Project in Time Series Analysis

> Life can only be understood backwards,
> but it must be lived forwards.

– Søren Kierkegaard (1844)

Arvid Gramer & Nils Romanus

How to predict the temperature in Svedala?

Model Architectures

Complexity

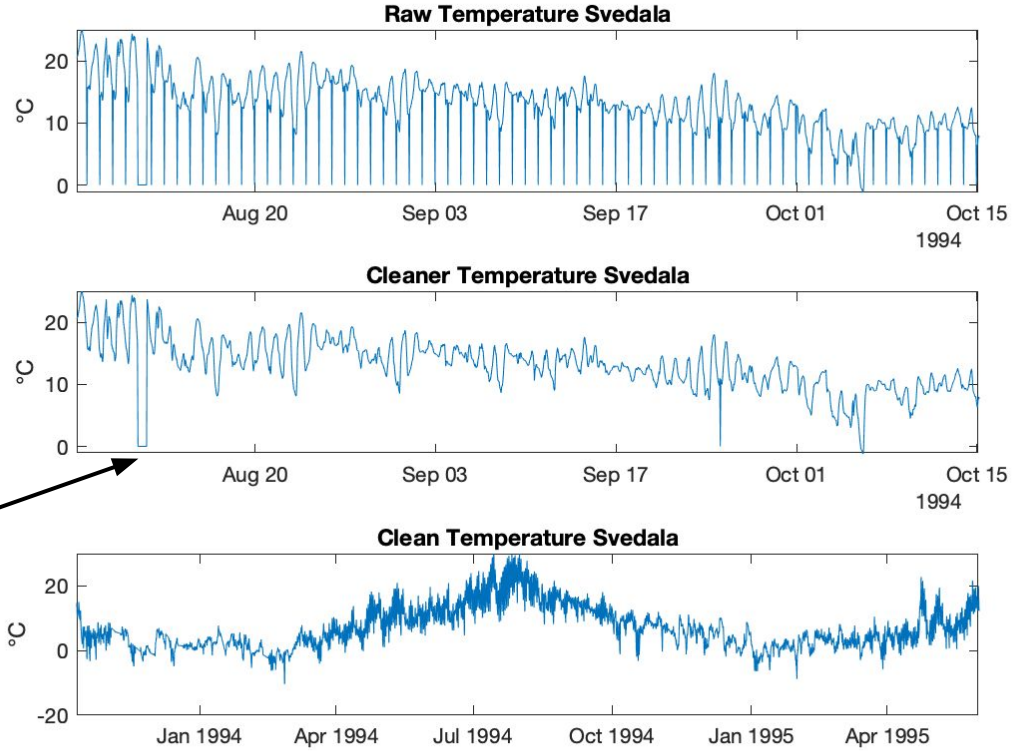
a. Naive Model
b. SARIMA
c. SARIMAX
d. Recursive Model
e. Prophet

SVEDALA KOMMUN

Real world data is messy

- Zeros at 23:00
- Strange zeros

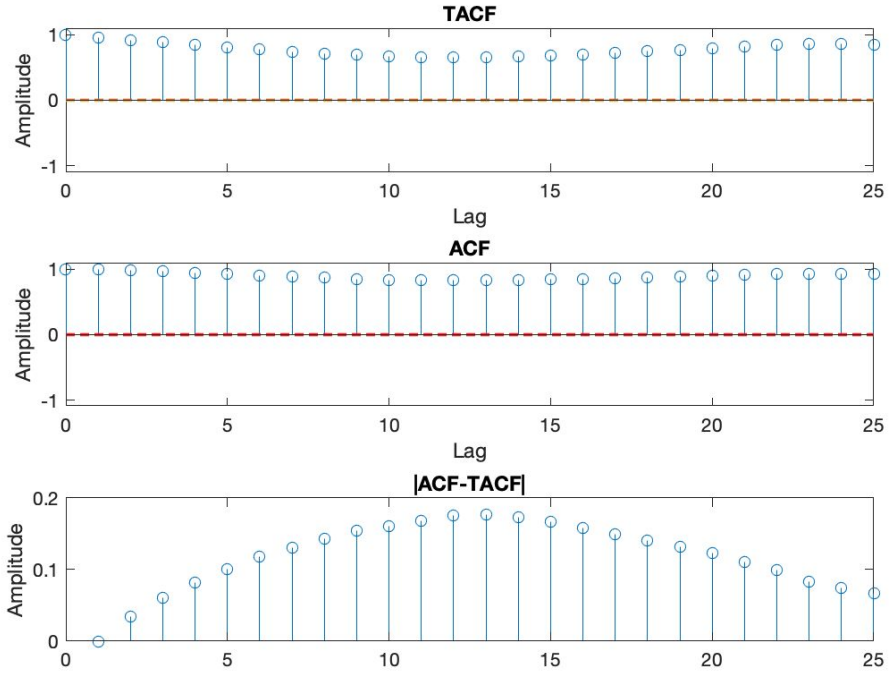
Is the temperature really 0 degrees during noon in August?



Real world data is messy

Handling Outliers

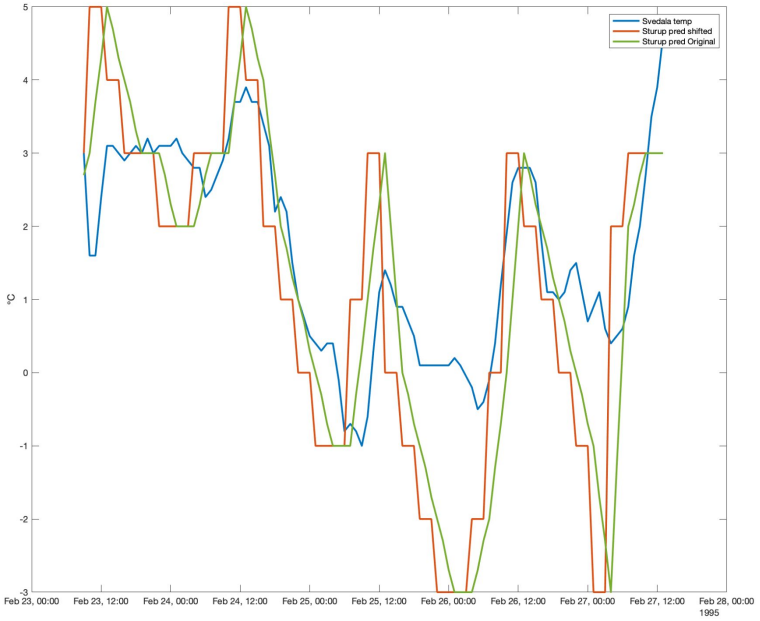
- Do we have any outliers?
- Do they impact the structure of our data?



Real world data is messy

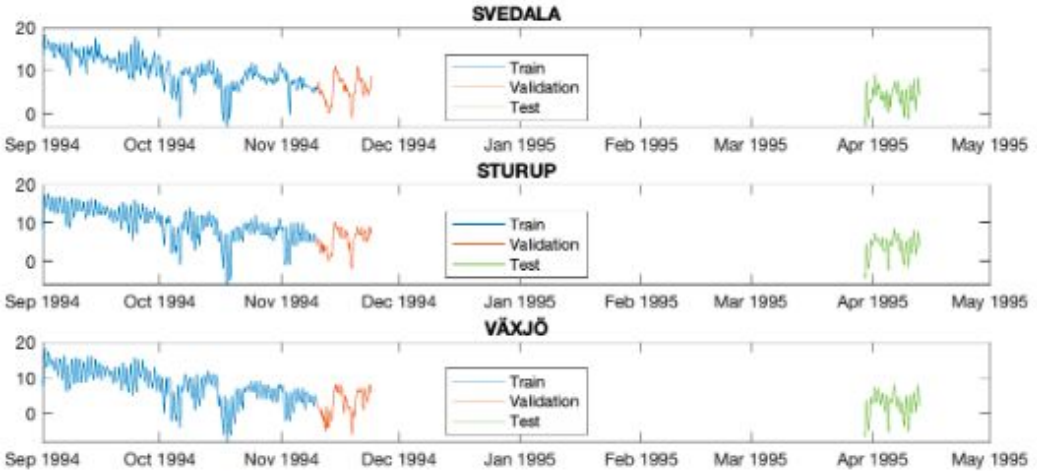
Shifting exogenous data

- Predictions are available three hour in advance.
- Interpolations would include future values.

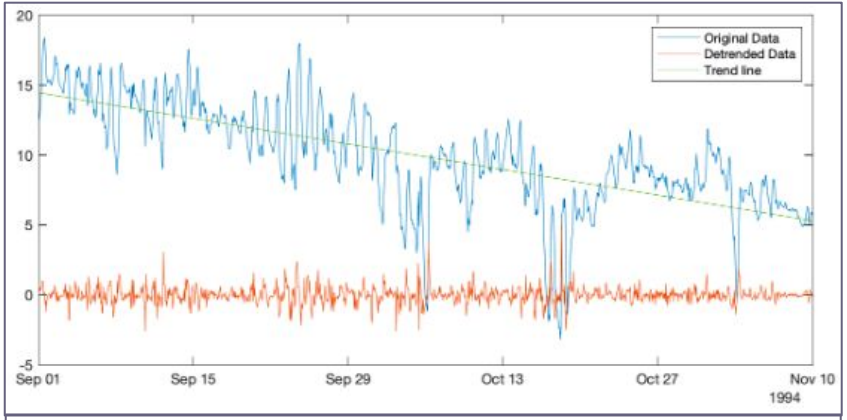


Split the data

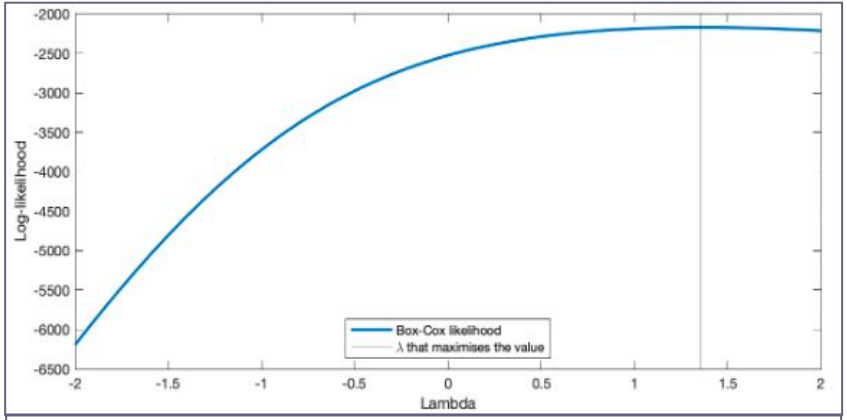
- Modeling
- Validation
- Test



Trends & Transformations

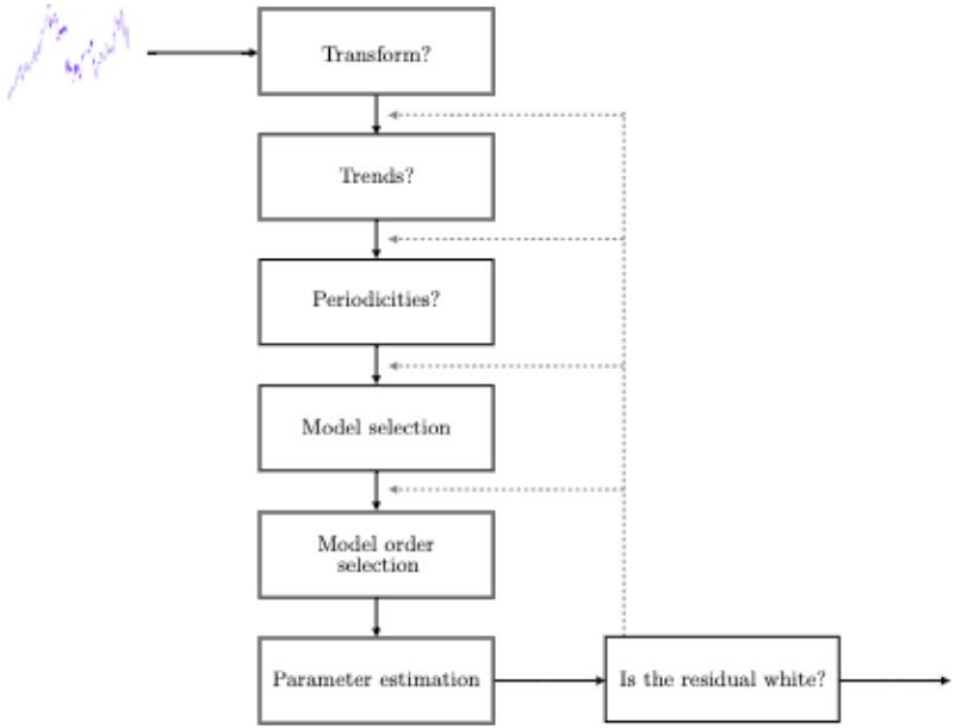


Data needs to be detrended



No need for power transforms

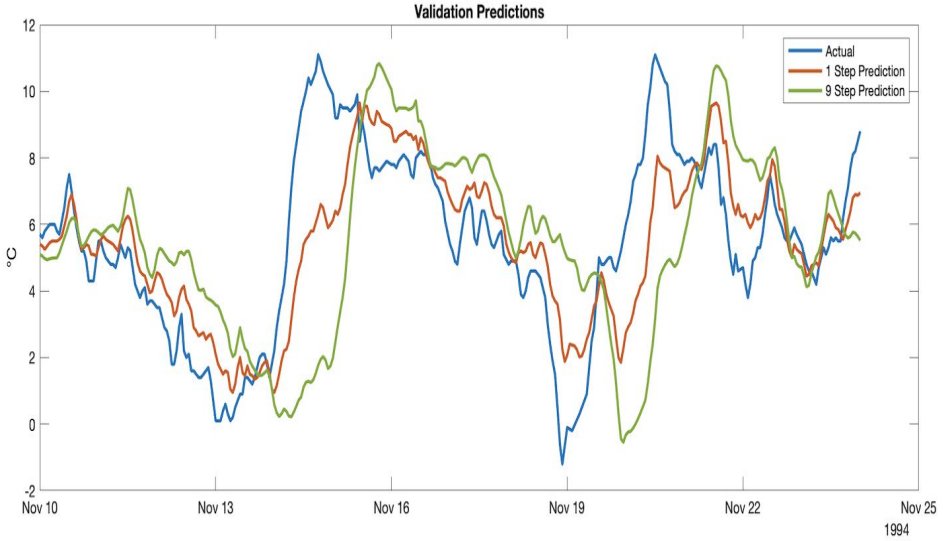
Use the TSA workflow!



Naïve predictor

- A fair guess, but not estimated using data
- Something in between the current temperature and the one yesterday, namely:

$$y_t = \frac{y_{t-1} + y_{t-24}}{2}$$

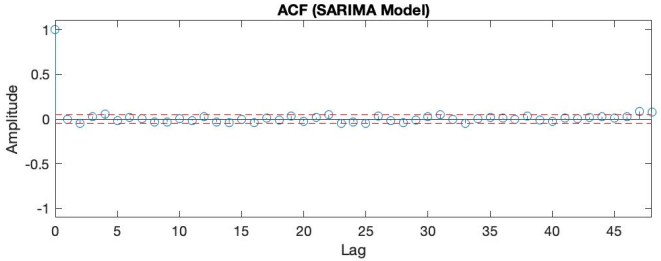


Data Set	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$	σ^2_y
Modeling	1.5267	5.3761	13.5078
Validation	3.1007	11.9268	7.5529

Variance of prediction residuals

Model without external input

- SARIMA
- Flexible Seasonality
- KISS



	a_1	a_2	a_{23}	a_{25}
A(z)	$-1.451(\pm 0.02)$	$0.5246(\pm 0.02)$	$-0.1295(\pm 0.01)$	$0.103(\pm 0.01)$

	c_1	c_{24}
C(z)	$-0.9661(\pm 0.01)$	$-0.02175(\pm 0.01)$

Table 4: Model A: $\nabla_1 A(z) y_t = C(z) e_t$

WOW! That is **white** as snow!

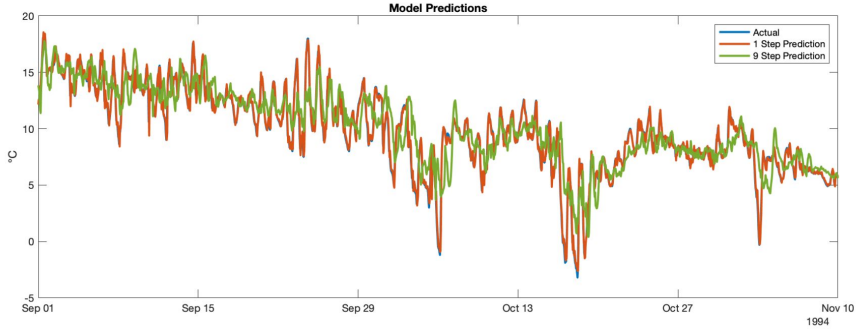
Model without external input

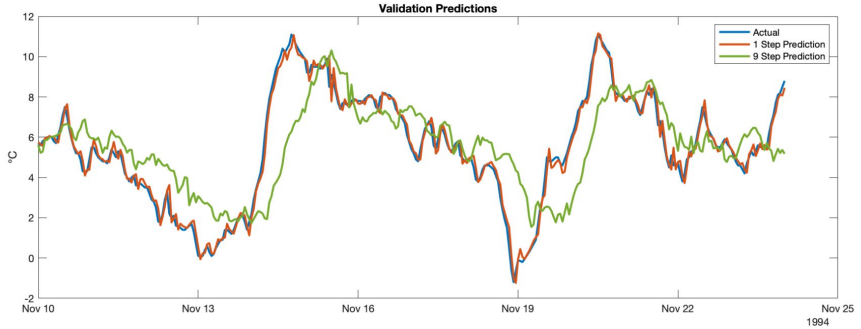
- Better than Naive
- No white prediction residuals

Data Set	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$	σ^2_y
Modeling	0.20722	3.6958	13.5078
Validation	0.12046	4.0546	7.5529

Data Set	$\epsilon_{t+1\|t}$	$\epsilon_{t+9\|t}$
Modeling	$40.08 \not< 36.42$	$2272.73 \not< 36.42$
Validation	$88.58 \not< 36.42$	$409.66 \not< 36.42$

Variance and whiteness of temperature prediction residuals

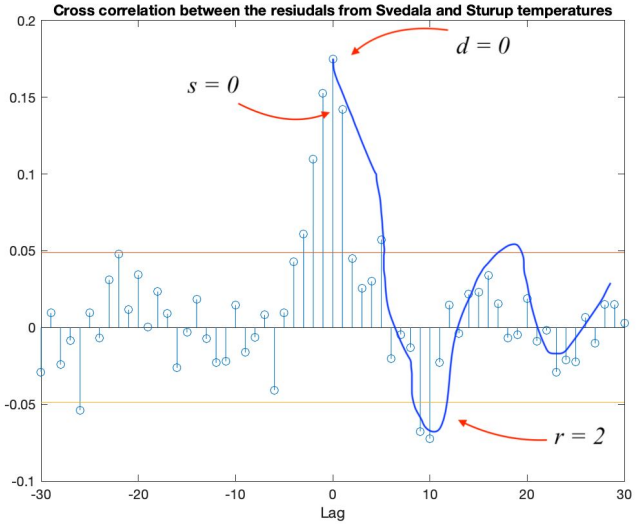




Predictions on validation set using model A

Lets use external input

- BJ / SARIMAX
- Use Sturup as external signal
- White modelling residual



Use BJ Scheme from book to get suitable model orders

Model with external input

- Worse than A on validation data
- White prediction residuals!
- Poor prediction of external signal

Svedala Predictions

Data Set	$\epsilon_{t+1\|t}$	$\epsilon_{t+9\|t}$
Modeling	$34.23 < 36.42$	$2312.21 \not< 36.42$
Validation	$33.13 < 36.42$	$403.87 \not< 36.42$

Data Set	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$	σ^2_y
Modeling	0.19962	3.5773	13.5078
Validation	0.10887	4.7321	7.5529

Variance and whiteness of temperature prediction residuals

Sturup (external) Predictions

Data Set	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$	σ^2_y
Modeling	1.0654	4.658	13.5078
Validation	0.51458	5.4261	7.5529

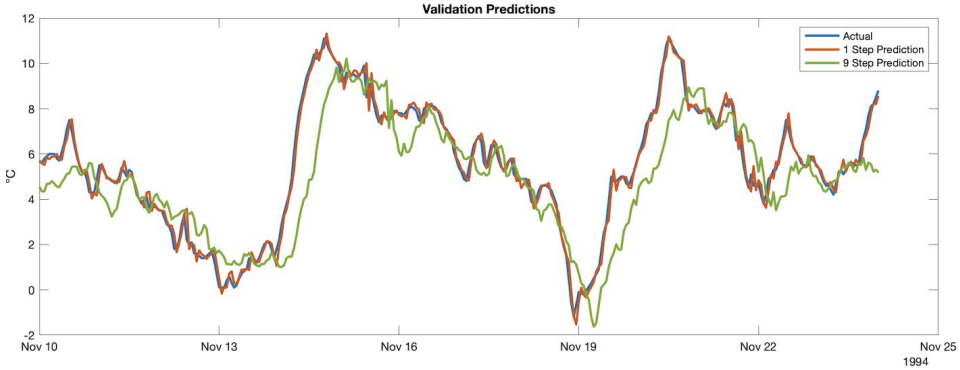
Residual variance of input signal prediction

Model using recursive method

Data Set	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$	σ^2_y
Modeling	0.19985	2.5626	13.5078
Validation	0.10701	2.1477	7.5529

Data Set	$\epsilon_{t+1\|t}$	$\epsilon_{t+9\|t}$
Modeling	$28.05 < 36.42$	$2125.63 \not< 36.42$
Validation	$26.70 < 36.42$	$387.93 \not< 36.42$

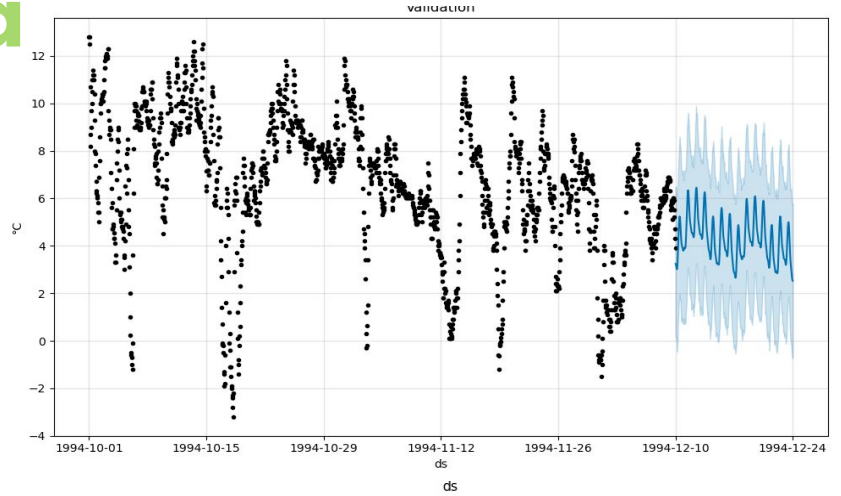
- Kalman filtering renders the best results
- White prediction residuals!

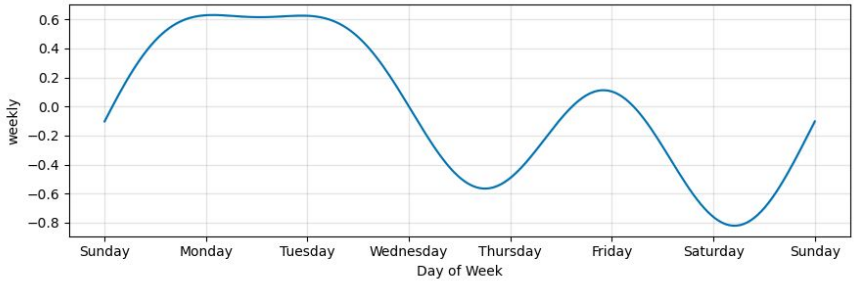


Temperature, one- and nine step predictions on validation set using recursive model

Prophet is bad at predicting the temperature in Svedala

- Worse than Naive Predictor
- Fits erroneous pattern
- Temperature does not care if it is Saturday or Monday





Of course it is warmer on mondays and tuesdays!

Börsöversikt > Meta

115,78 USD

−209,35 (−64,33 %) ↓ senaste året

19 dec. 11:42 GMT−5 • Ansvarsfriskrivning

PROPHET

Forecasting at scale.

FORECAST THIS!

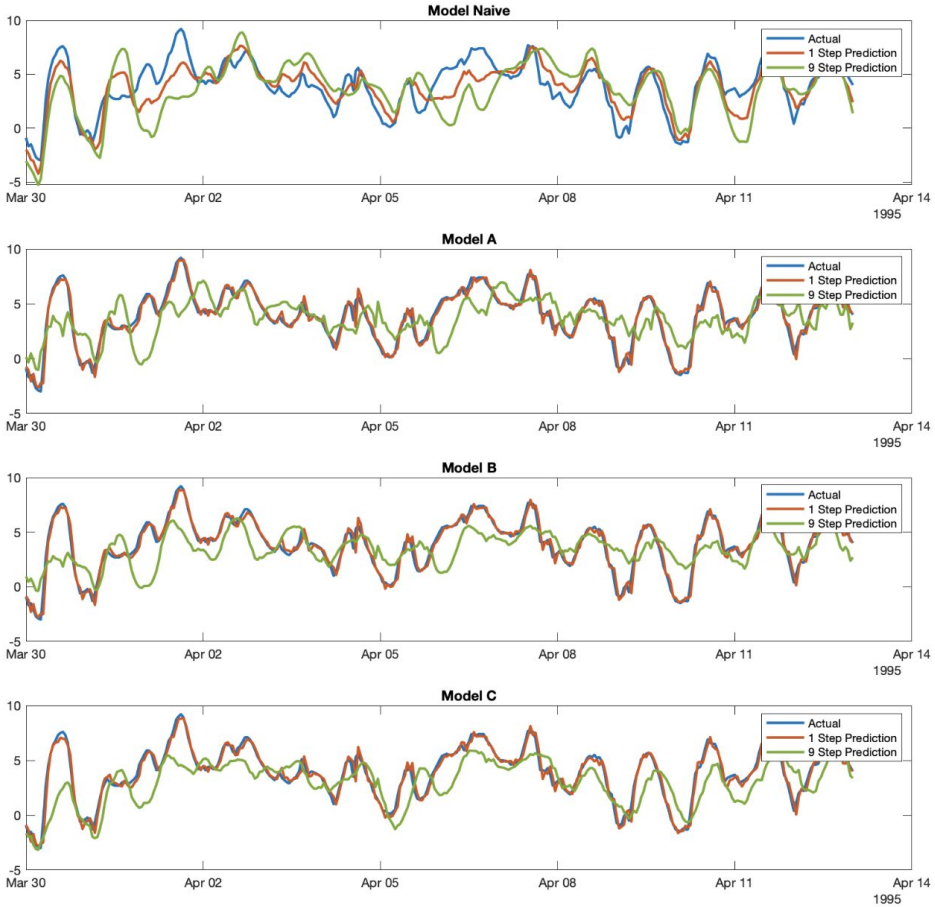
1d | 5d | 1mån | 6mån | I år | 1år | 5år | Högst

Testing

- Models perform in reverse complexity order

Model	$\sigma^2_{t+1\|t}$	$\sigma^2_{t+9\|t}$
Naive	1.8144	6.3152
A	0.23977	4.5599
B	0.23098	3.946
C	0.21926	3.0591
σ^2_y	5.8738	-

Residual variance on test set



Temperature, one- and nine-step predictions on test set using all models

Q&A