Forecasting Models with LamaH dataset

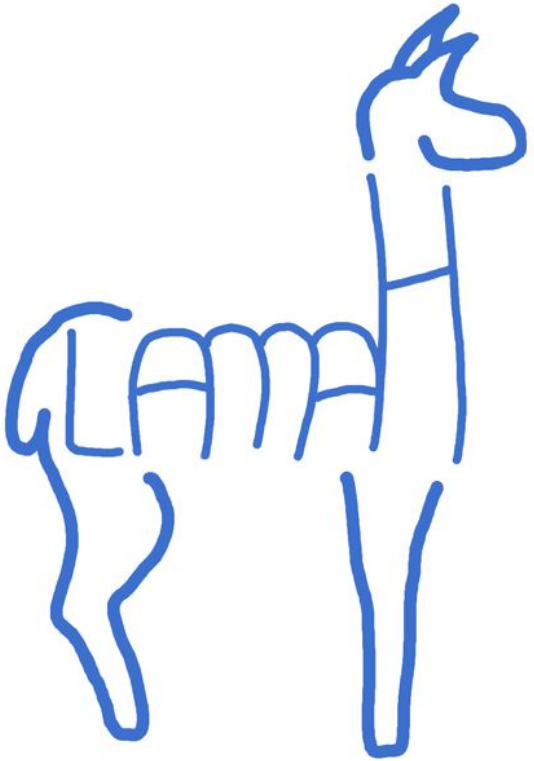
AI/ML in the Era of Climate Change

Jonas Unruh 12331457
Leon Beccard 12133103

Content

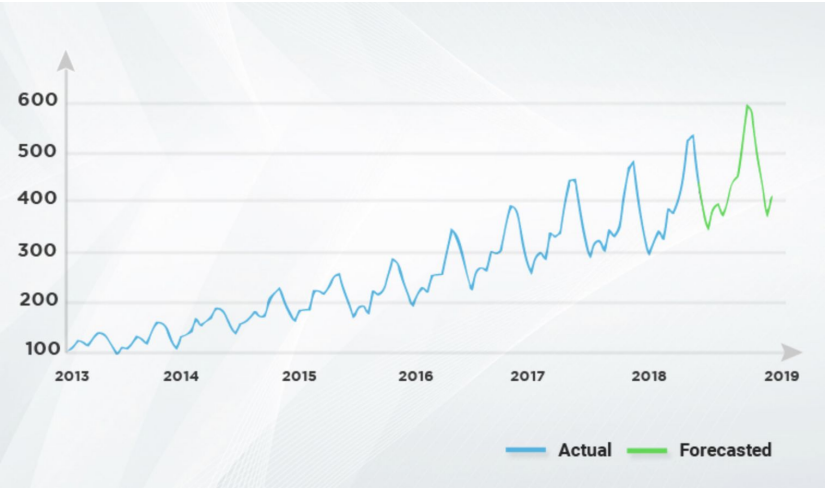
Introduction - LamaH data

- Hydrological data
- Gathered on rivers
- Covers 170.000 km² in Central Europe
- Wide variety of values
- Wide variety of geographical locations

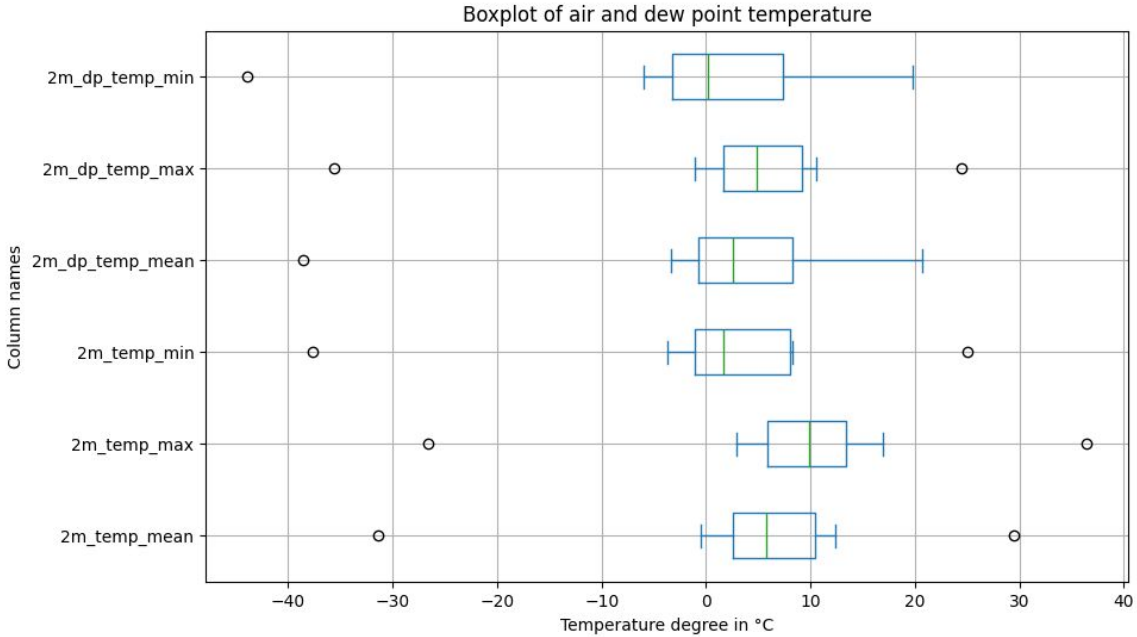


Introduction - Time Series Modeling

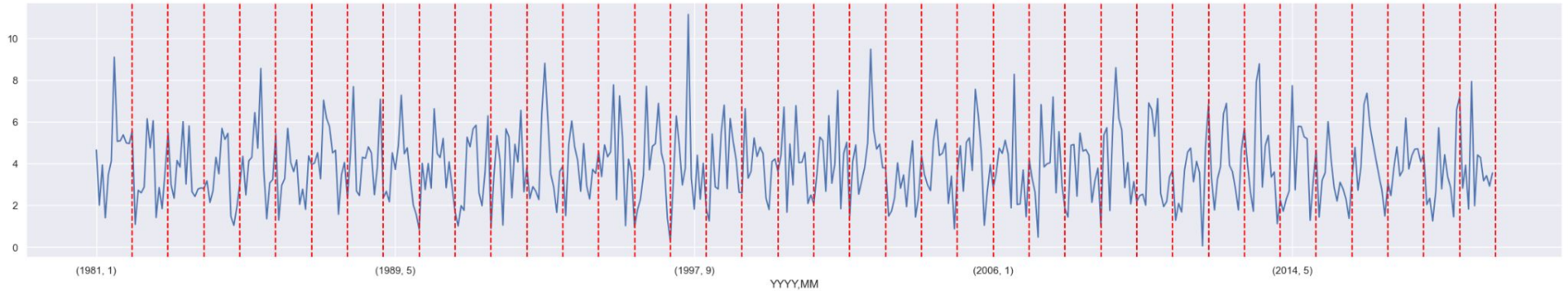
- Predict future events based on past
- Consideration:
  - Trend
  - Seasonality
- Different models
- Combination of models



Data Analysis - Findings



Data Analysis - Findings



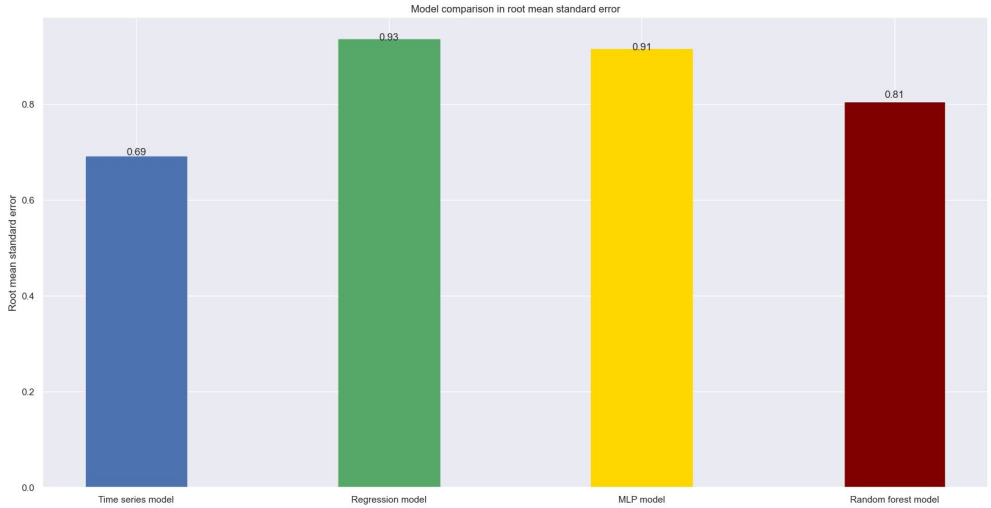
Experiments and Results - Setup

- Models
  - Autoregressive Model
  - Regression Model
  - Random Forest Model
  - MLP Model
- Task
  - Predict next days precipitation
- Preprocessing Standard
  - Adding area attributes
  - Setup next day prediction
  - One Dataframe
  - Normalization
- Preprocessing AR
  - Datetime Index
  - One Dataframe of targets

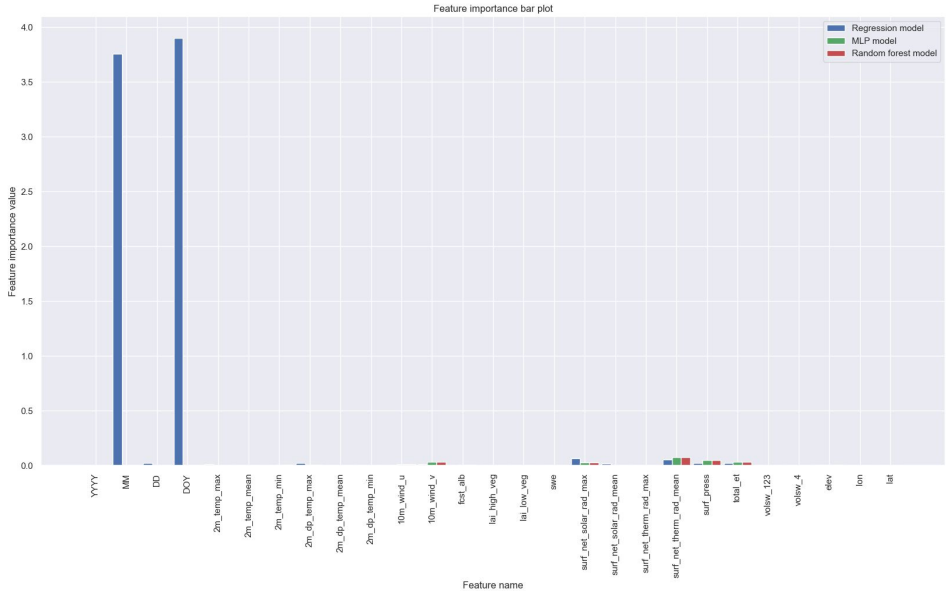
Experiments and Results - Tests

- Train basic models
- Prediction only for the next day
- Evaluate with RMSE
- Feature Importance through permutation

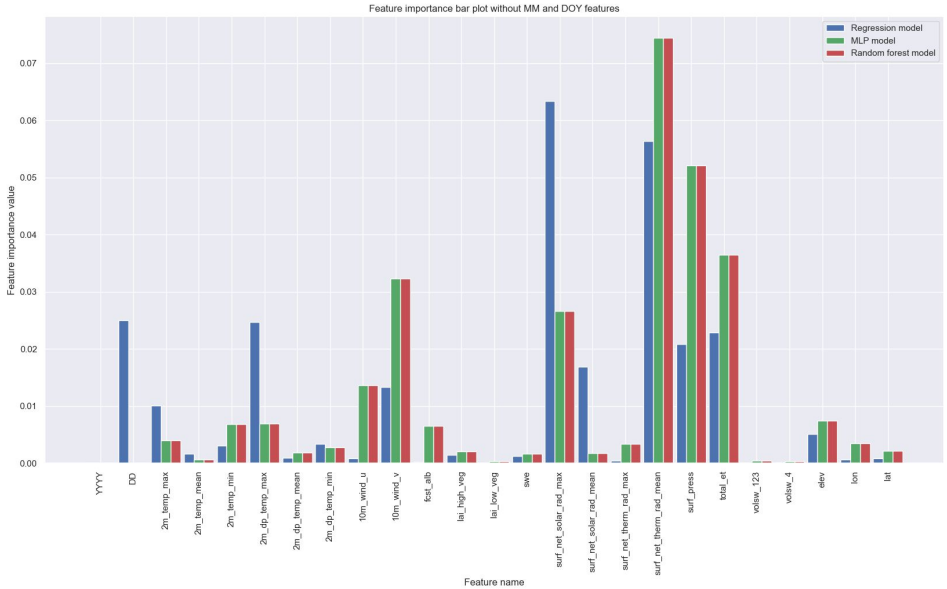
Experiments and Results - RMSE



Experiments and Results - Feature Importance



Experiments and Results - Feature Importance



Conclusion

- Feature Importance hard to compare
- Modeling time series is different
- Hyperparameter tuning computationally intensive
- Improvements:
  - More complex models
  - SARIMA, RNN
  - Make use of all features in time series model