Time Series Prediction of the Tide Gauge in St. Pauli

- 08.01.2024 -

Janika Rhein, Ludwig Meder and Hannes Sandberg

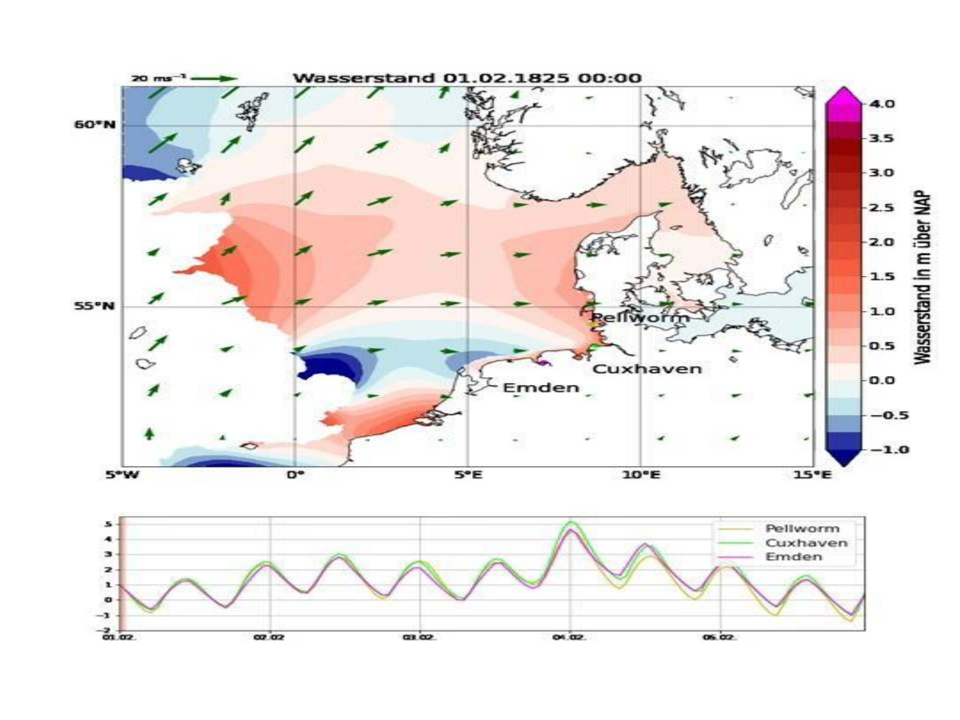
Introduction

- River Elbe is affected by tides
- Mean high water St. Pauli: 2.13 m > NHN
- Flooded parking lot: 2.90 m > NHN
- Flooded fish auction house: 3.50 m > NHN
- At German Bay,
  water level mainly depends on
  - astronomical tides
  - wind speed
  - wind direction

NHN - Normal Height Null

https://commons.wikimedia.org/wiki/File:Sturmflut_Fischmarkt_Hamburg.jpg

Introduction - Visualization of historical storm surge



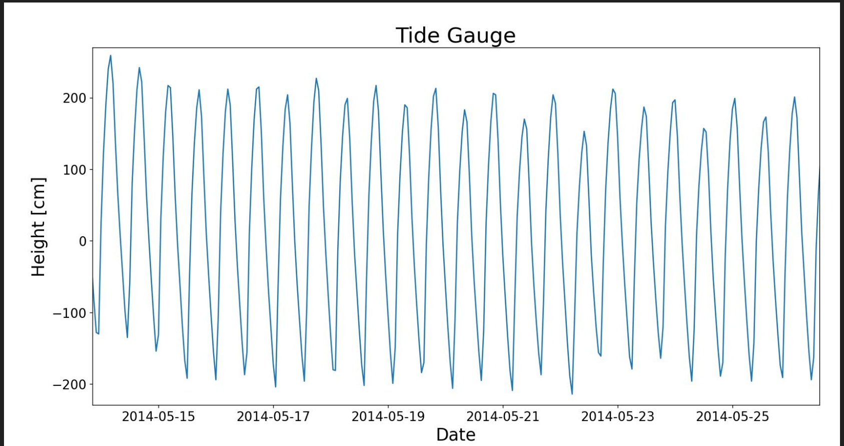
- tidal wave travels counterclockwise through North Sea
- wind measurements from Helgoland are representative for German Bay

Literature Review

- Descriptions of time series forecasts architectures by Tensor Flow
- Paper:
    - Long Short-Term Memory Recurrent Neural Network for Tidal Level Forecasting by Yang et al. (2020)
    - Compares LSTM model with 6 different models
    - tidal water level of 21 years
    - LSTM had best performance with RMSE = 4.9cm for forecasting period of 30 days

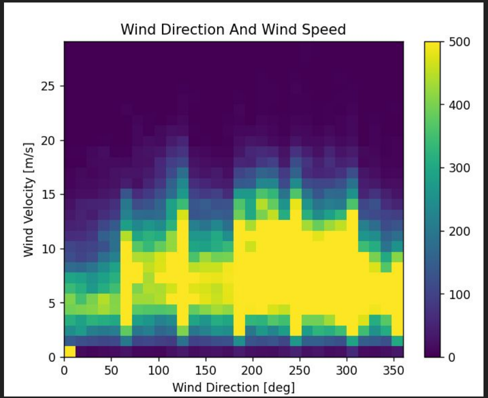
RMSE = root mean square error

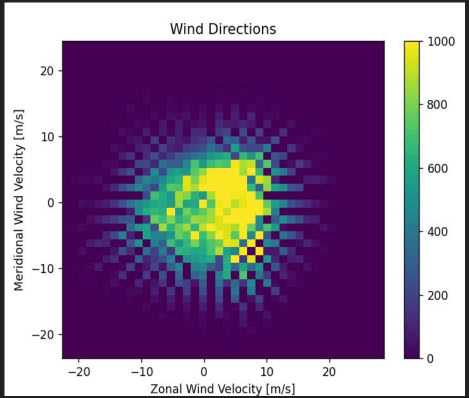
Dataset characteristics



Water Level Data

- **Data Availability**: Continuous minutely data
- **Commencement**: Since 1999
- **Accessibility**: Freely accessible at hydroonline.hpanet.de





Wind Data

- **Data Availability**: Continuous hourly data
- **Commencement**: Since 1959
- **Accessibility**: Freely accessible at opendata.dwd.de

Dataset characteristics

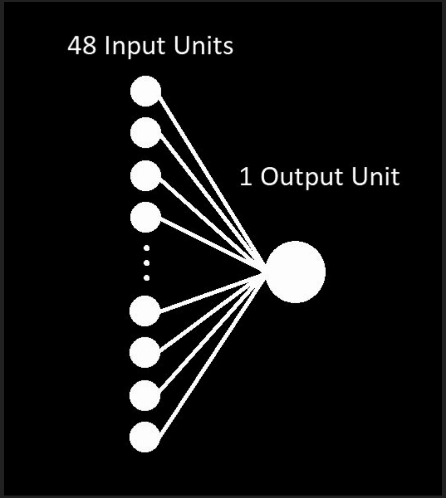
Preprocessing:

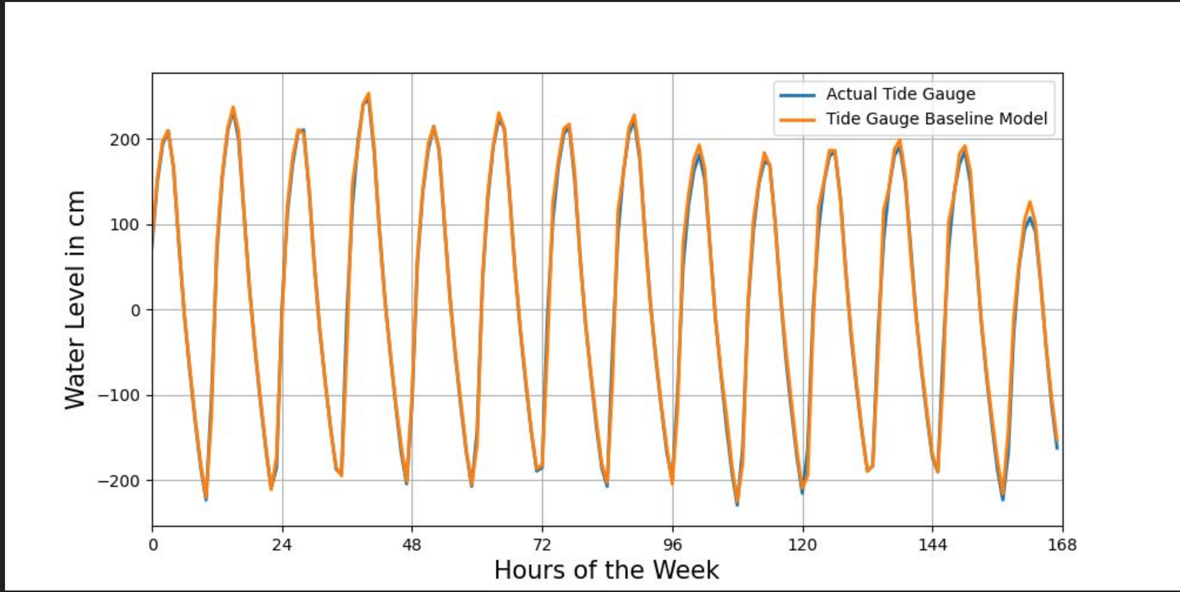
- Reduction of tide resolution from minutely to hourly
- Cleaning of extreme outliers
- Gradient cleaning to handle gradient outliers
- Linear interpolation of gaps smaller than 6 hours
- Vectorization of wind data
- Combining data into continuous data frames

Result:

	water_level [cm]	wind_v [m/s]	wind_u [m/s]	ds_number
0	86.0	2.187967e+00	-12.408578	0
1	203.0	3.796424e+00	-10.430588	0
2	281.0	1.788576e+00	-10.143520	0
3	322.0	1.788576e+00	-10.143520	0
4	328.0	-1.726752e-15	-9.400000	0
...	...	...	...	...
196515	42.6	-5.335137e+00	-6.358169	86
196516	130.6	-3.850000e+00	-6.668396	86
196517	185.6	-5.515520e+00	-4.628071	86
196518	224.6	-6.894400e+00	-5.785088	86
196519	235.6	-6.684991e+00	-7.966862	86

Baseline Model





RMSE_1hour = 12.32 cm
RMSE_6hour = >20 cm

Model Definition and Evaluation

Hyperparameter used for all models:

- optimizer: Adam
- loss function: mean squared error
- dense layer activation function: ReLu
- Input: 144 units consisting of 48 time steps
  - tidal water level
  - wind speed u
  - wind speed v
- Output: 6 units (time steps) tidal water level
  activation function: linear

Model Definition and Evaluation

Model	Special Architecture	Epochs	Time per Epoch [s]	RMSE [cm]
Simple NN	Input: Dense Layer: 64 units, activation relu Output: Dense Layer 6 units, activation = linear	20	ca. 19	16.9
NN with hidden layers	2 hidden Dense layers with 128 and 256 units	20	ca. 14	15.0
Complex NN	3 hidden Dense layers with 128, 256 and 512 units	60		14.6
Complex NN with dropout	4 dropouts with 10, 15, 20 and 20%	60	ca. 39	19.5
Simple LSTM	Input: LSTM layer: 64 units Output: Dense layer: 6 units	20	ca. 99	15.1
More complex LSTM	LSTM layer: 128 units 2 hidden Dense layers with 64 and 128 units	60	ca. 9.5min	8.4
Convolution	Conv layer in front of more complex LSTM	60	ca. 8min	8.8

NN - Neural Network

LSTM - Long Short Term Memory

RMSE - Root Mean Square Error

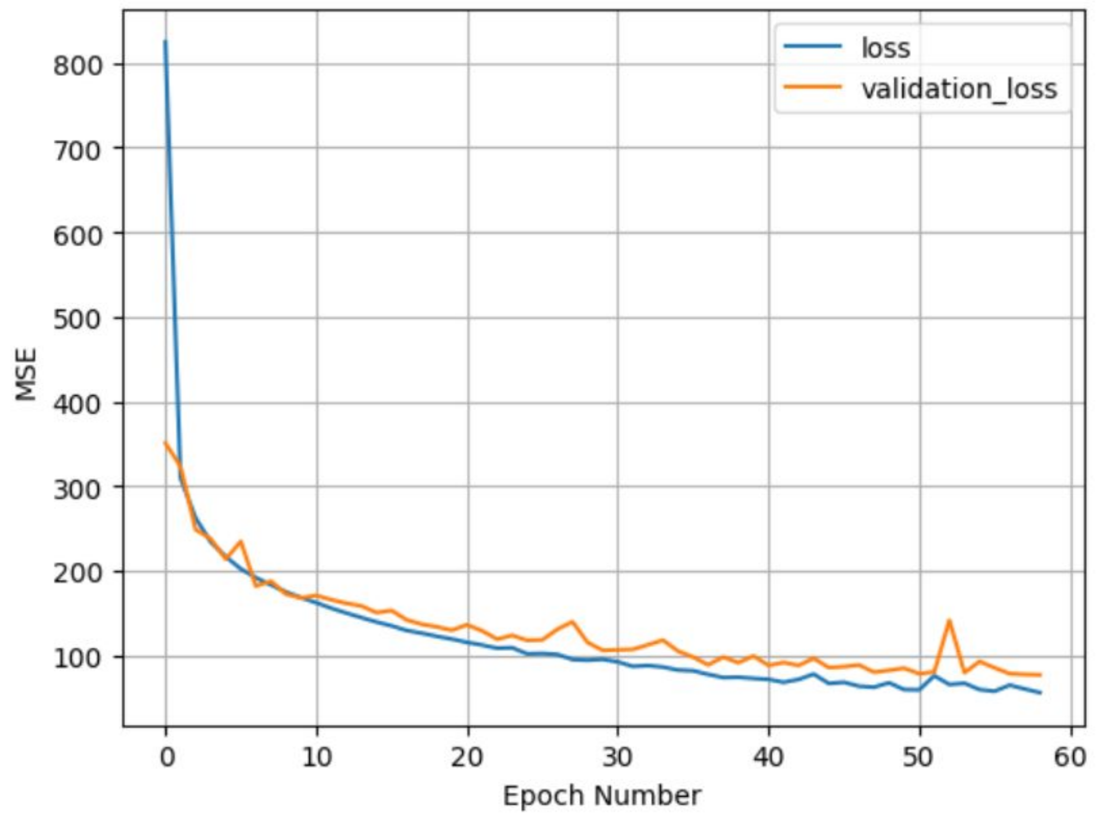
Model Definition and Evaluation

```
lstm_v2 = tf.keras.models.Sequential([
    tf.keras.layers.Conv1D(filters=32, kernel_size=6,
                           strides=1, padding="causal",
                           activation="relu",
                           input_shape=[input_length,features]),
    tf.keras.layers.LSTM(units = 64, return_sequences = True),
    tf.keras.layers.LSTM(units = 128),
    tf.keras.layers.Dense(units = 64),
    tf.keras.layers.Dense(units = 128),
    tf.keras.layers.Dense(units = output_length)
])
```

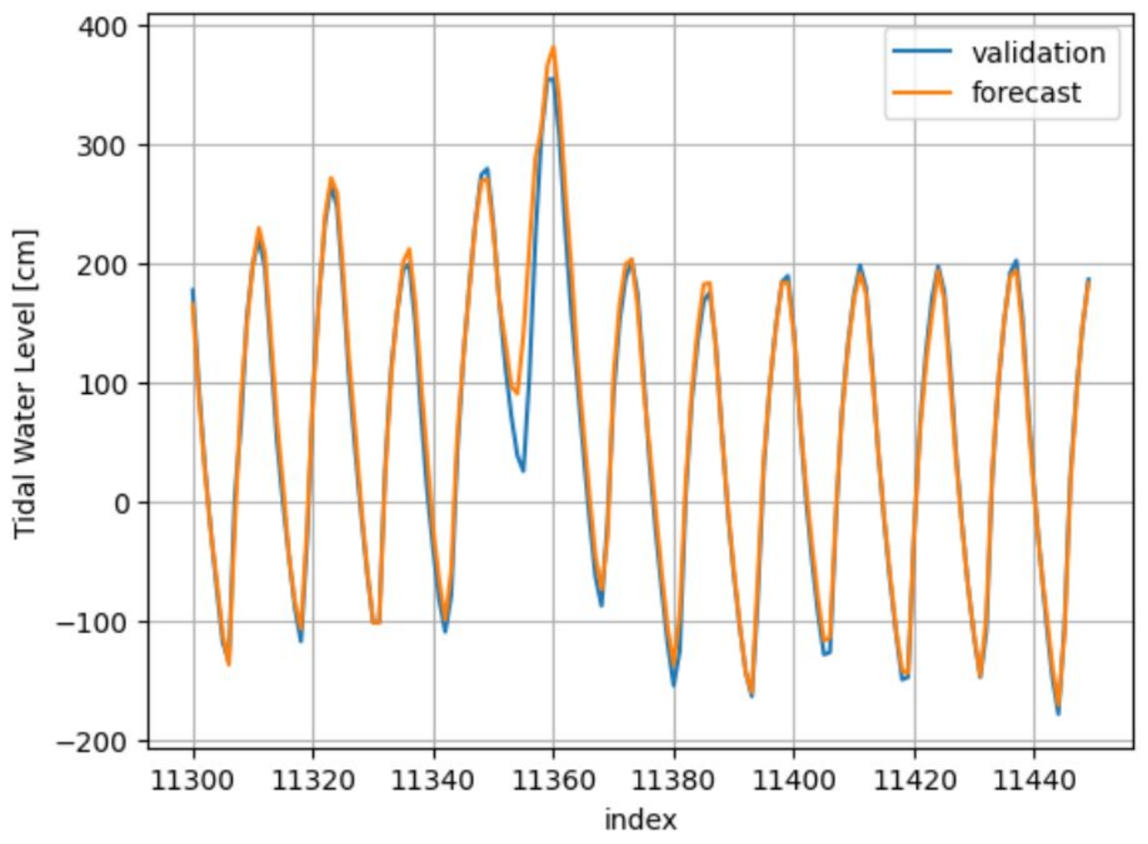
Trainable/Total parameters: 141606

Model Definition and Evaluation

RMSE = 8.8 cm



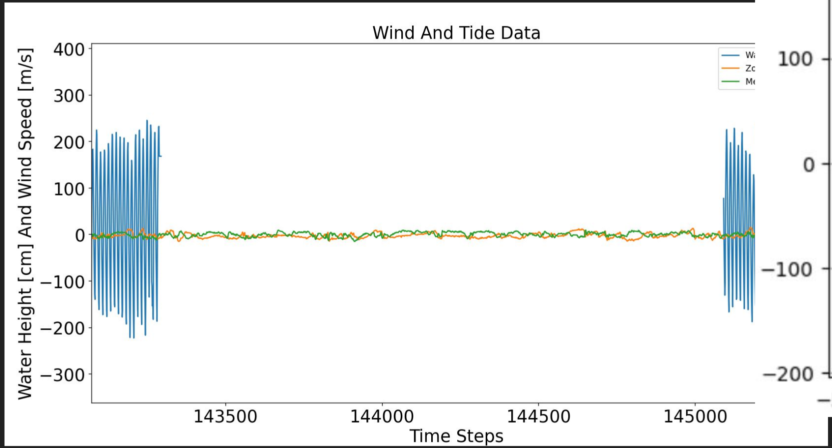
Results

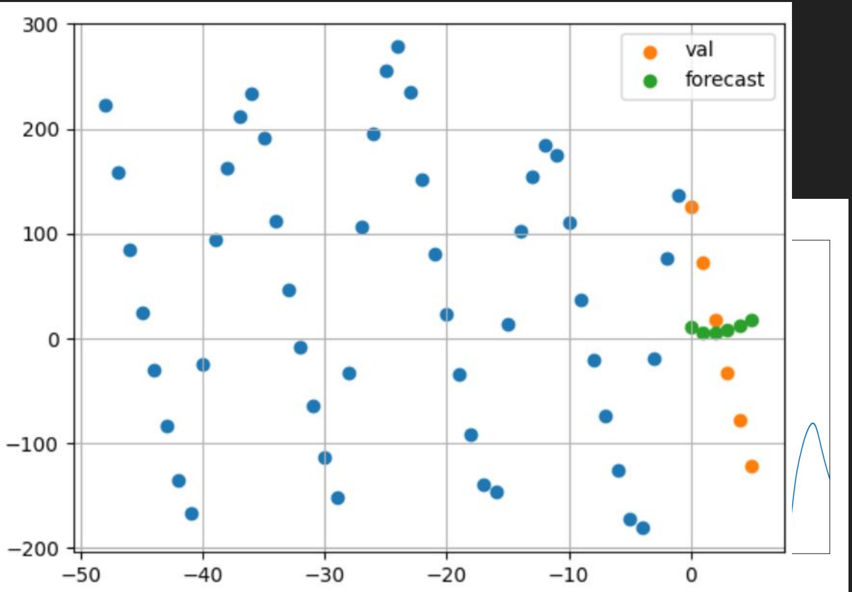


validation
forecast
Tidal Water Level [cm]
400
300
200
100
0
−100
−200
11300
11320
11340
11360
11380
11400
11420
11440
index

Challenges and Errors

1. Preprocessing dataset
2. Storm events - Undersampling: 2408 windows instead of 192202
   RMSE: 99.4cm!
3. Training time





Discussion

- Overall RMSE = 8 cm
  difference of mean flood water level to overflooding: 77 cm
- Possible improvements - considering:
  - more wind locations (Shetland Islands)
  - precipitation and melting snow of river’s drainage area
- Improving undersampling:
  - more overflooding events
  - decrease threshold of overflooding from 3.50m to 2.90m

Conclusion and Future Work

- Considering another wind location: Shetland Islands (German Bay boundary)
- Using 6h wind forecast as additional input parameters
- Improving model architecture
- Undersampling: Lower threshold

References

Animation:
https://blogs.helmholtz.de/kuestenforschung/2023/07/28/erstmalige-rekonstruktion-der-sturmflut-vom-februar-1825/ (last access: 07.01.2024)

Paper: Yang, Cheng-Hong, Chih-Hsien Wu, and Chih-Min Hsieh. "Long short-term memory recurrent neural network for tidal level forecasting." IEEE Access 8 (2020): 159389-159401.

Other sources:
Federal Maritime and Hydrographic Agency (BSH)
German Weather Service (DWD)

Questions?