Forecasting charging station occupancy using supervised learning algorithms

Adrian Ostermann, Yann Fabel, Kim Ouan

01.10.2021

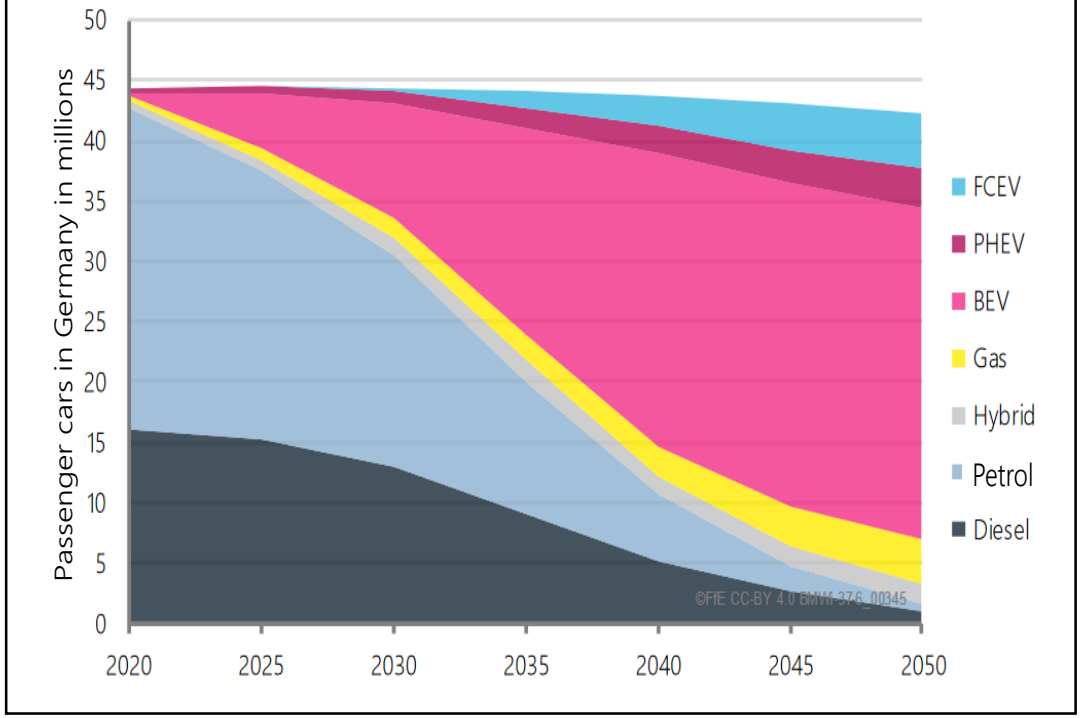
Federal Ministry
for Economic Affairs
and Energy

Agenda

1. Motivation
2. Methodology
3. ACN Data
4. Evaluation
5. Summary & Outlook

Motivation

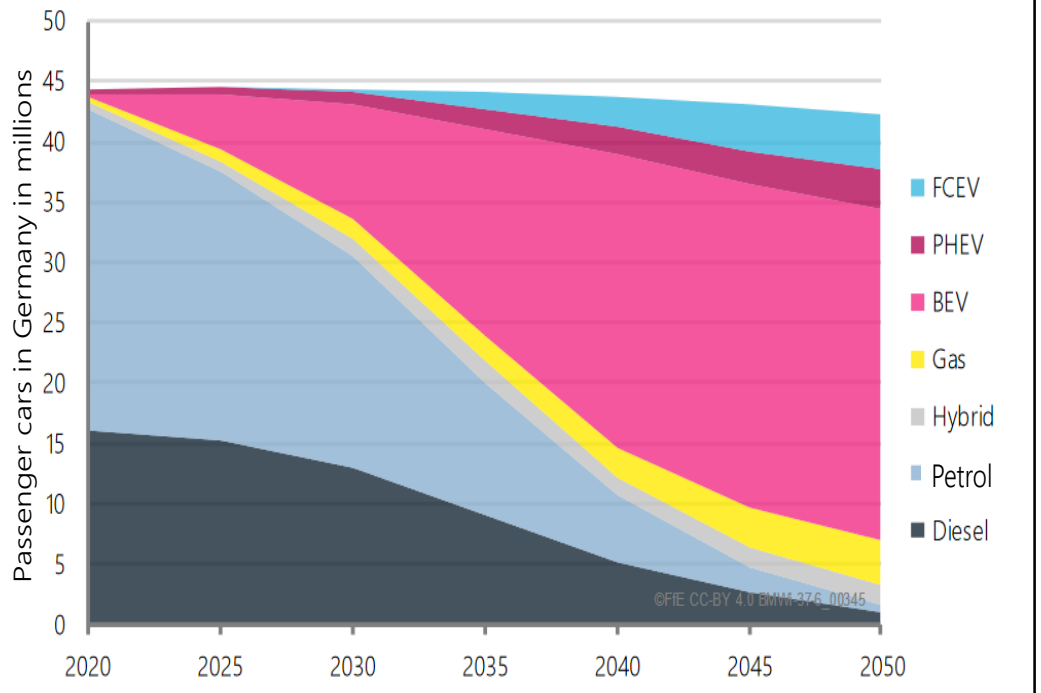
Development car types in GER [1]



FCEV: fuel cell electric vehicles
PHEV: plugin-in-hybrid
BEV: battery-powered electric vehicles
CP: Charging Point

Motivation

Development car types in GER [1]



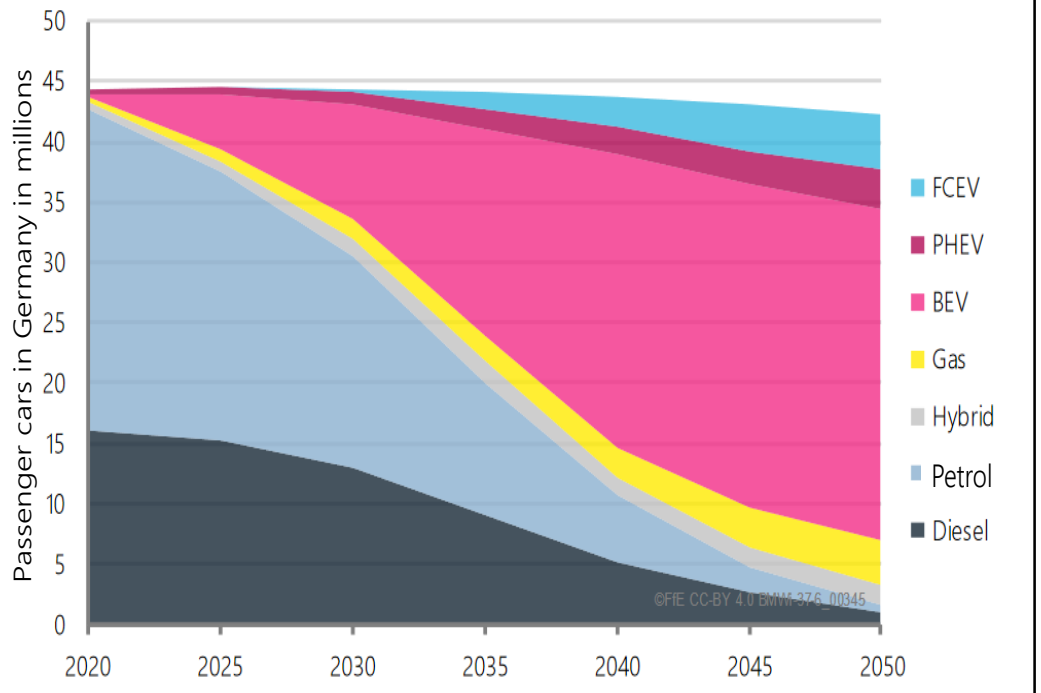
Development charging points in GER [2] [3]

- National Centre for Charging Infrastructure 2030:
  - 440.000 to 843.000 public CP
  - 2.5 Mio to 2.7 Mio. CP at work
- Directive 2014/94/EU:
  - 10:1 ratio of EV to publicly accessible CPs

FCEV: fuel cell electric vehicles
PHEV: plugin-in-hybrid
BEV: battery-powered electric vehicl
CP: Charging Point

Motivation

Development car types in GER [1]



Development charging points in GER [2] [3]

- National Centre for Charging Infrastructure 2030:
  - 440.000 to 843.000 public CP
  - 2.5 Mio to 2.7 Mio. CP at work
- Directive 2014/94/EU:
  - 10:1 ratio of EV to publicly accessible CPs

Charging behaviour in GER [2]

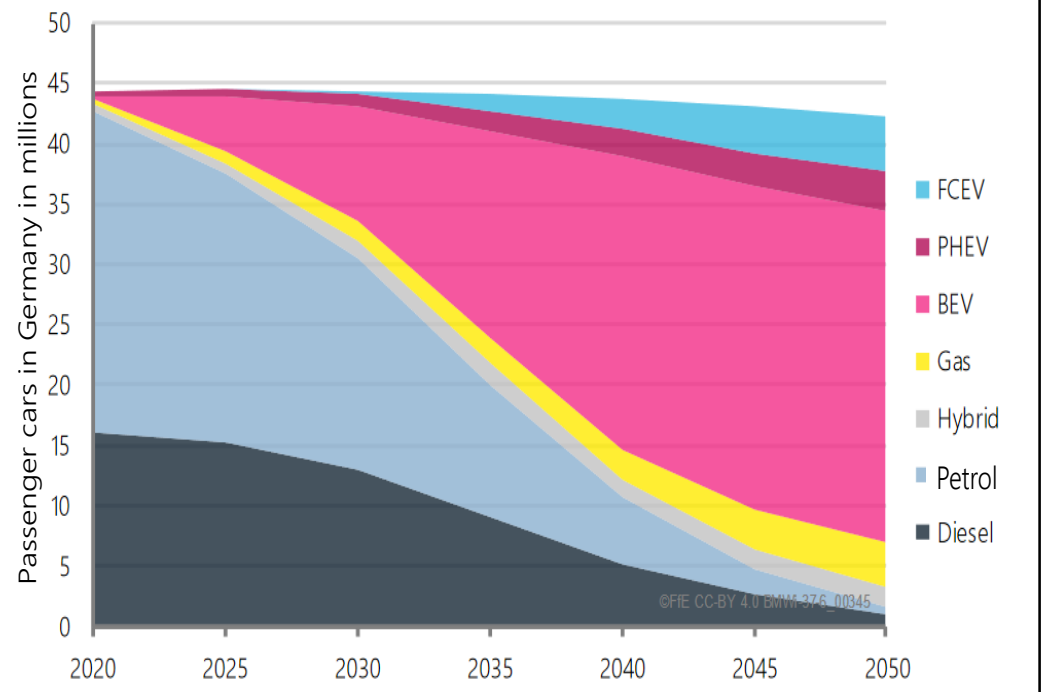
- 20 % of charging done at work
- 12 to 24 % (depending on charging behaviour development) at public CP
- 70 % of EV Users use a public CP at least once per month

->Available infrastructure is one of the main constraints for users

FCEV: fuel cell electric vehicles
PHEV: plugin-in-hybrid
BEV: battery-powered electric vehicl
CP: Charging Point

Motivation

Development car types in GER [1]



Development charging points in GER [2] [3]

- National Centre for Charging Infrastructure 2030:
  - 440.000 to 843.000 public CP
  - 2.5 Mio to 2.7 Mio. CP at work
- Directive 2014/94/EU:
  - 10:1 ratio of EV to publicly accessible CPs

Charging behaviour in GER [2]

- 20 % of charging done at work
- 12 to 24 % (depending on charging behaviour development) at public CP
- 70 % of EV Users use a public CP at least once per month

->Available infrastructure is one of the main constraints for users

With the prediction of CP occupancy, we aim to support:

a) EV users to plan their charging progress

b) scheduling CP maintenance

c) energy manager anticipate utilization

FCEV: fuel cell electric vehicles
PHEV: plugin-in-hybrid
BEV: battery-powered electric vehicl
CP: Charging Point

Methodology

ACN Dataset

Caltech	JPL
• 54 EVSE	• 52 EVSE
• Public	• Workplace

Data Exploration

- Describe data
- Per site & station:
  - Occupancy durations
  - Usage patterns

Data Preparation

- Feature Selection
- Target variable: occupation
- Different Feature Sets for comparison
- Train / Test split
- Data preprocessing

Supervised Learning

- Binary: available or occupied
- 4 different machine learning algorithms for classification: Logistic Regression, KNN, Random Forest & XGBoost

Hyperparameter Tuning

- Grid search
- Different parameter ranges
- Optimization for AUC-Score

Model Evaluation

- Performance for different metrics
- Comparison:
  - Feature Sets
  - Location: work & public
  - Mode: Individual CP training vs. all

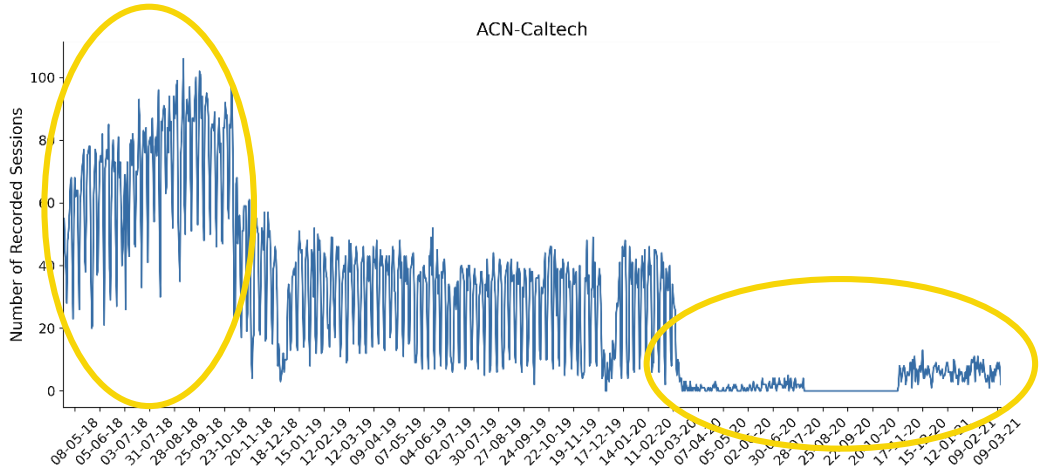
ACN: Adaptive Charging Network
KNN: K-Nearest-Neighbor
AUC: Area under the Curve

How to predict charging station occupancy?

ACN Data

Caltech (public)

- 54 Charging Points
- More than 25.000 charging sessions
- Garage at Caltech campus near campus gym
- Public charging



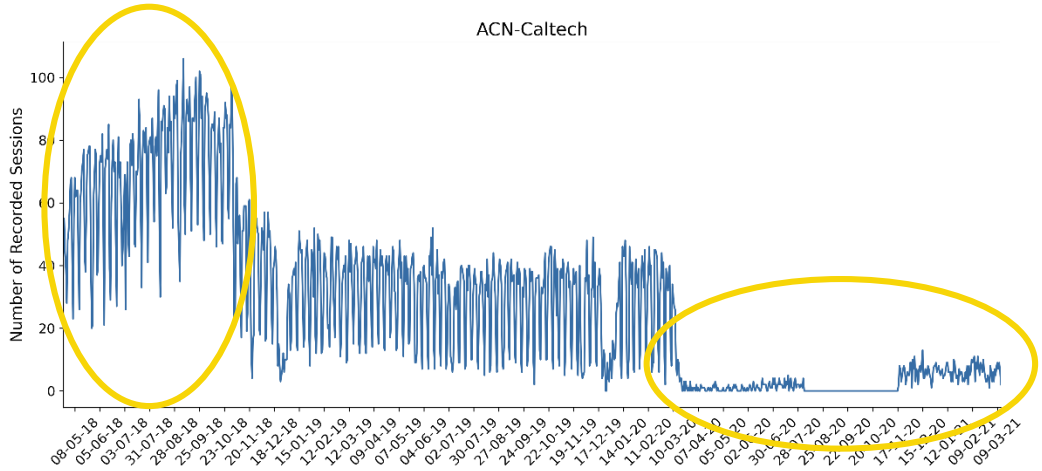
- Exclude free charging and corona period
- 01.12.2018 - 29.02.2020

JPL (work)

ACN Data

Caltech (public)

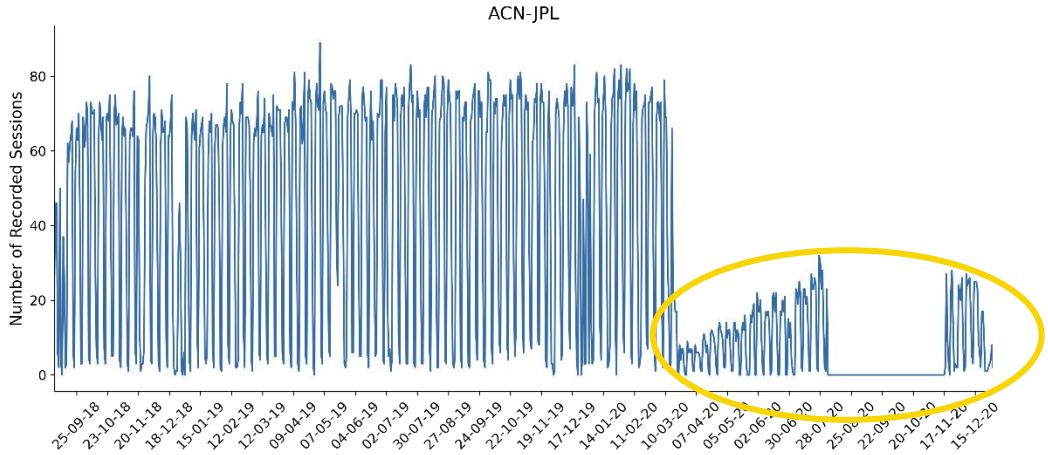
- 54 Charging Points
- More than 25.000 charging sessions
- Garage at Caltech campus near campus gym
- Public charging



- Exclude free charging and corona period
- 01.12.2018 - 29.02.2020

JPL (work)

- 52 Charging Points
- More than 25.000 charging sessions
- Garage at JPL campus with restricted access
- Workplace charging

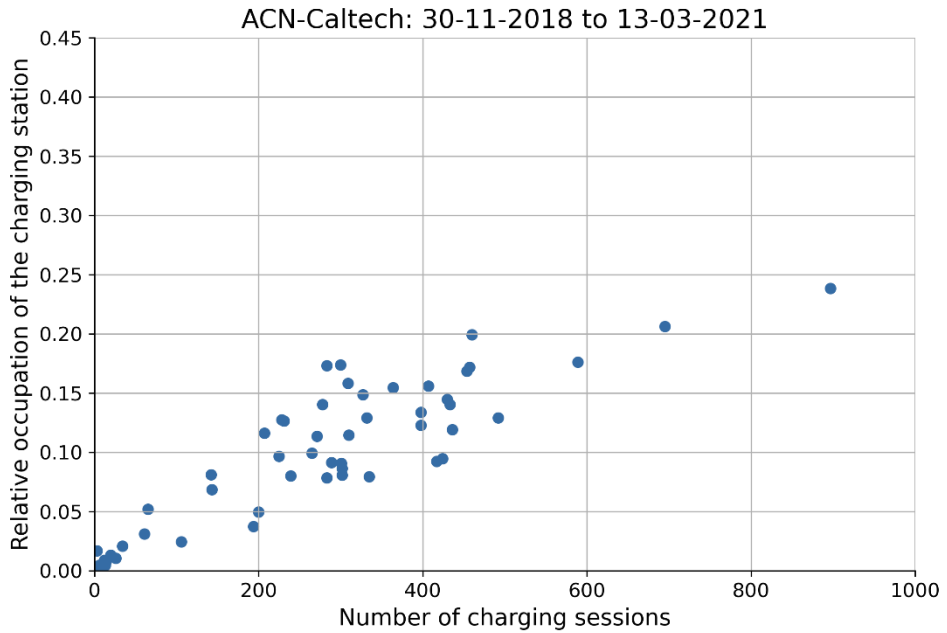


- Exclude corona period
- 05.09.2018 - 29.02.2020

Two datasets: public & work place charging sessions for more than one year

ACN Data

Caltech (public)

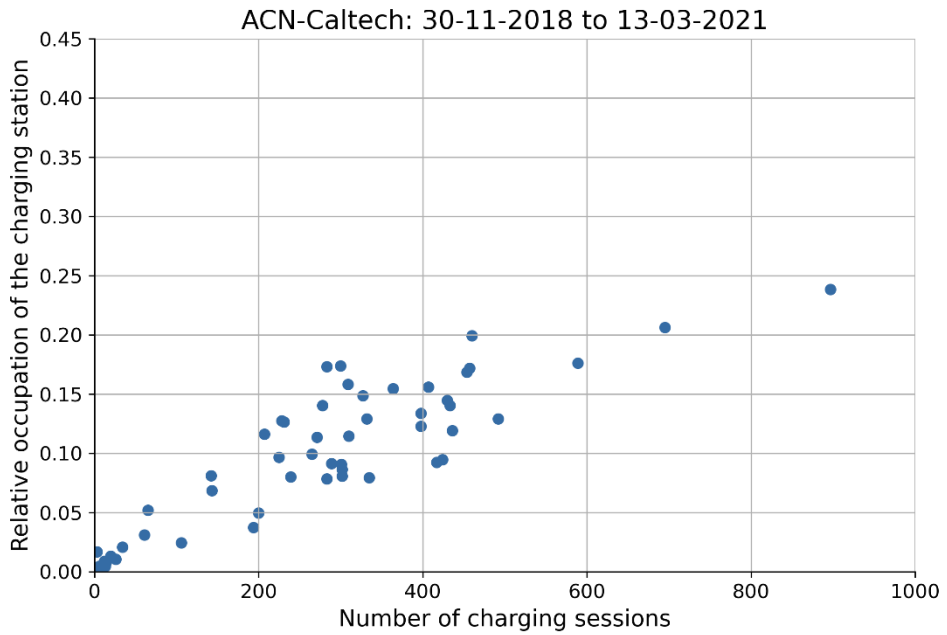


- More heterogeneous
- Relative occupation low, some stations less than 200 charging sessions
- Occupied - 14 %, Not Occupied - 86 %

JPL (work)

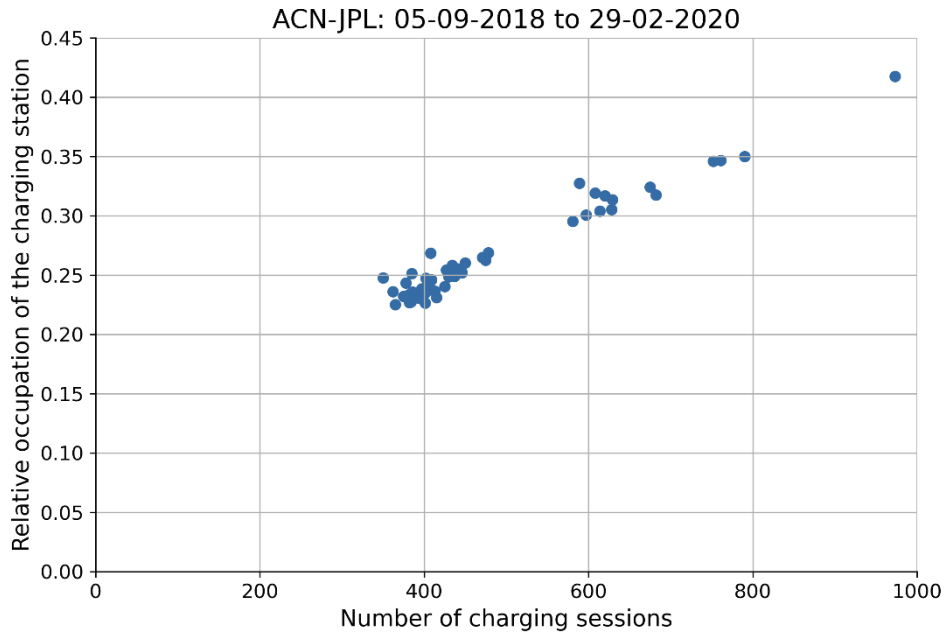
ACN Data

Caltech (public)



- More heterogeneous
- Relative occupation low, some stations less than 200 charging sessions
- Occupied - 14 %, Not Occupied - 86 %

JPL (work)



- More homogeneous
- Relative occupation high, all stations more than 350 charging sessions
- Occupied - 27 %, Not Occupied - 73 %

Caltech more imbalanced than JPL

Evaluation – Individual vs. All

Individual

- Use grid search for each charging station to determine hyperparameters where model performs best on charging station specific data set
- One model per algorithm for each charging station per site

All

- Use grid search to determine hyperparameters where model performs best on whole data set
- One model per algorithm per site

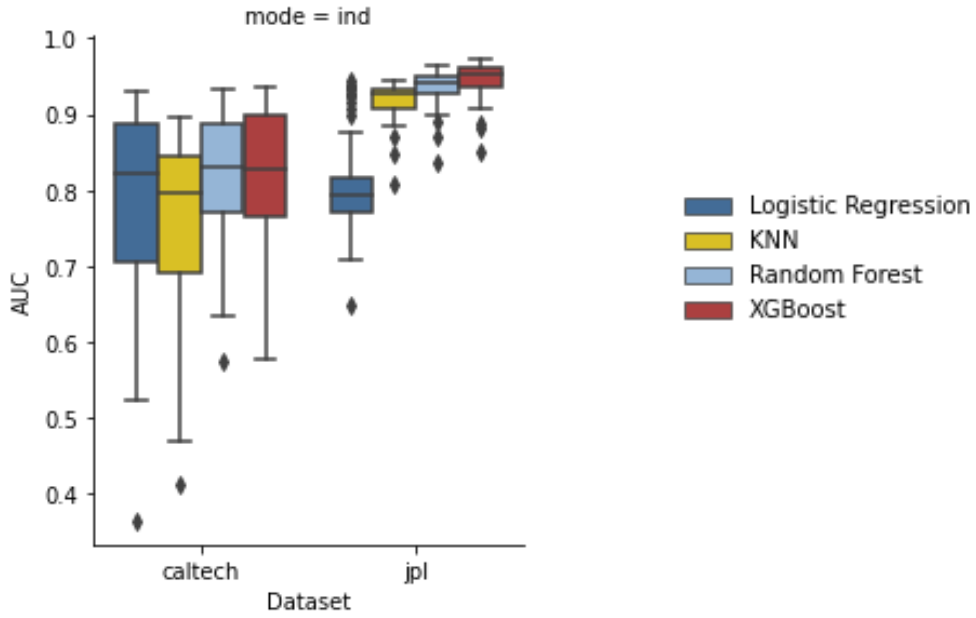
Evaluation – Individual vs. All

Individual

- Use grid search for each charging station to determine hyperparameters where model performs best on charging station specific data set
- One model per algorithm for each charging station per site

All

- Use grid search to determine hyperparameters where model performs best on whole data set
- One model per algorithm per site



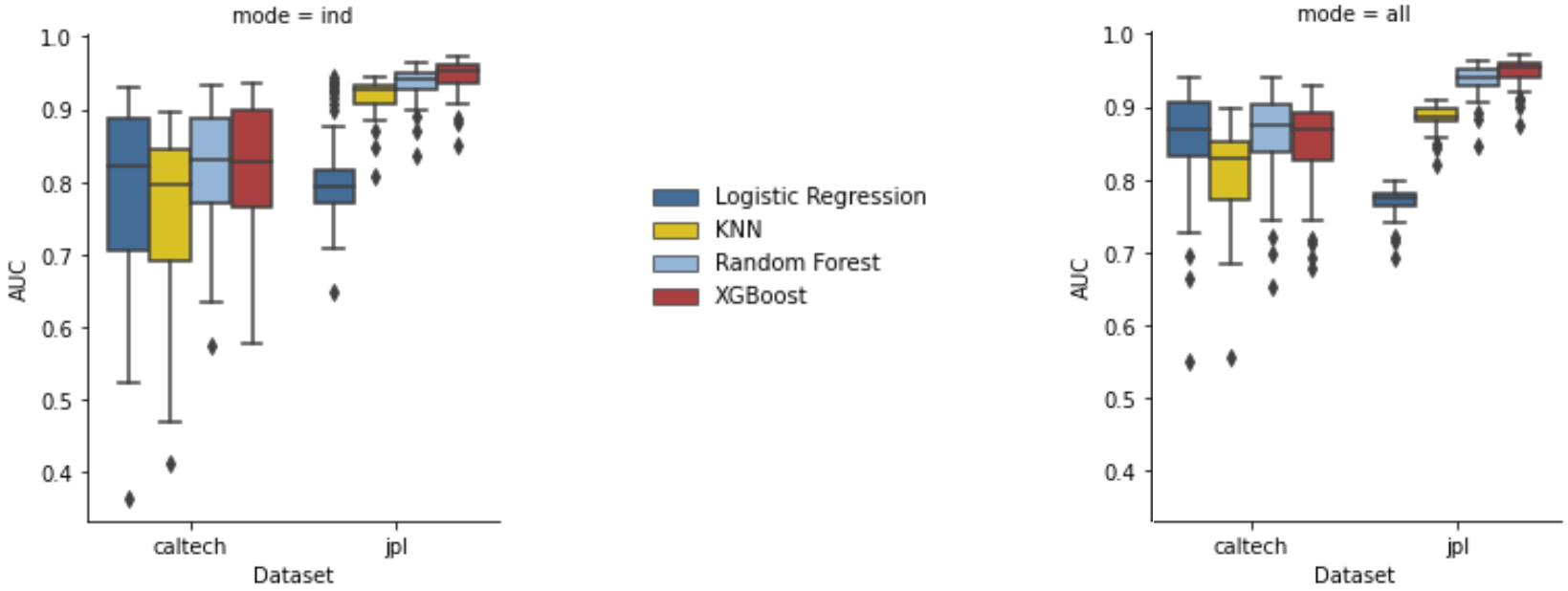
Evaluation – Individual vs. All

Individual

- Use grid search for each charging station to determine hyperparameters where model performs best on charging station specific data set
- One model per algorithm for each charging station per site

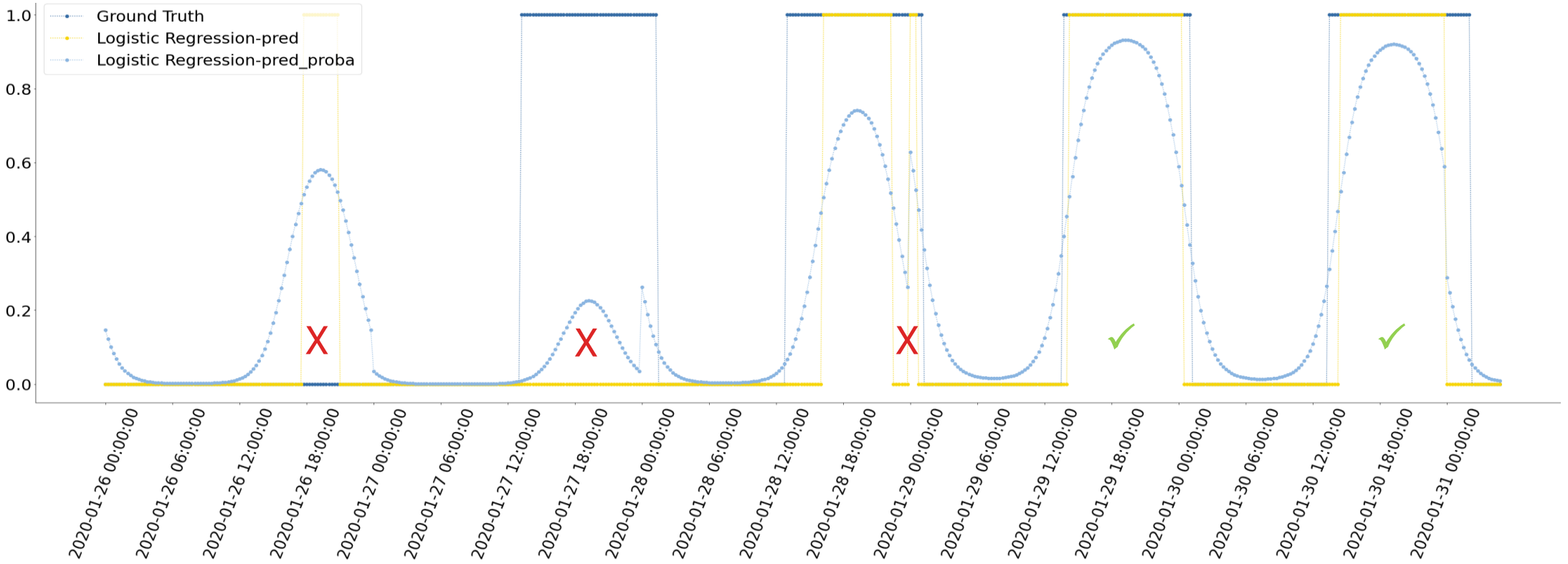
All

- Use grid search to determine hyperparameters where model performs best on whole data set
- One model per algorithm per site



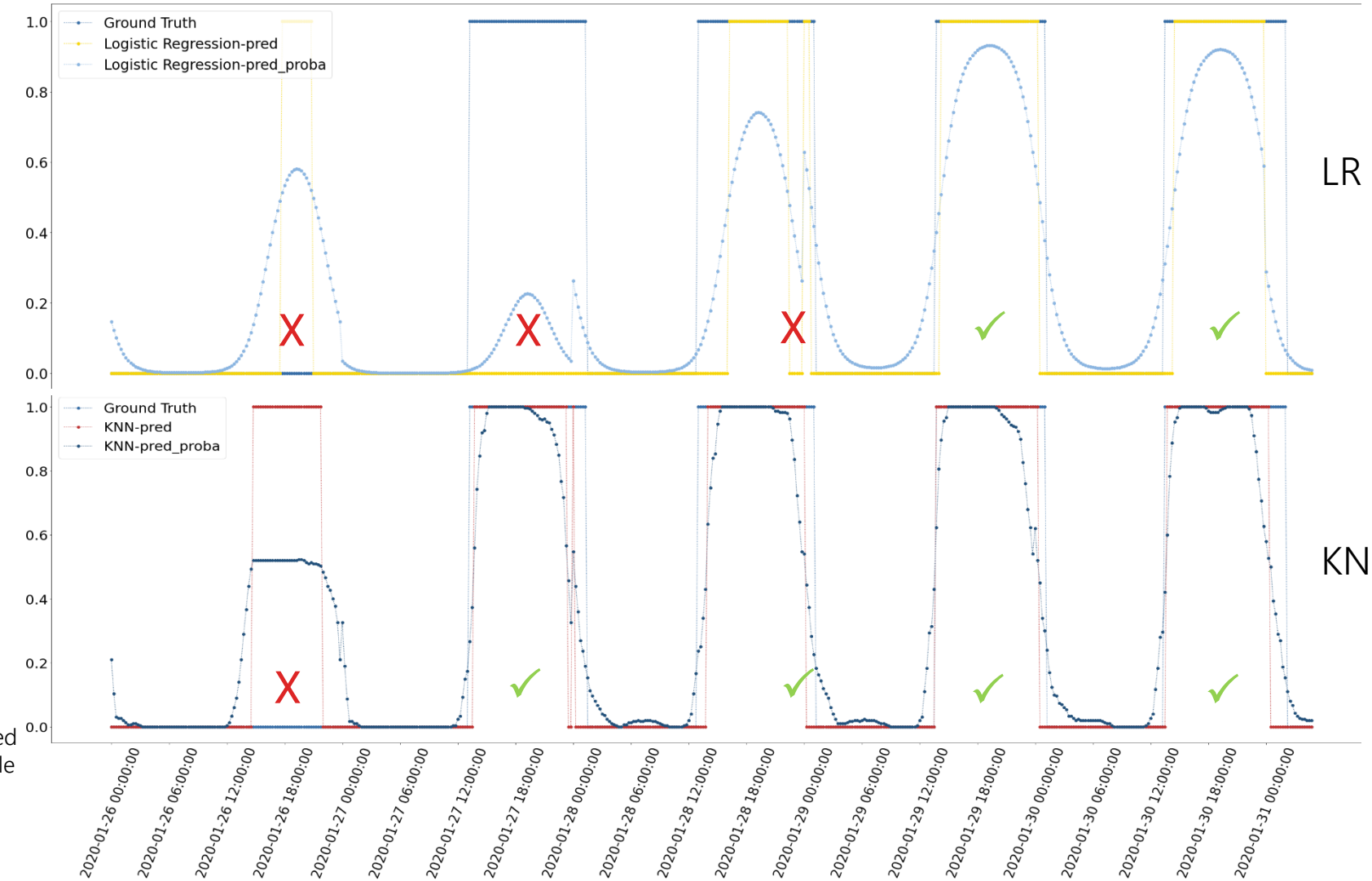
XGBoost performs best on JPL-data, Random Forest on Caltech-data
Individual mode has higher variation
All mode performs better on Caltech-data

Evaluation – Logistic Regression (LR)
JPL Sample Week 26.01.2020 (Sun) – 31.01.2020 (Fri)



1 = Occupied
0 = Available

Evaluation – KNN
JPL Sample Week 26.01.2020 (Sun) – 31.01.2020 (Fri)

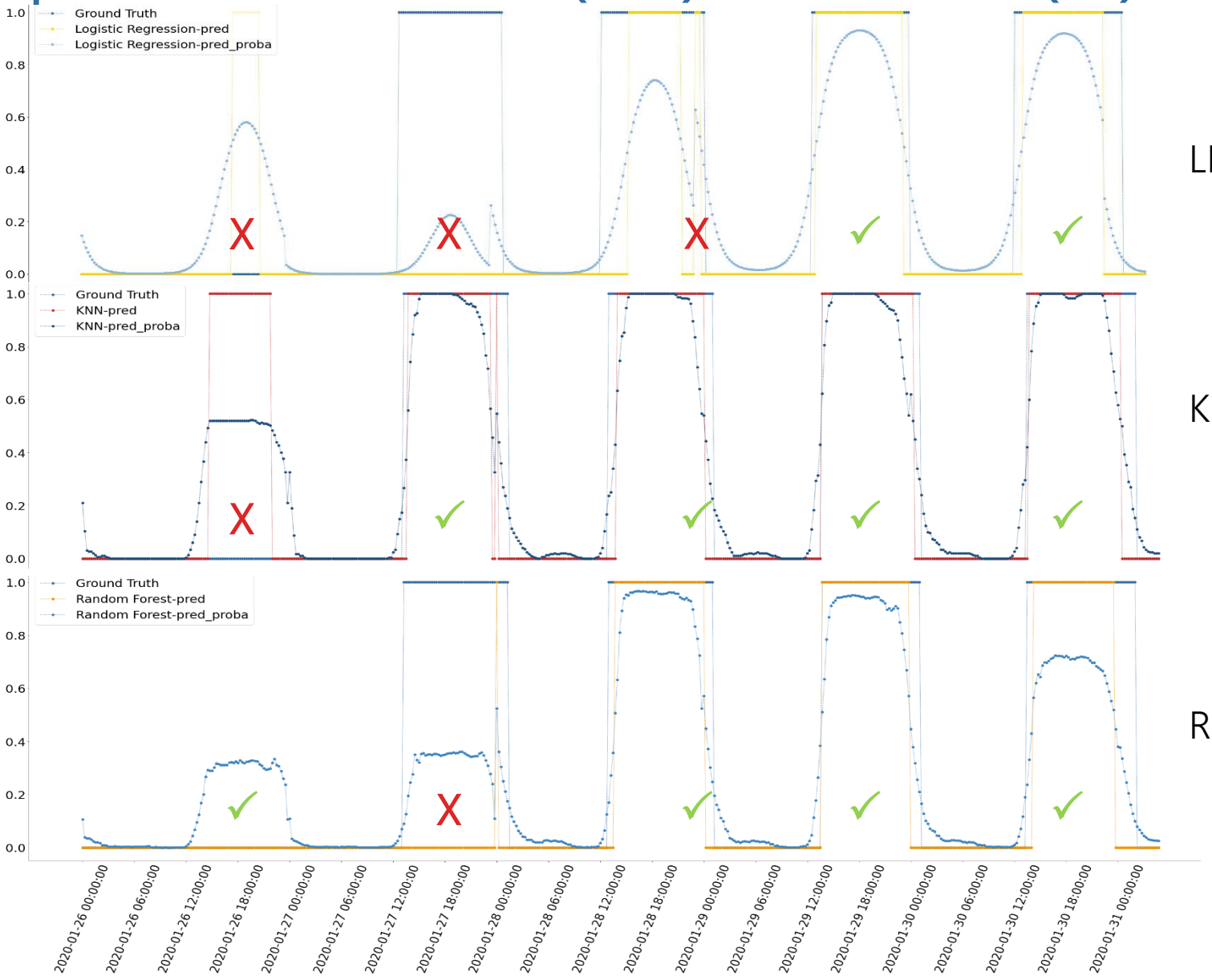


LR

KNN

1 = Occupied
0 = Available

Evaluation – Random Forest
JPL Sample Week 26.01.2020 (Sun) – 31.01.2020 (Fri)



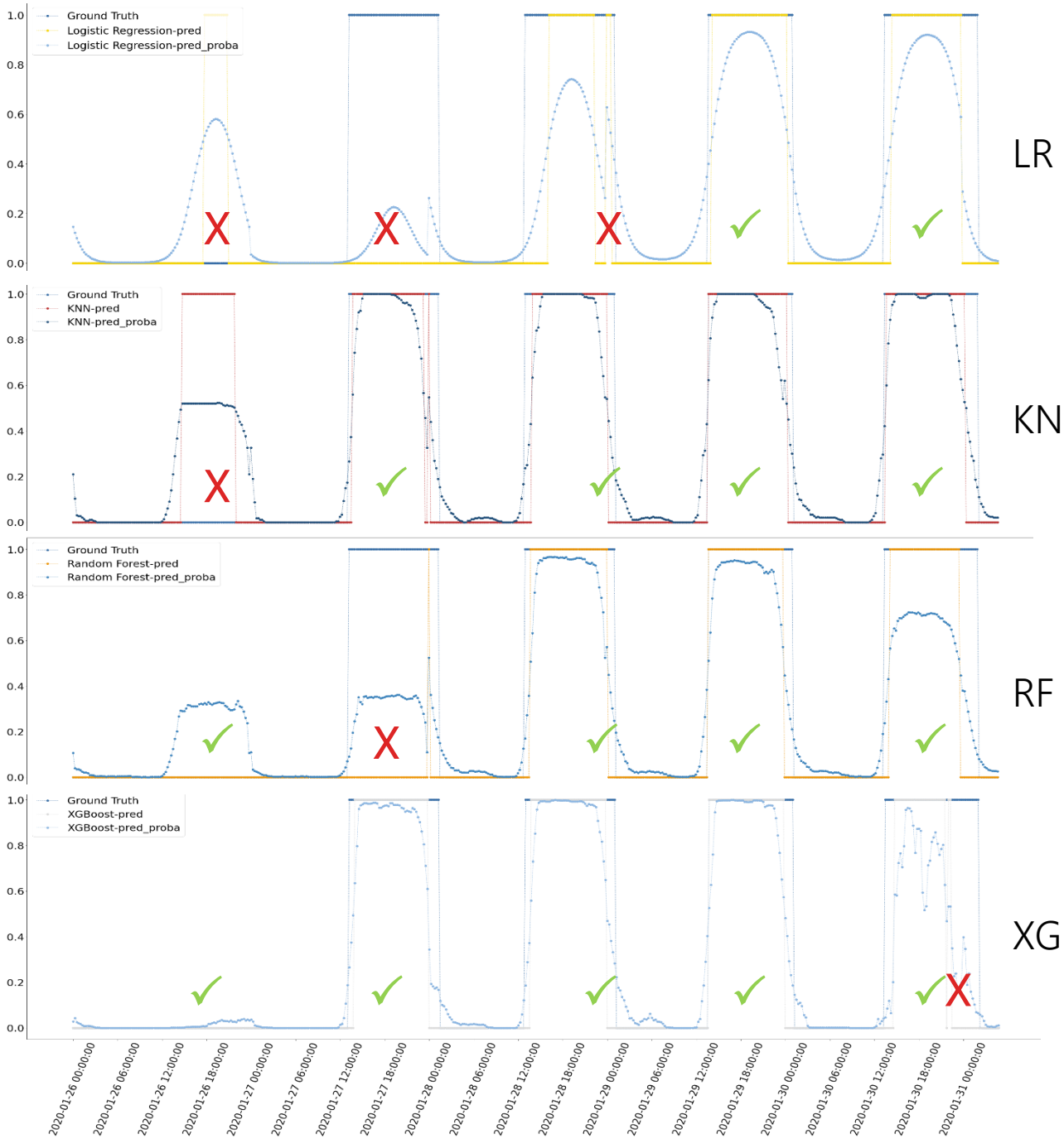
LR

KNN

RF

1 = Occupied
0 = Available

Evaluation
JPL XGBoost



1 = Occupied
0 = Available

Summary & Outlook

Summary

- Random Forest and XGBoost perform best on our data
- Higher relative occupancy rate leads to better prediction
- Considering all data from one site can lead to better prediction for a single charging station

Outlook

- Comparison of different input feature sets
- Optimization and comparison for different metrics

Adrian Ostermann, M. Sc.

Research Center for Energy Economics (FfE)

Phone: +49(0)89 15 81 21 – 50

Email: aostermann@ffe.de

Forschungsstelle für Energiewirtschaft

Am Blütenanger 71 – 80995 München

Tel.: +49(0)89 15 81 21 – 0

Email: info@ffe.de

Internet: www.ffe.de

Twitter: @FfE_Muenchen

References

[1] Fattler, Steffen; Conrad, Jochen; Regett, Anika et al.: Dynamis Hauptbericht - Dynamis - Dynamische und intersektorale Maßnahmenbewertung zur kosteneffizienten Dekarbonisierung des Energiesystems - Online: https://www.ffe.de/dynamis. München: Forschungsstelle für Energiewirtschaft e.V., 2019. DOI: 10.34805/ffe-144-19

[2] Windt, Alexander et al.: Ladeinfrastruktur nach 2025/2030: Szenarien für den Markthochlauf. Berlin: Nationale Leitstelle Ladeinfrastruktur, 2020.

[3] Directive 2014/94/EU of the European Parliament and of the Council on the deployment of alternative fuels infrastructure. Ausgefertigt am 2014-10-22, Version vom 2017-11-18; Brussels: European Parliament and the Council, 2017.