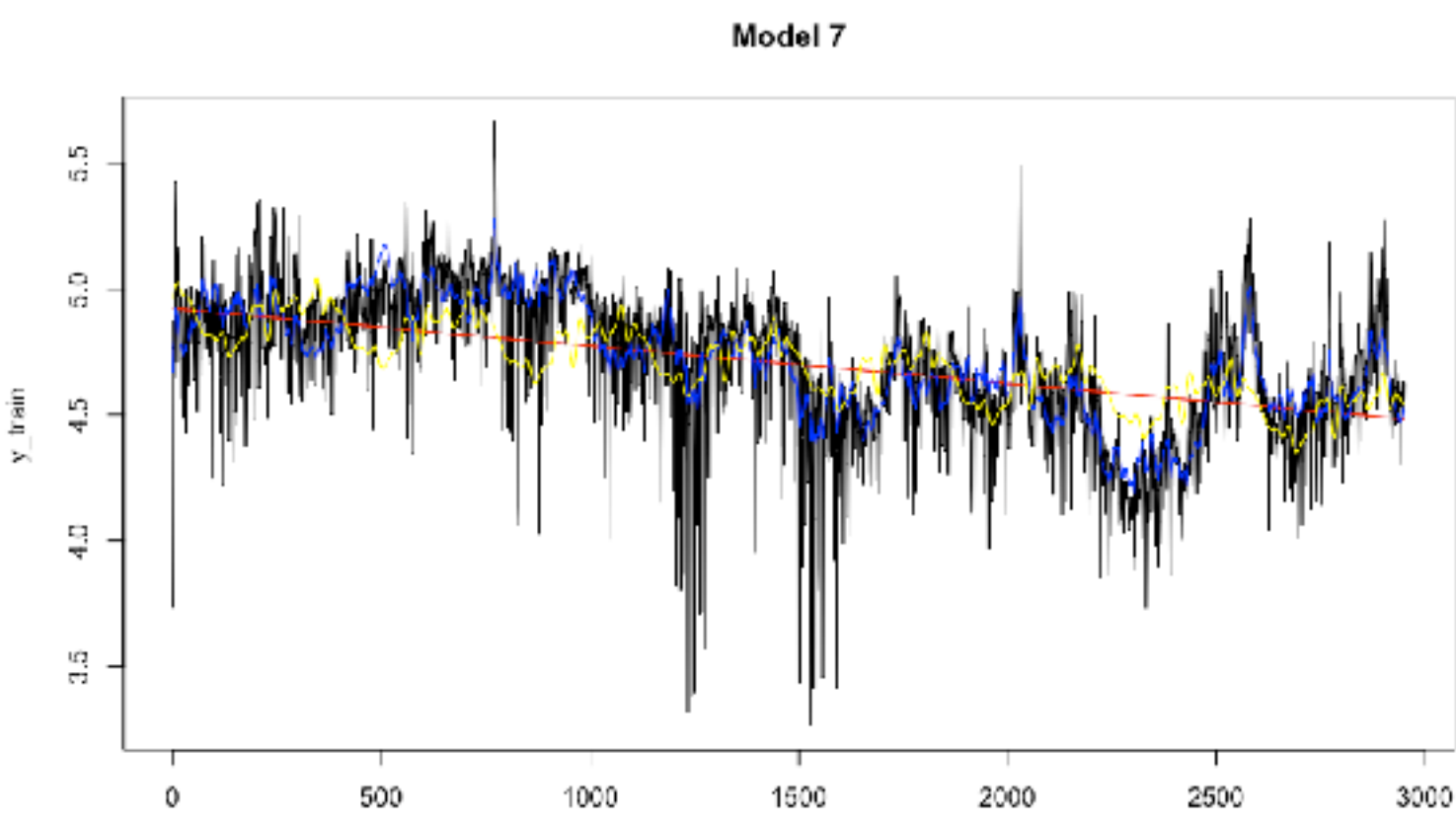
KNN

LSTM RNN

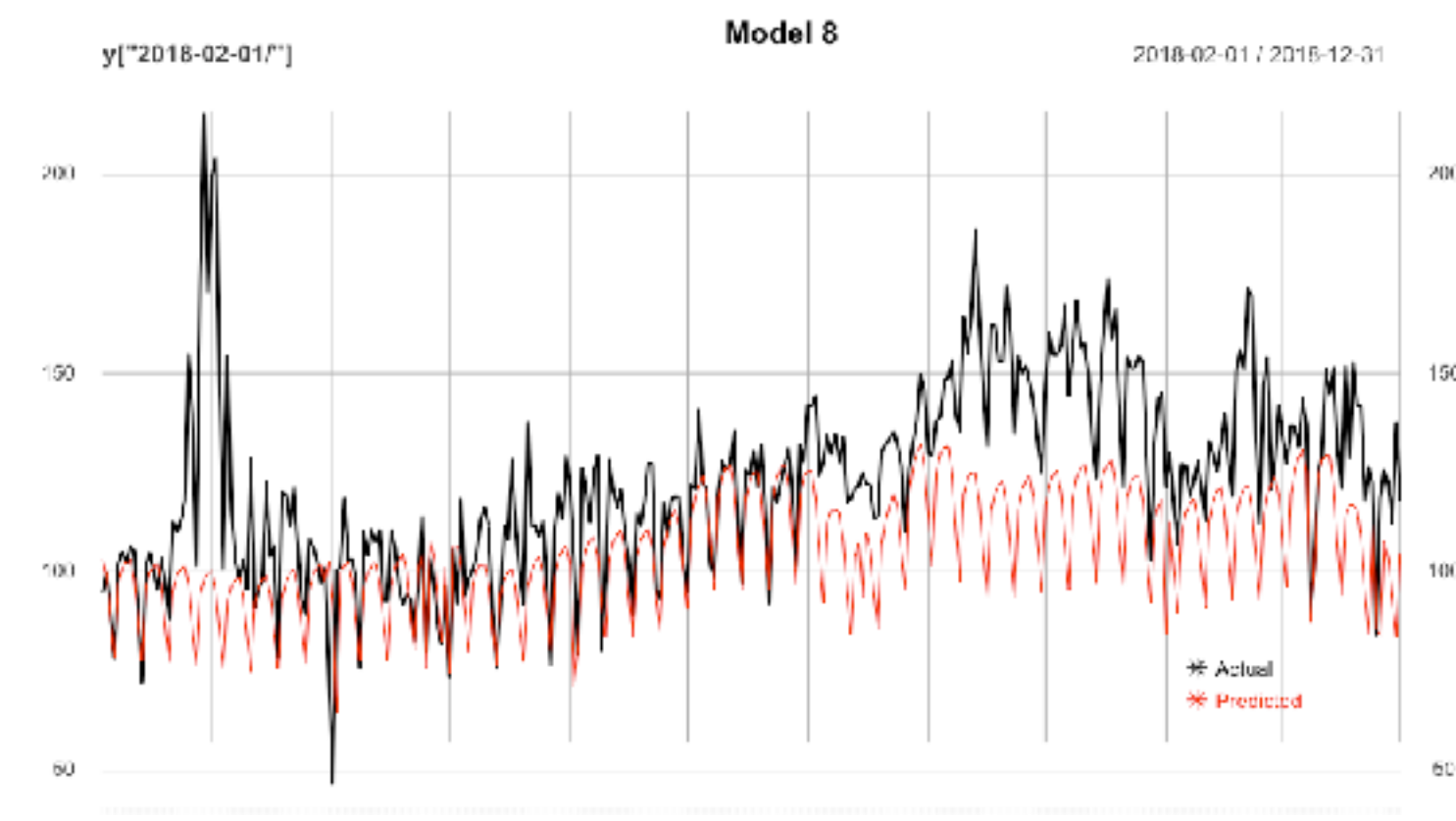
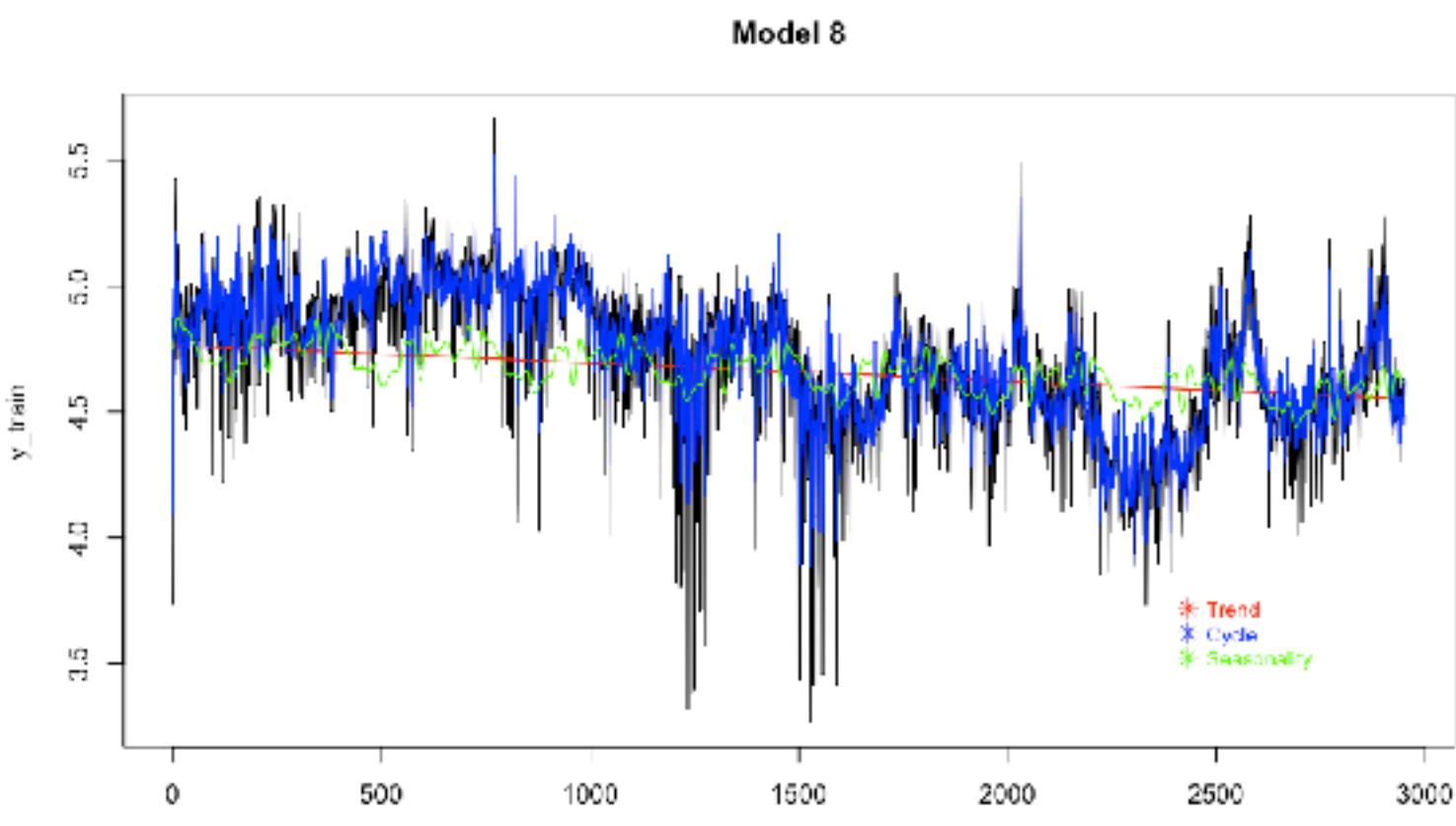
COMPARISON AND CONCLUSIONS

UCM MODELS 7-8

IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)



LLT + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)



MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON
AND
CONCLUSIONS

MODEL 8
UPDATE FUNCTION

U C M

```
1149 # Update function
1150 updt8 <- function(pars, model) {
1151   model$Q[1, 1, 1] <- exp(pars[1]) #level
1152   model$Q[2, 2, 1] <- exp(pars[2]) #slope
1153   model$Q[3, 3, 1] <- exp(pars[4]) #seas dummy
1154   diag(model$Q[4:35, 4:35, 1]) <- exp(pars[5]) #seas trig
1155   model$Q[36, 36, 1] <- model$Q[37, 37, 1] <- exp(pars[3]) #cycle
1156   rho <- ext_sigmoid(pars[7]) * 0.99
1157   per <- ext_sigmoid(pars[8], 1825, 2555)
1158   lam <- 2*pi/per
1159   vpsi <- model$Q[36, 36, 1] / (1 - rho^2)
1160   rho_co <- rho*cos(lam)
1161   rho_si <- rho*sin(lam)
1162   s <- 365
1163   model$T[23:24, 23:24, 1] <- c(cos((2*pi) / s), -sin((2*pi) / s),
1164                                   sin((2*pi) / s), cos((2*pi) / s))
1165   model$T[25:26, 25:26, 1] <- c(cos((2*pi) / s * 2), -sin((2*pi) / s * 2),
1166                                   sin((2*pi) / s * 2), cos((2*pi) / s * 2))
1167   model$T[27:28, 27:28, 1] <- c(cos((2*pi) / s * 3), -sin((2*pi) / s * 3),
1168                                   sin((2*pi) / s * 3), cos((2*pi) / s * 3))
1169   model$T[29:30, 29:30, 1] <- c(cos((2*pi) / s * 4), -sin((2*pi) / s * 4),
1170                                   sin((2*pi) / s * 4), cos((2*pi) / s * 4))
1171   model$T[31:32, 31:32, 1] <- c(cos((2*pi) / s * 5), -sin((2*pi) / s * 5),
1172                                   sin((2*pi) / s * 5), cos((2*pi) / s * 5))
1173   model$T[33:34, 33:34, 1] <- c(cos((2*pi) / s * 6), -sin((2*pi) / s * 6),
1174                                   sin((2*pi) / s * 6), cos((2*pi) / s * 6))
1175   model$T[35:36, 35:36, 1] <- c(cos((2*pi) / s * 7), -sin((2*pi) / s * 7),
1176                                   sin((2*pi) / s * 7), cos((2*pi) / s * 7))
1177   model$T[37:38, 37:38, 1] <- c(cos((2*pi) / s * 8), -sin((2*pi) / s * 8),
1178                                   sin((2*pi) / s * 8), cos((2*pi) / s * 8))
1179   model$T[39:40, 39:40, 1] <- c(cos((2*pi) / s * 9), -sin((2*pi) / s * 9),
1180                                   sin((2*pi) / s * 9), cos((2*pi) / s * 9))
1181   model$T[41:42, 41:42, 1] <- c(cos((2*pi) / s * 10), -sin((2*pi) / s * 10),
1182                                   sin((2*pi) / s * 10), cos((2*pi) / s * 10))
1183   model$T[43:44, 43:44, 1] <- c(cos((2*pi) / s * 11), -sin((2*pi) / s * 11),
1184                                   sin((2*pi) / s * 11), cos((2*pi) / s * 11))
1185   model$T[45:46, 45:46, 1] <- c(cos((2*pi) / s * 12), -sin((2*pi) / s * 12),
1186                                   sin((2*pi) / s * 12), cos((2*pi) / s * 12))
1187   model$T[47:48, 47:48, 1] <- c(cos((2*pi) / s * 13), -sin((2*pi) / s * 13),
1188                                   sin((2*pi) / s * 13), cos((2*pi) / s * 13))
1189   model$T[49:50, 49:50, 1] <- c(cos((2*pi) / s * 14), -sin((2*pi) / s * 14),
1190                                   sin((2*pi) / s * 14), cos((2*pi) / s * 14))
1191   model$T[51:52, 51:52, 1] <- c(cos((2*pi) / s * 15), -sin((2*pi) / s * 15),
1192                                   sin((2*pi) / s * 15), cos((2*pi) / s * 15))
1193   model$T[53:54, 53:54, 1] <- c(cos((2*pi) / s * 16), -sin((2*pi) / s * 16),
1194                                   sin((2*pi) / s * 16), cos((2*pi) / s * 16))
1195   model$T[55:56, 55:56, 1] <- c(rho_co, -rho_si,
1196                                   rho_si, rho_co)
1197   model$P1inf[55, 55] <- model$P1inf[56, 56] <- 0
1198   model$P1[55, 55] <- model$P1[56, 56] <- vpsi
1199   model$H[1, 1, 1] <- exp(pars[6])
1200
1201 }
```

MODELS	
ARIMA	
UCM	
KNN	
LSTM RNN	
COMPARISON AND CONCLUSIONS	

UCM COMPARISON			
MODEL	COMPONENTS	TRAIN MAE	VAL MAE
1	LLT + DUMMY SEASONALITY (7)	0.073	0.255
2	IRW + CYCLE + DUMMY SEASONALITY (7)	0.062	0.273
3	IRW + CYCLE + TRIG DETERMINISTIC SEASONALITY (365)	0.106	0.349
4	LLT + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.064	0.309
5	IRW + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.080	1.906
6	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG DETERMINISTIC SEASONALITY (365)	0.063	0.322
7	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.063	0.323
8	LLT + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	1.9e-15	0.158

MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON AND CONCLUSIONS

K N N

MULTIPLE INPUT MULTIPLE OUTPUT

HYPERPARAMETERS

- Lags = 1:1500
- K = 2

EVALUATION

- MAE = 0.243

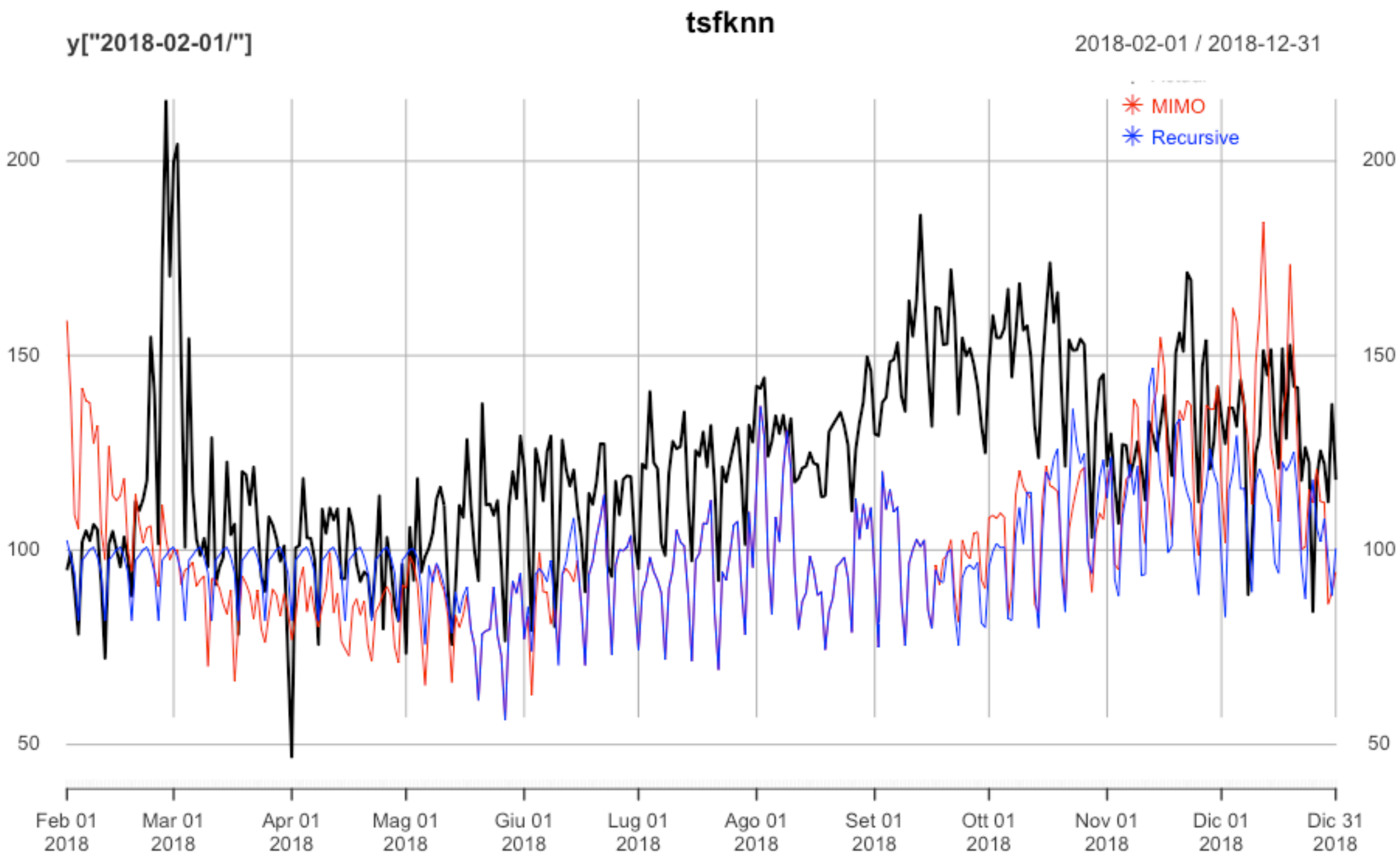
RECURSIVE

HYPERPARAMETERS

- Lags = 1:1500
- K = 2

EVALUATION

- MAE = 0.229



MODELS

ARIMA

UCM

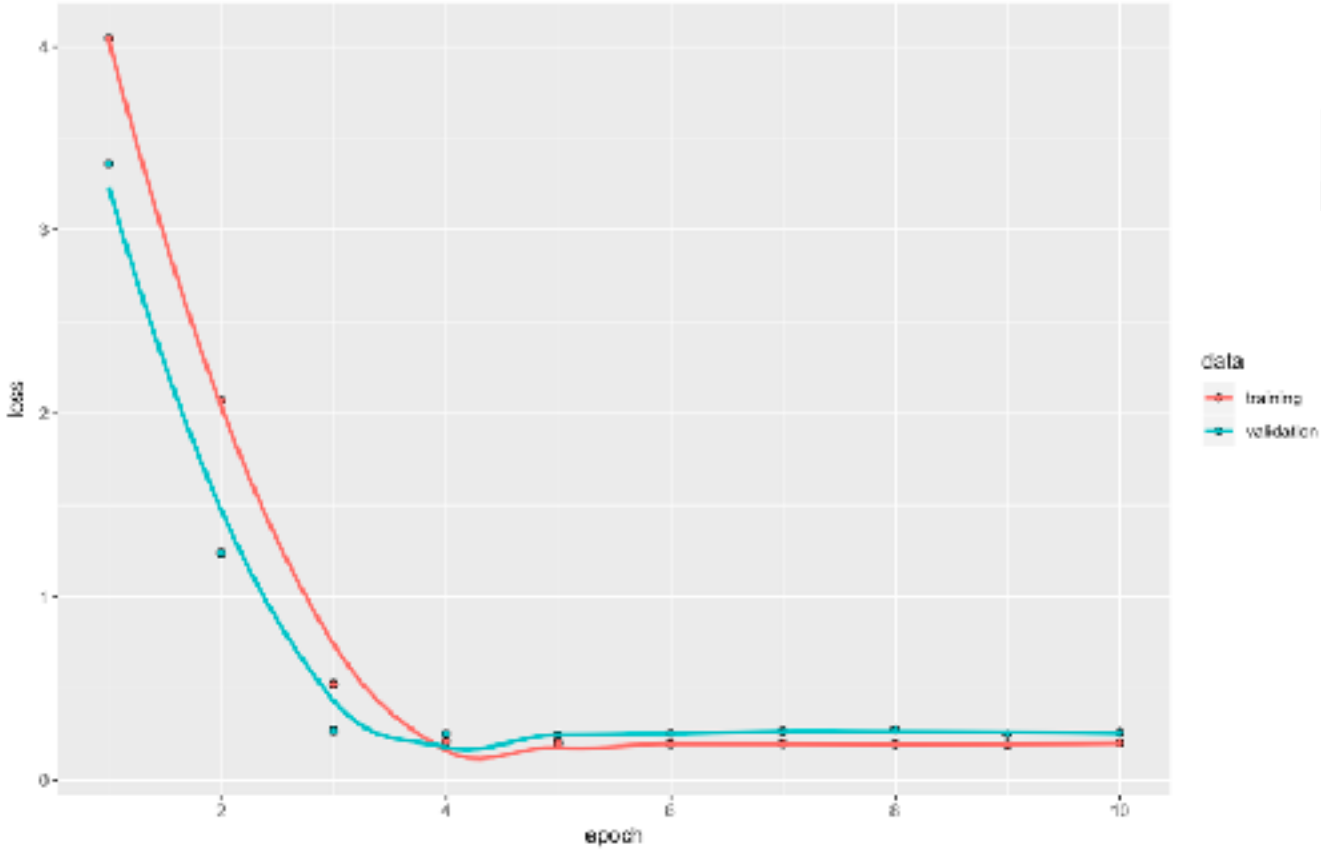
KNN

LSTM RNN

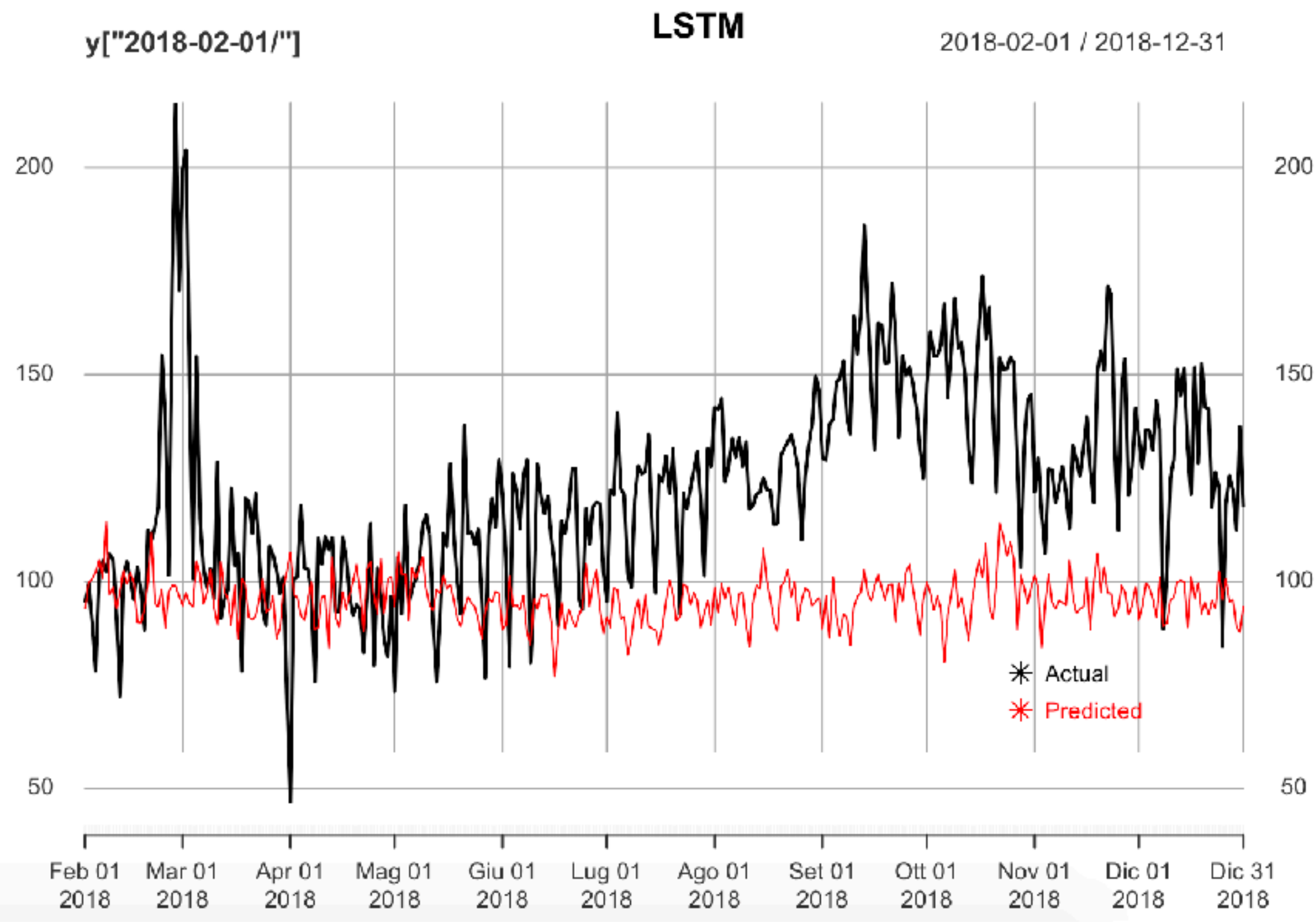
COMPARISON AND CONCLUSIONS

HYPERPARAMETERS

- Lookback = 1500
- Batch size = 128
- LSTM Layer = 1
- LSTM neurons = 32
- Dropout = 0.3
- Recurrent Dropout = 0.3
- Activation function = “tanh”
- Dense Layer = 1 (334 neurons
- Optimizer = “rmsprop£
- Loss = “mae”
- Steps per epoch = 50
- Epochs = 10



LSTM



EVALUATION

- MAE = 0.26

MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON AND CONCLUSIONS

COMPARISON			
MODEL	COMPONENTS	TRAIN MAE	VAL MAE
ARIMA	/	0.110	0.236
UCM 1	LLT + DUMMY SEASONALITY (7)	0.073	0.255
UCM 2	IRW + CYCLE + DUMMY SEASONALITY (7)	0.062	0.273
UCM 3	IRW + CYCLE + TRIG DETERMINISTIC SEASONALITY (365)	0.106	0.349
UCM 4	LLT + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.064	0.309
UCM 5	IRW + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.080	1.906
UCM 6	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG DETERMINISTIC SEASONALITY (365)	0.063	0.322
UCM 7	IRW + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	0.063	0.323
UCM 8	LLT + CYCLE + DUMMY SEASONALITY (7) + TRIG SEASONALITY (365)	1.9e-15	0.158
KNN MIMO	/	NA	0.243
KNN REC	/	NA	0.229
LSTM	/	0.2	0.26

MODELS

ARIMA

UCM

KNN

LSTM RNN

COMPARISON
AND
CONCLUSIONS

CONCLUSIONS

ADVANTAGES

- ARIMA captures the trend but less oscillations for many steps ahead
- UCM: the freedom of KFAS package allows to design specific models
- KNN: extremely fast to train
- RNN: design of ad-hoc models

ISSUES

- ARIMA: can't handle yearly seasonality for daily data
- UCM: identification of the best model
- KNN: distance metric
- RNN: training costs

NEXT STEPS

- ARIMA: handle yearly seasonality for daily data
- UCM: use of alternative models
- KNN: distance metric that weights more recent observations
- RNN: deeper network

THANKS FOR YOUR ATTENTION



Raffaele Anselmo - 846842