TFT

Temporal Fusion Transformer

Jonas Koch, Malte Lehna, Leon Franken

Kontakt: Jonas Koch | +49 561 7294-1756 | Jonas.koch@iee.fraunhofer.de

Kognitive Energiesysteme

Temporal Fusion Transformer (TFT) were investigated to be used in time-series forecasting for several use cases in the energy context within the Competence Center for Cognitive Energy Systems. In this case, the usability for probabilistic wind power forecasts was investigated.

The focus was placed on the simultaneous training and prediction of several wind parks within one model in order to make spatio-temporal dependencies available to the model and to evaluate the performance.

TFT (Temporal Fusion Transformer)

- Based on the principles of Transformer Networks and optimized for time-series forecasting
- Additionally to NWP forecasts, **known future information** such as day-of-week, time-of-day can be incorporated into the forecast -> temporal patterns can be better identified
- Easy to group/tag information belonging to one location while training several
- In addition to the heterogeneous utilization of input data, the TFT also offers the possibility of
  - multiple-horizon forecasts and
  - probabilistic forecasts.
- Attention can help to understand the model better (XAI)

Model Inputs

- Wind power generation of one or several wind parks
- NWP parameter (wind-speed and direction, temperature, irradiance, …)
- Lags and leads of parameters
- Known future information
  - Categorical variables (time-of-day, day-of-week, …)

Feature importance

- Helps to explain the model and analyze the problem
- Helpful to reduce the number of input features
  - Required training period reduced

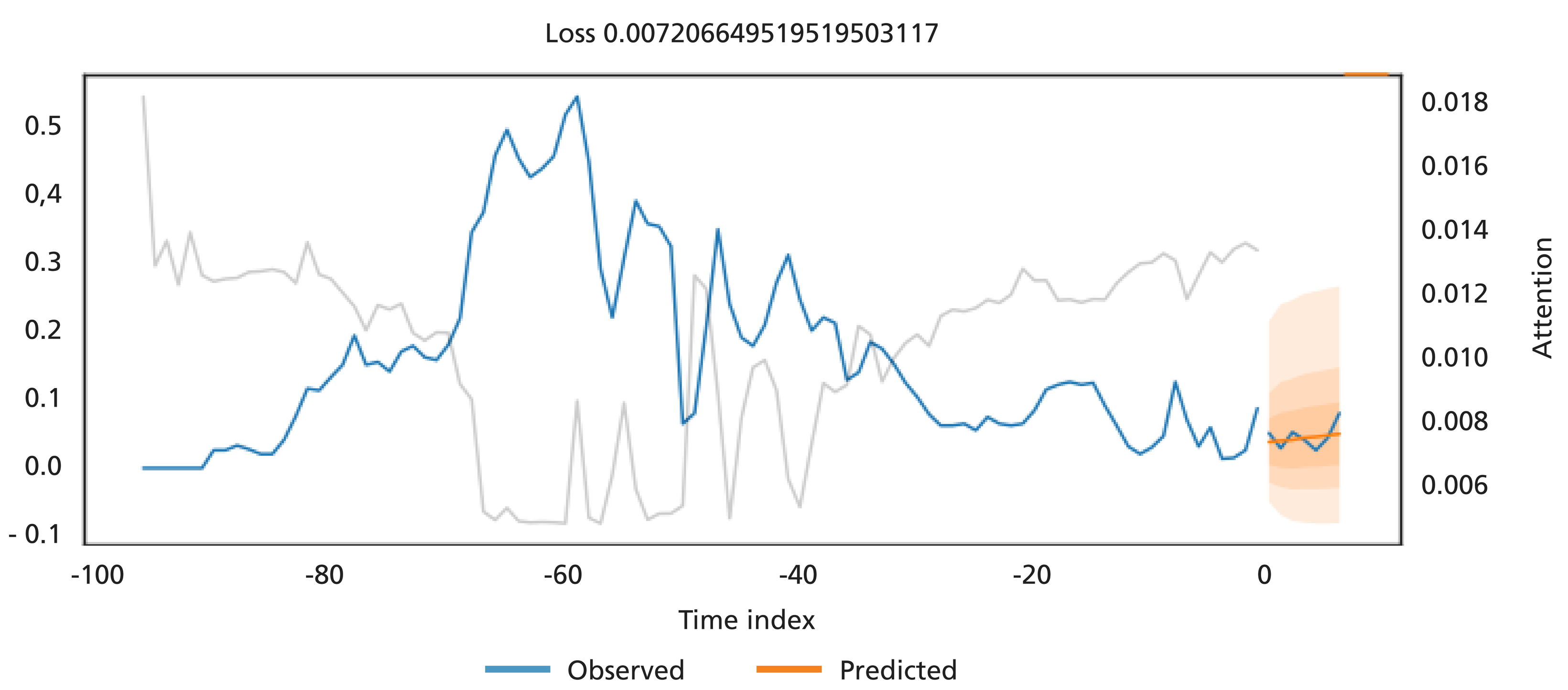

Figure 1: Attention plot showing the time-dependent importance of the variables (here for one time-step. Measurements in blue and attention in grey.

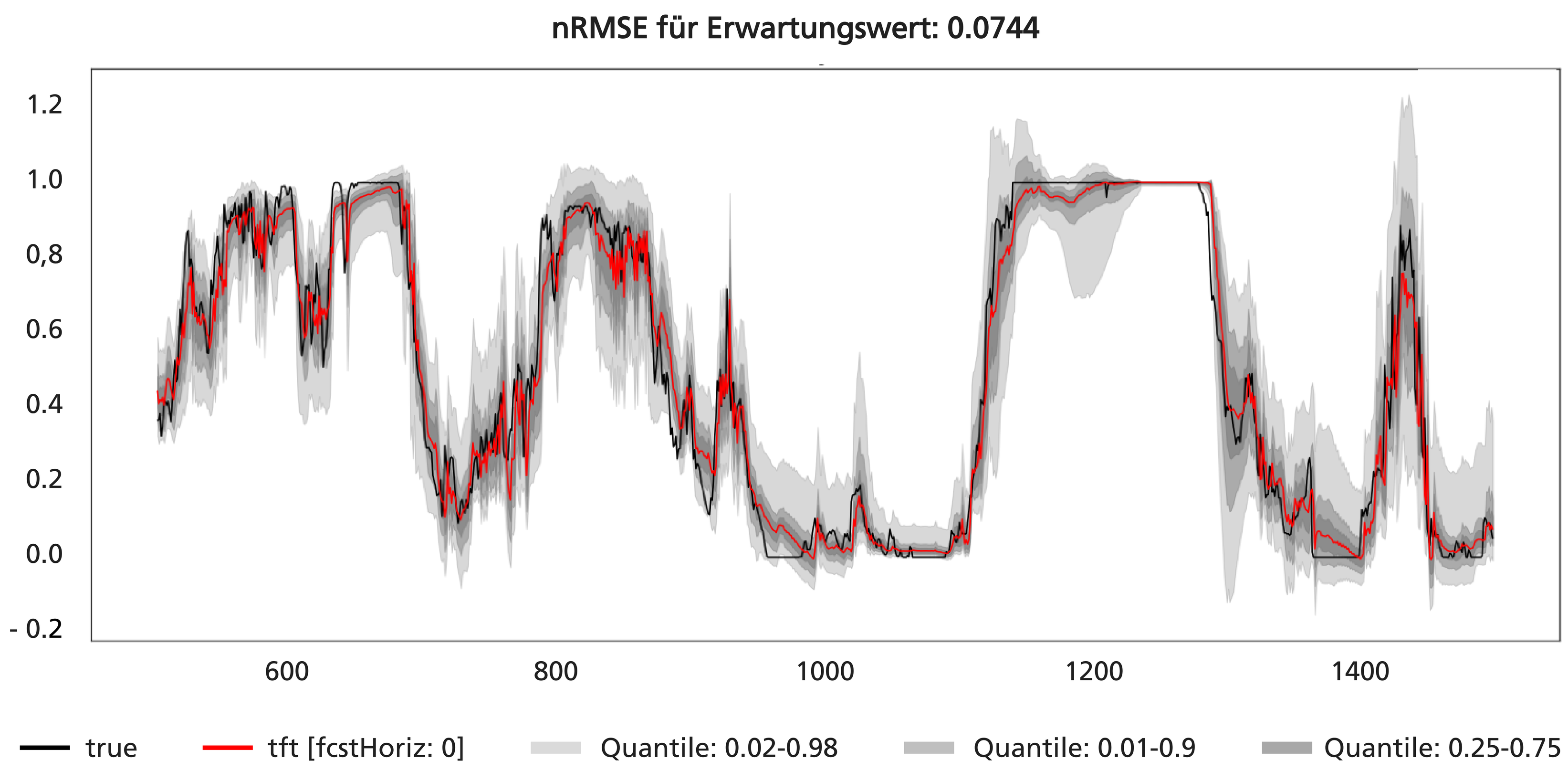

Figure 2: Probabilistic forecast example for one wind park

Results

Performance

- Outperforms the operative model for very short-term horizons
- Forecasts not "pulled towards the mean"
- Decreased performance for longer horizons
- Probabilistic forecast reflects uncertainties very well
- Profits if several parks are used for the model

Additional findings

- Multi-horizon forecasting trade-off between model reduction and performance
- Training time of TFT magnitudes longer than e.g. ELM
- TFT needs a lot of training data
- Without further steps (transfer-learning) not yet scalable for a huge amount of parks

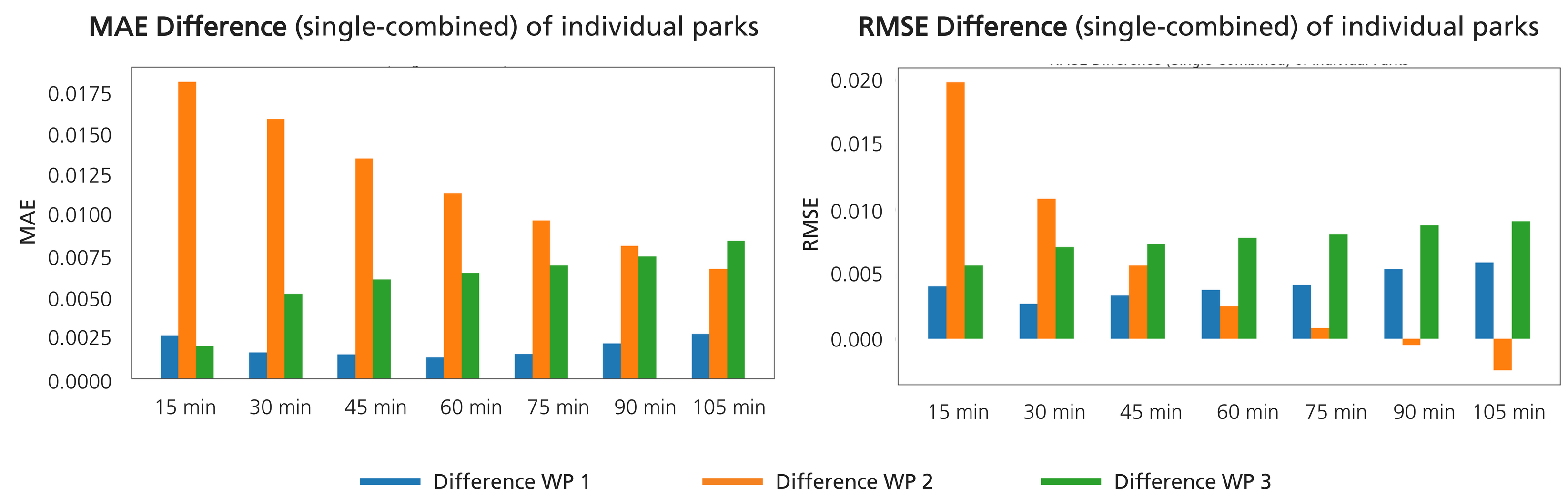

Figure 3: Decreasing performance gain with offsite information for three wind parks over the forecast horizons

Conclusion

The results show that the TFT is a method capable of generating well-performing forecasts. Especially the gain in the usage of offsite information is considerable but depends on the specific park and decreases quickly with the forecast horizon. The caveat of the model is its complexity and in comparison to methods like ELM considerably longer training times.

Supported by:

HESSEN
Hessisches Ministerium für Wissenschaft und Kunst

SPOTLIGHTS

Fraunhofer IEE