DEMOCRITUS UNIVERSITY OF THRACE
DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Federated Learning for 5G Base Station Traffic Forecasting

Vasileios Perifanis
PhD Student
Dept. of Electrical and Computer Engineering
Democritus University of Thrace
Greece

Team Euclid

Joint work Nikolaos Pavlidis, Remous-Aris Koutsiamanis and Pavlos Efraimidis

November 30, 2022

2022 ITU AI/ML in 5G Challenge Play-offs

Challenge Overview

- Traffic prediction is of great importance towards the 5G era.
    - Smart and efficient infrastructure planning

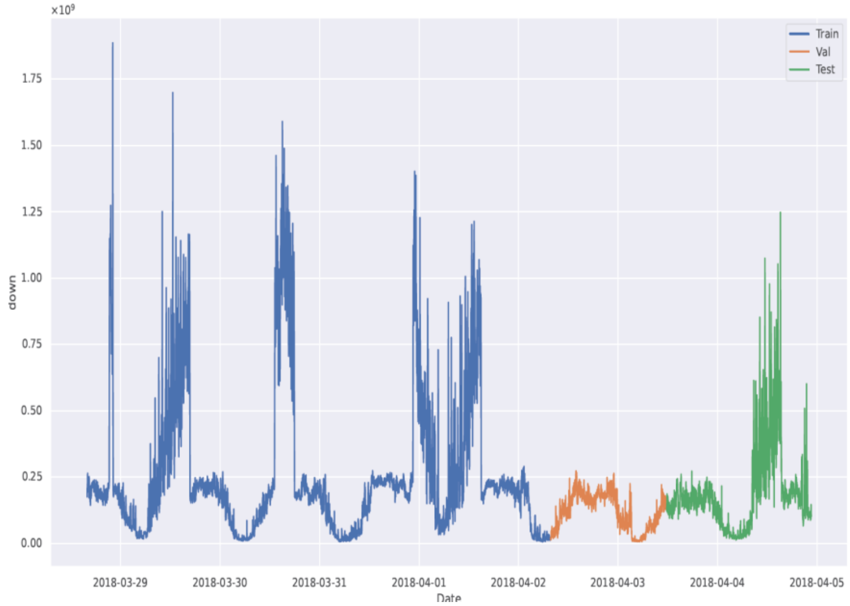
- Suitable Deep Learning models: MLPs, RNNs, CNNs.

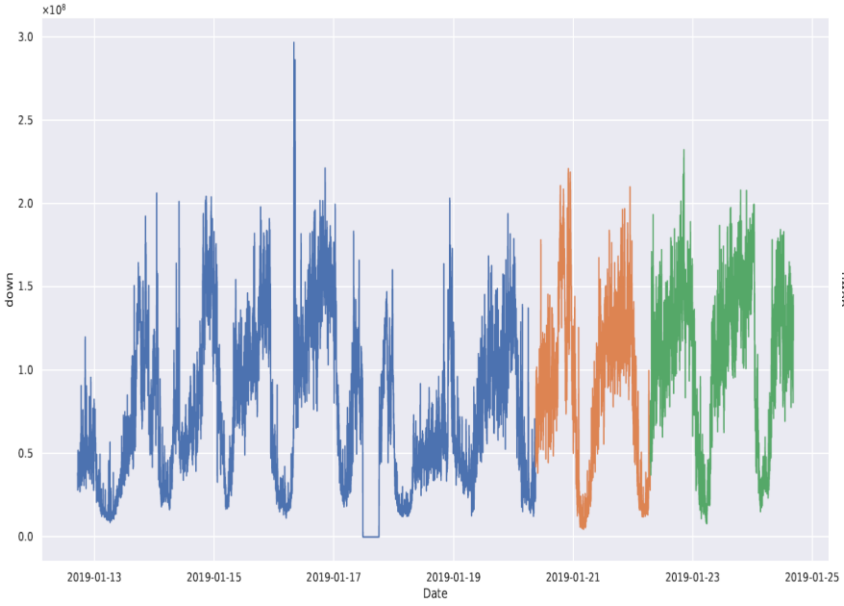
- Traditional approaches build forecasting models using combined base station data.
    - Privacy violations
    - Business competitiveness and confidentiality issues
    - Solution: Collaborative Learning without exchanging data

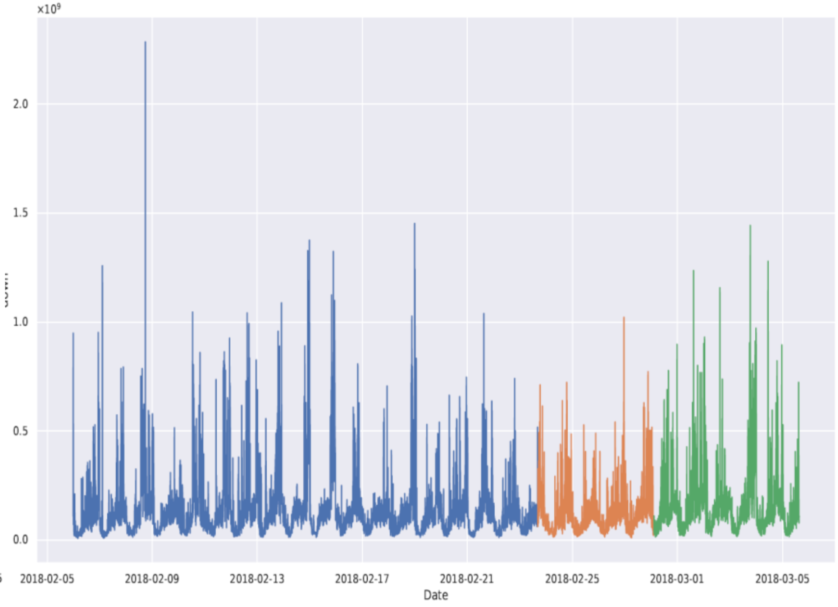
- Challenge:
    - Goal: Build a Federated Model Architecture
    - Input: 10 past observations (11 measurements/observation)
    - Output: Next-step prediction (5 measurements)

Dataset

- Collection of measurements from three base stations in the metropolitan area of Barcelona, Spain.
  - ElBorn: 5421 samples (7 days)
  - LesCorts: 8615 samples (12 days)
  - PobleSec: 19909 samples (28 days)

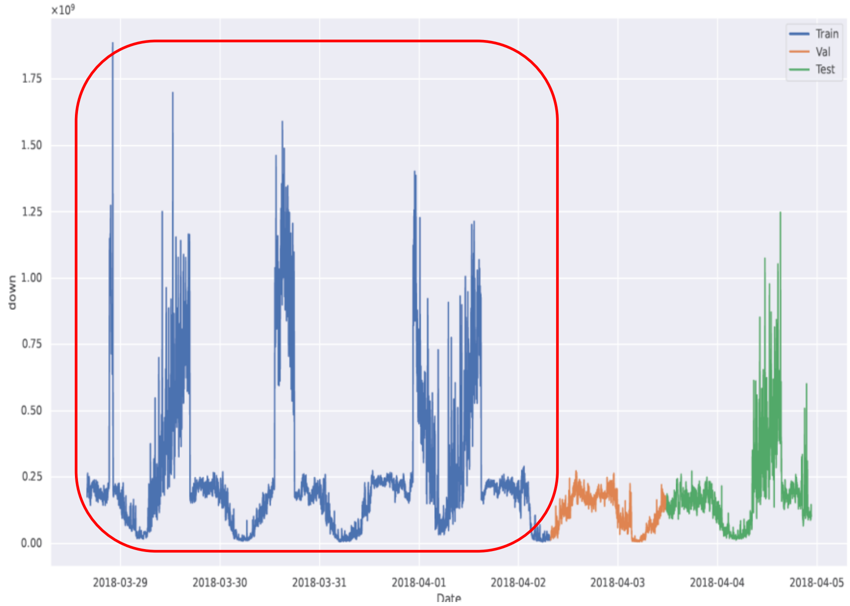


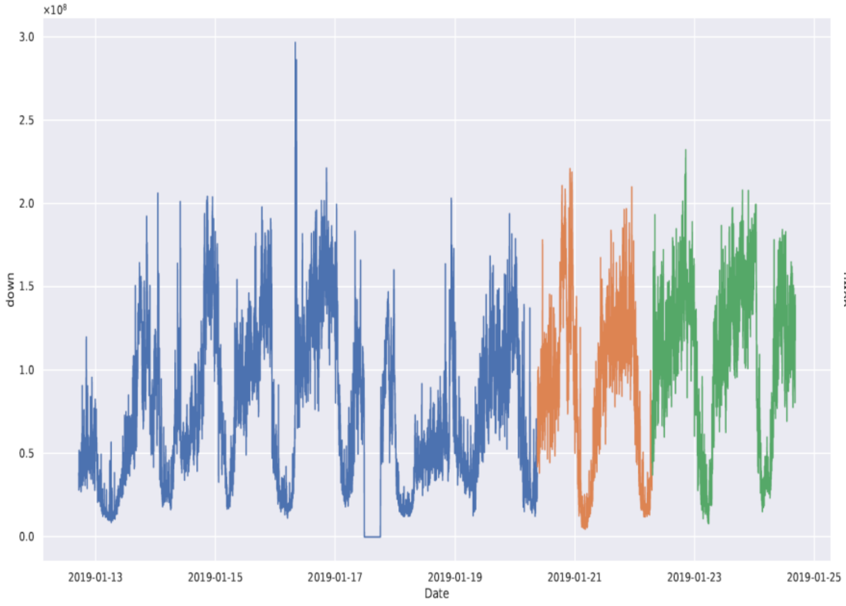


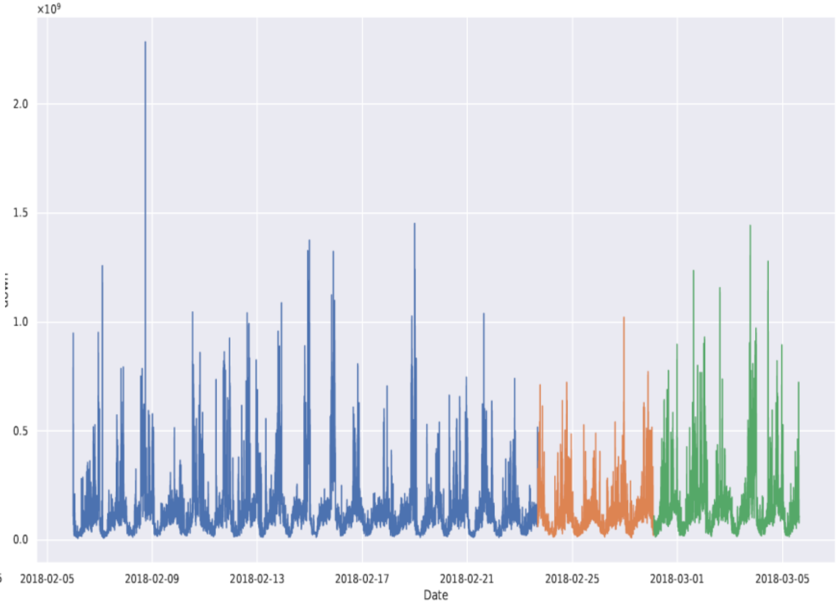


Dataset

- Collection of measurements from three base stations in the metropolitan area of Barcelona, Spain.
    - ElBorn: 5421 samples (7 days)
    - LesCorts: 8615 samples (12 days)
    - PobleSec: 19909 samples (28 days)

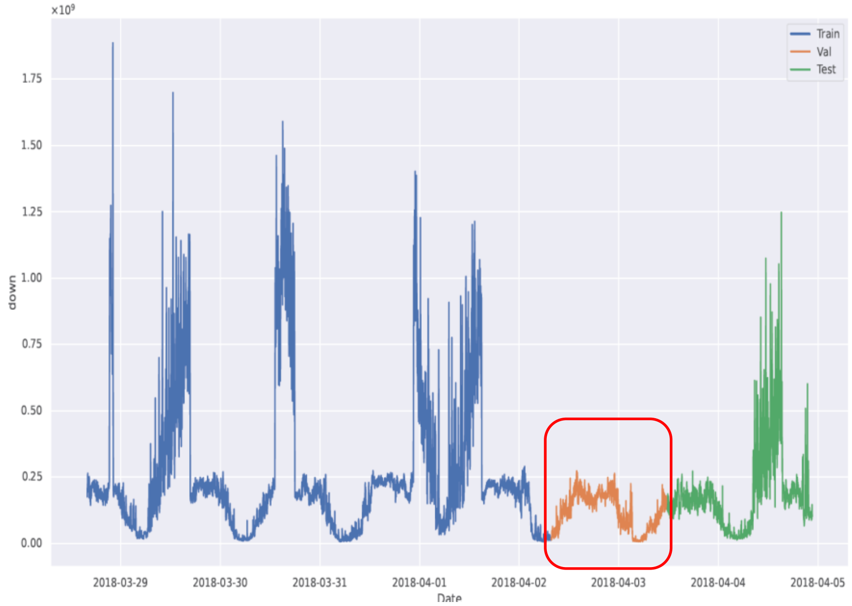


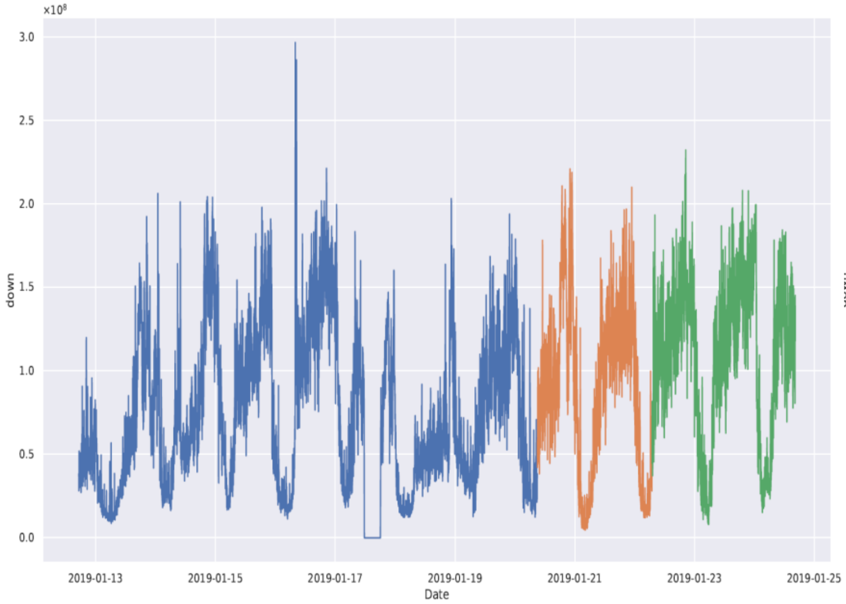


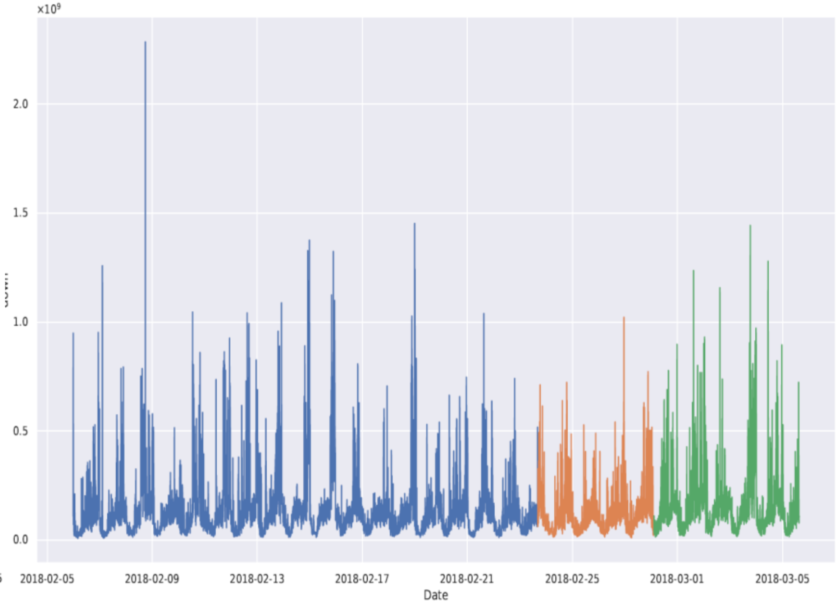


Dataset

- Collection of measurements from three base stations in the metropolitan area of Barcelona, Spain.
  - ElBorn: 5421 samples (7 days)
  - LesCorts: 8615 samples (12 days)
  - PobleSec: 19909 samples (28 days)

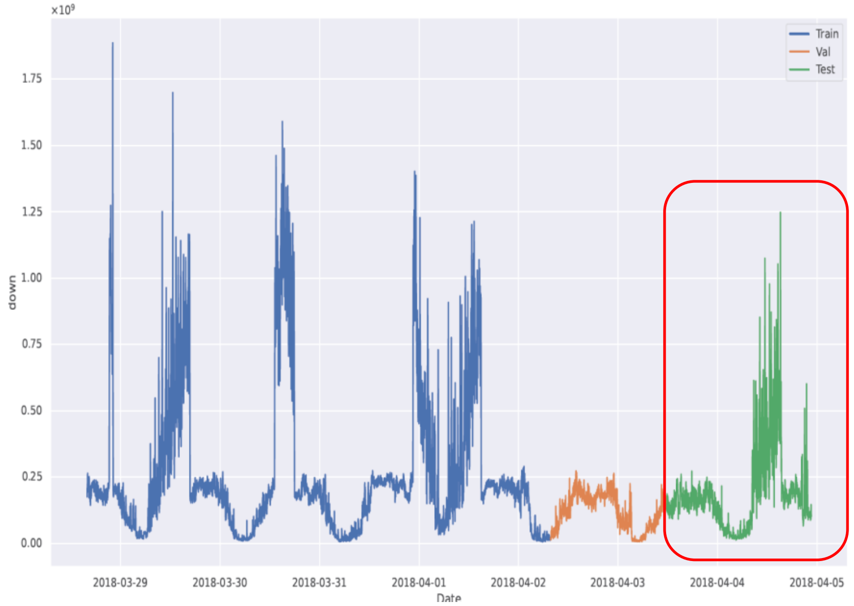


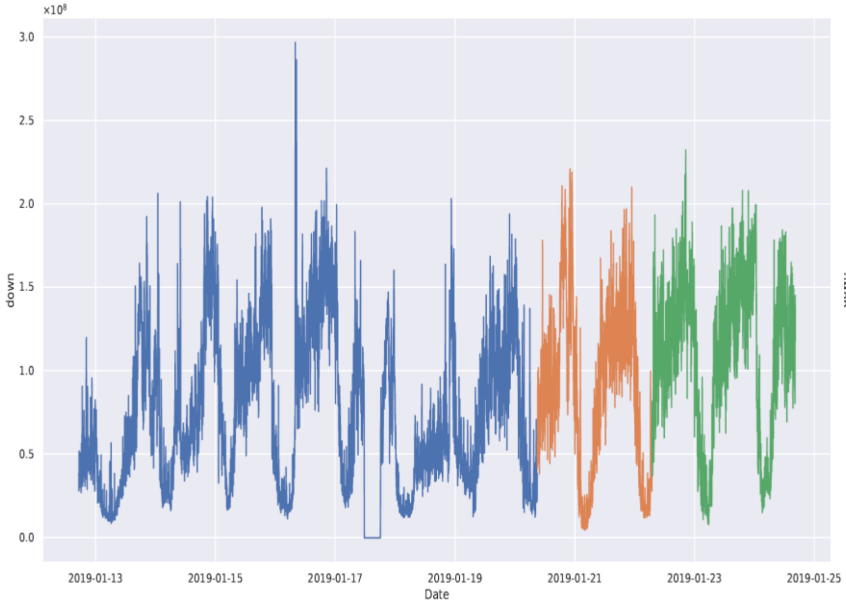


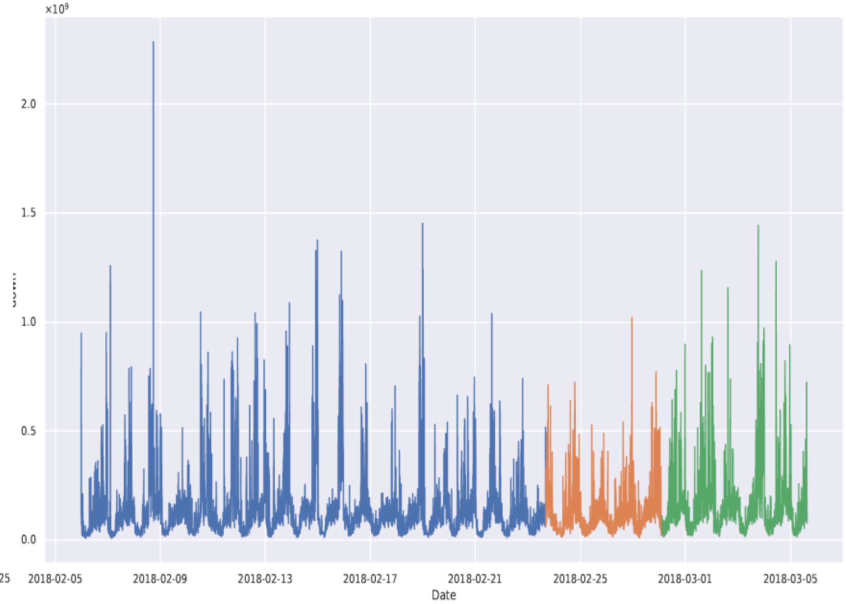


Dataset

- Collection of measurements from three base stations in the metropolitan area of Barcelona, Spain.
  - ElBorn: 5421 samples (7 days)
  - LesCorts: 8615 samples (12 days)
  - PobleSec: 19909 samples (28 days)

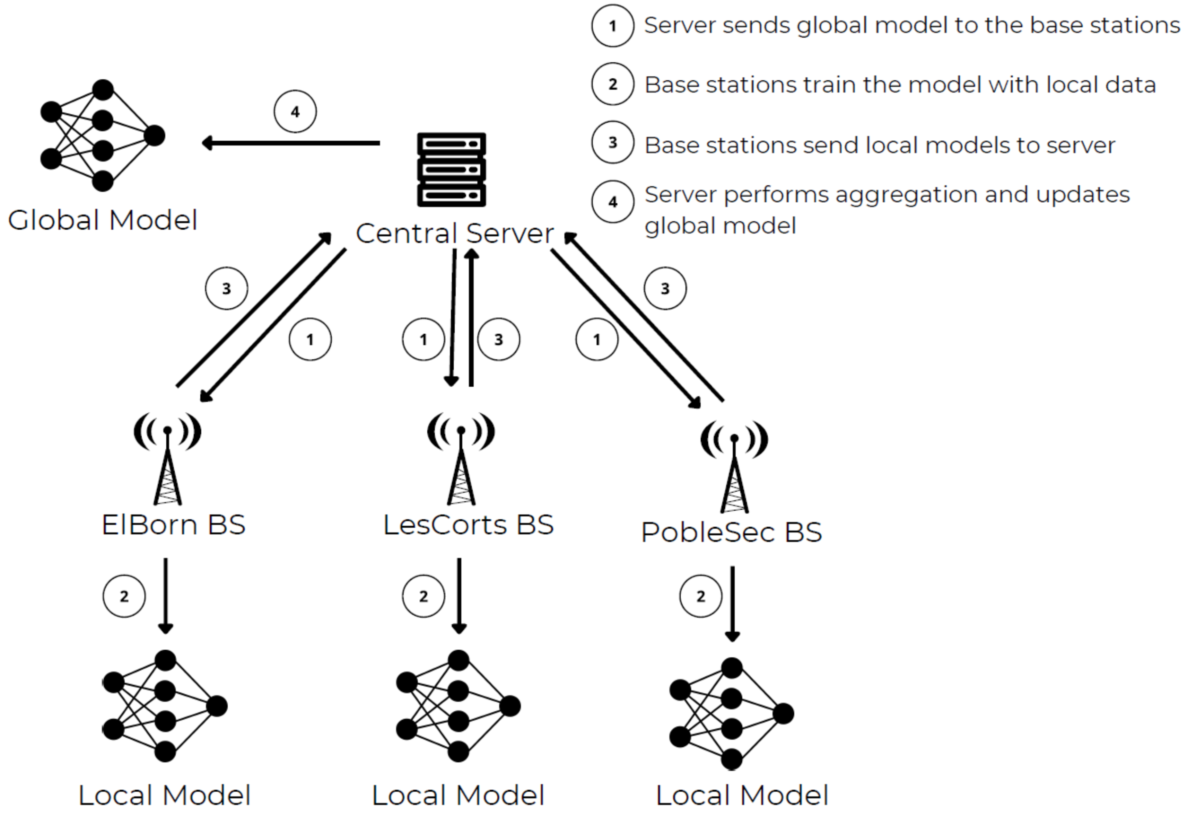






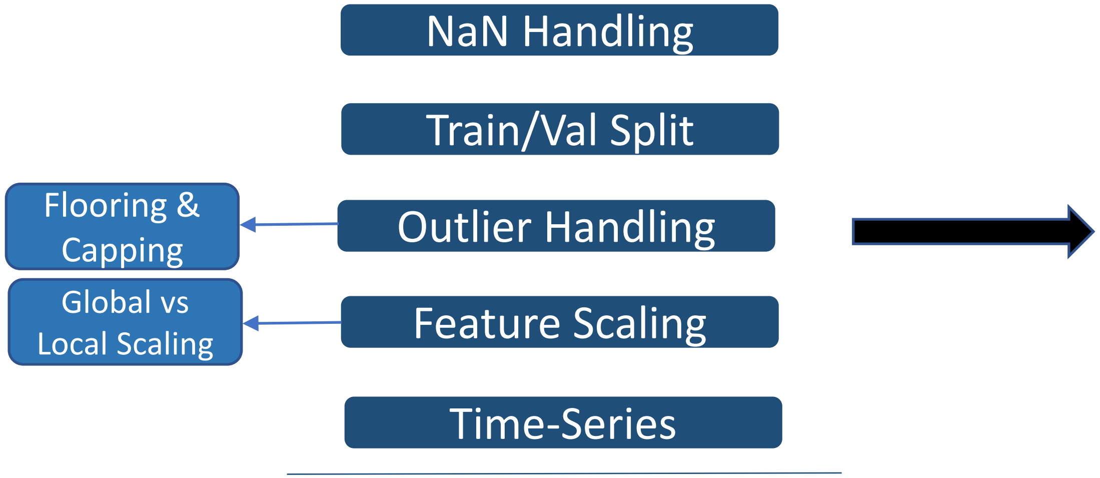
Federated Learning

Pre-processing

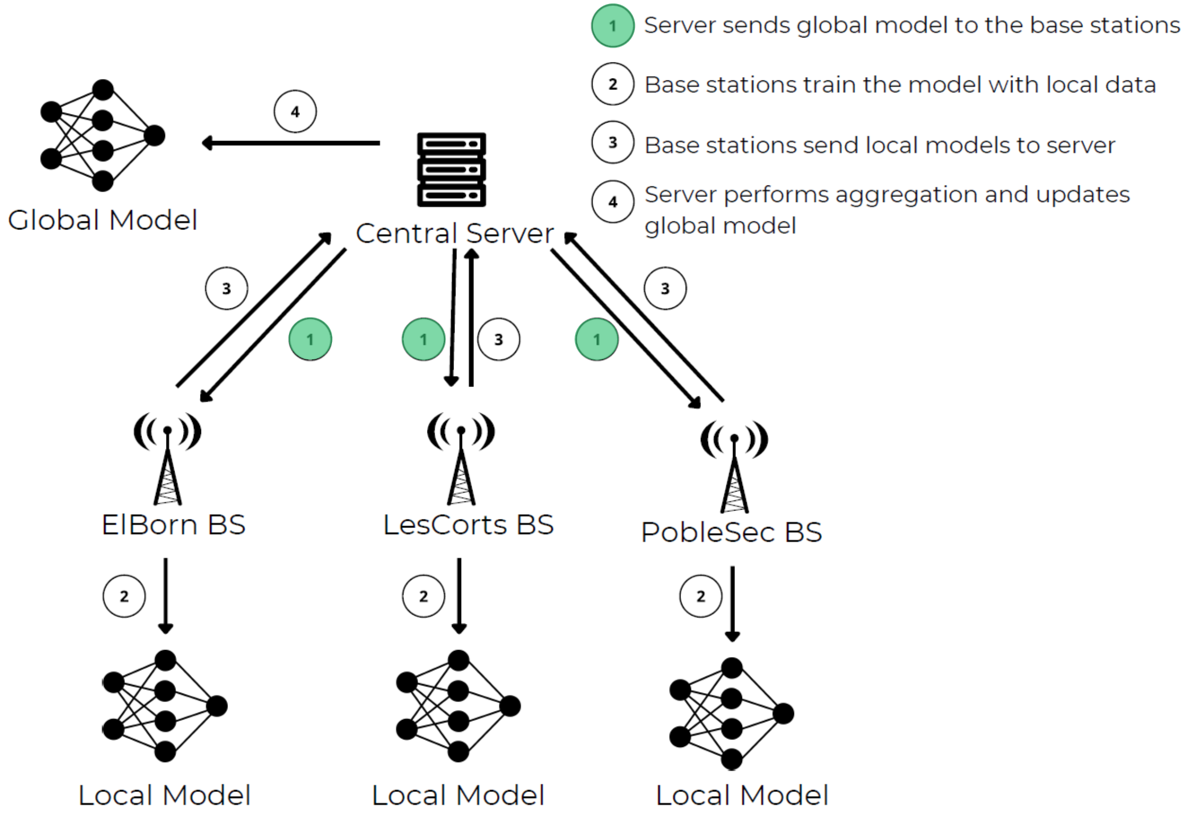


- Unique Characteristics
  - Feature and target distribution skew
  - Quantity skew
  - Temporal skew

Federated Learning

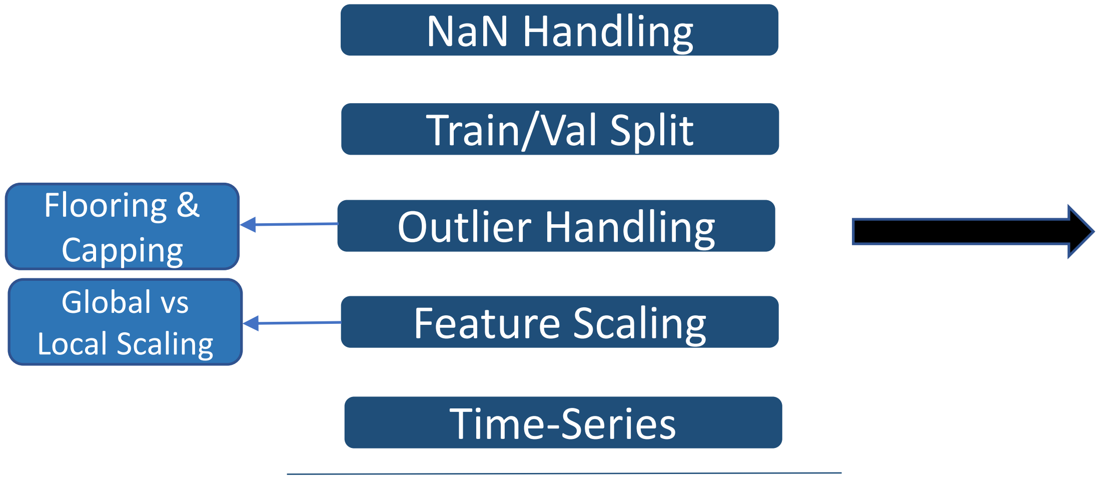


Pre-processing

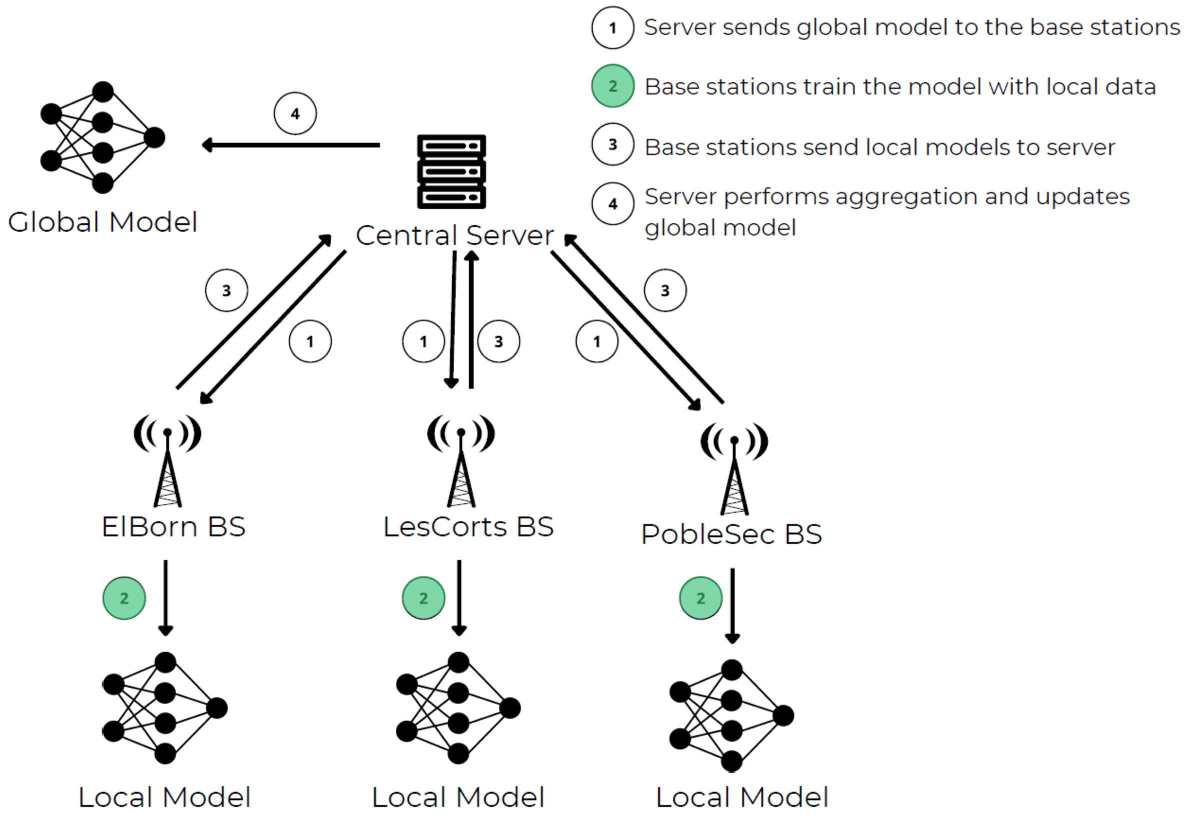


- Unique Characteristics
  - Feature and target distribution skew
  - Quantity skew
  - Temporal skew

Federated Learning



Pre-processing



- Unique Characteristics
  - Feature and target distribution skew
  - Quantity skew
  - Temporal skew

Federated Learning

NaN Handling

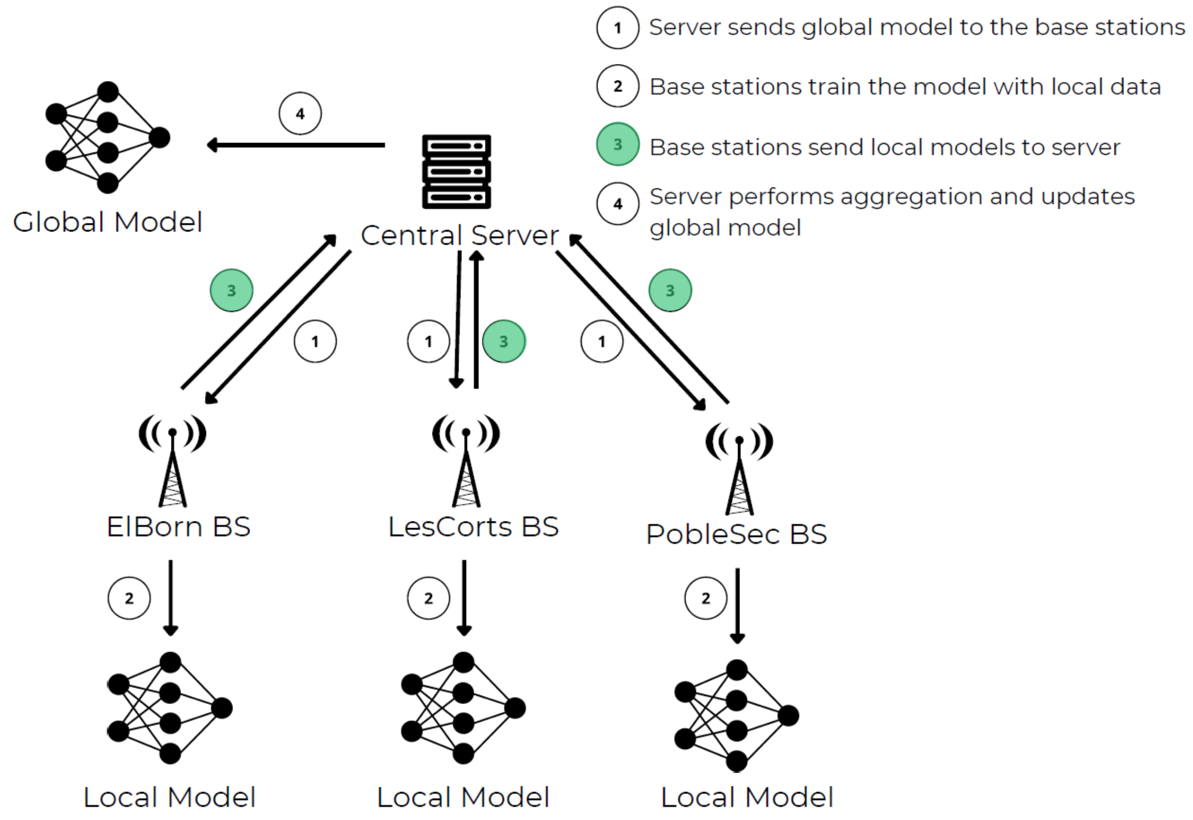
Train/Val Split

Outlier Handling → Flooring & Capping

Feature Scaling → Global vs Local Scaling

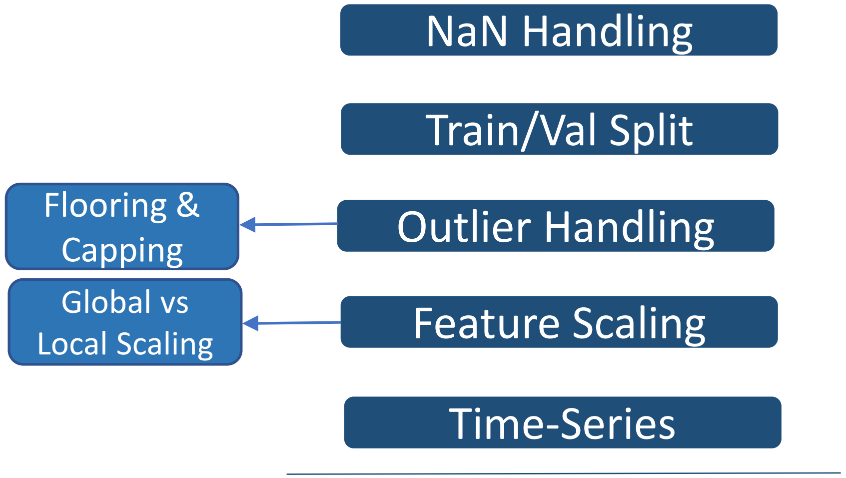
Time-Series

Pre-processing

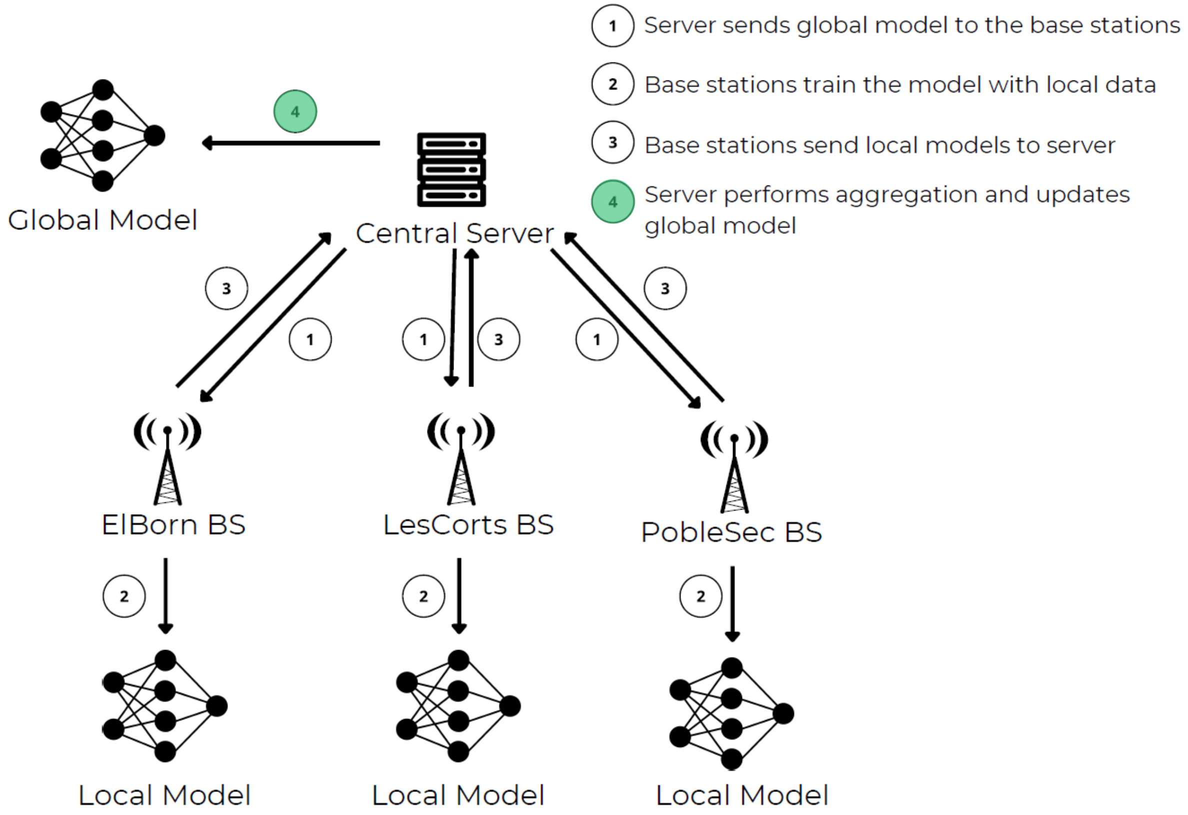


- Unique Characteristics
  - Feature and target distribution skew
  - Quantity skew
  - Temporal skew

Federated Learning



Pre-processing



- Unique Characteristics
  - Feature and target distribution skew
  - Quantity skew
  - Temporal skew

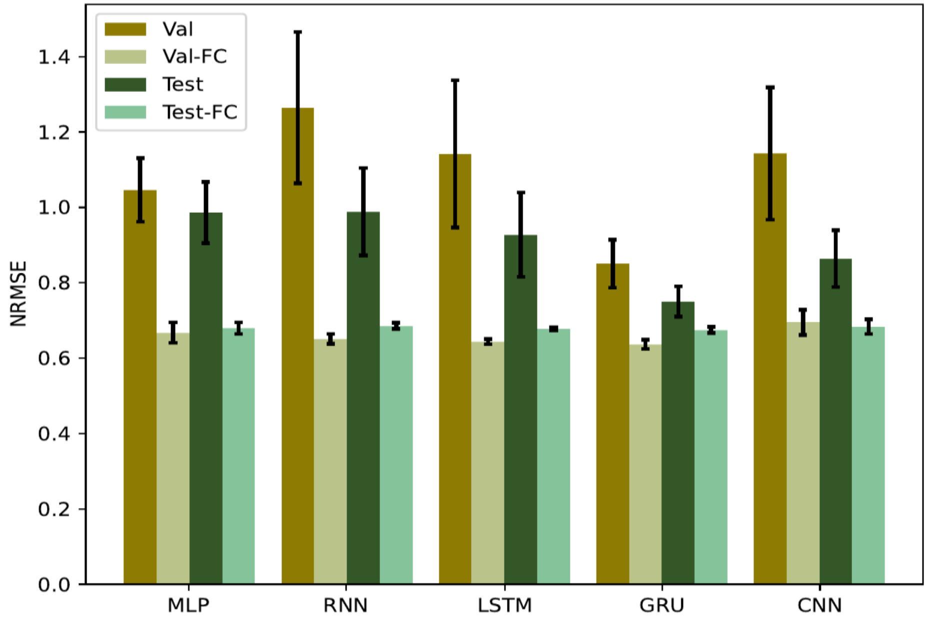
Learning Setting

- Deep Learning models
  - MLP, RNN, LSTM, GRU, 3D-CNN
- Aggregation Algorithms
  - Avg, Median, FedAvg, FedProx, FedAvgM, FedNova, FedAdagrad, FedYogi, FedAdam
- Evaluation Metrics
  - NRMSE regarding the uplink and downlink measurements

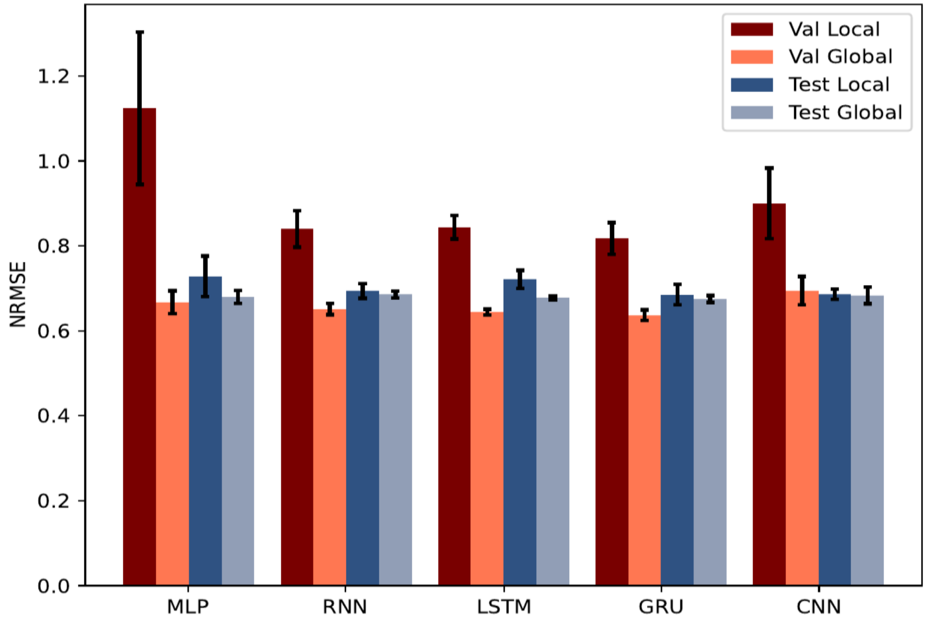
$$NRMSE = \frac{1}{\bar{y}} RMSE = \frac{1}{\bar{y}} \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}$$

  - Energy and CO_2 consumption
- Simulation
  - Python programming language with PyTorch [GitHub].

Results – Influence of Pre-processing

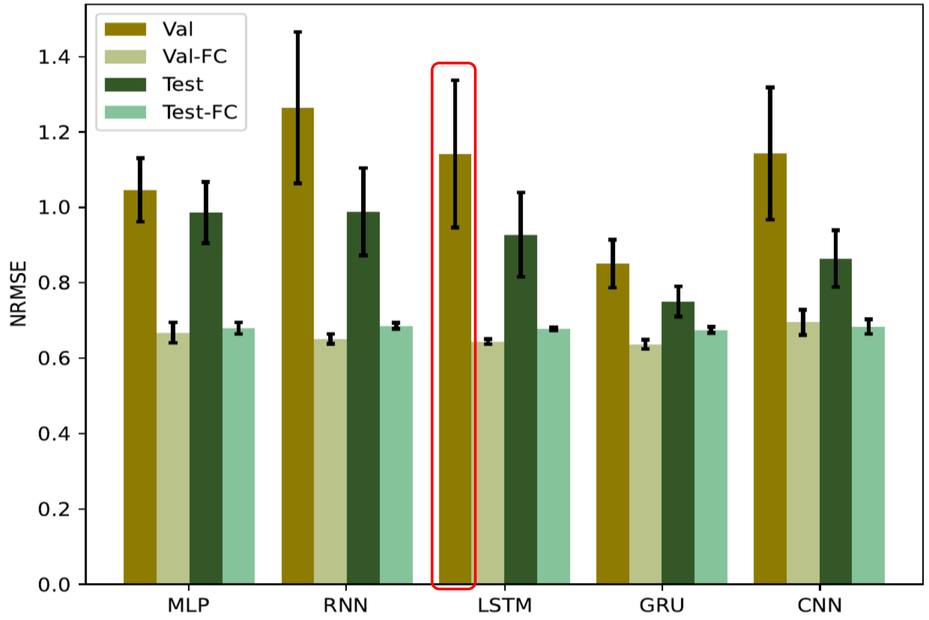


Averaged validation and test NRMSE using the raw data and with outliers handling.

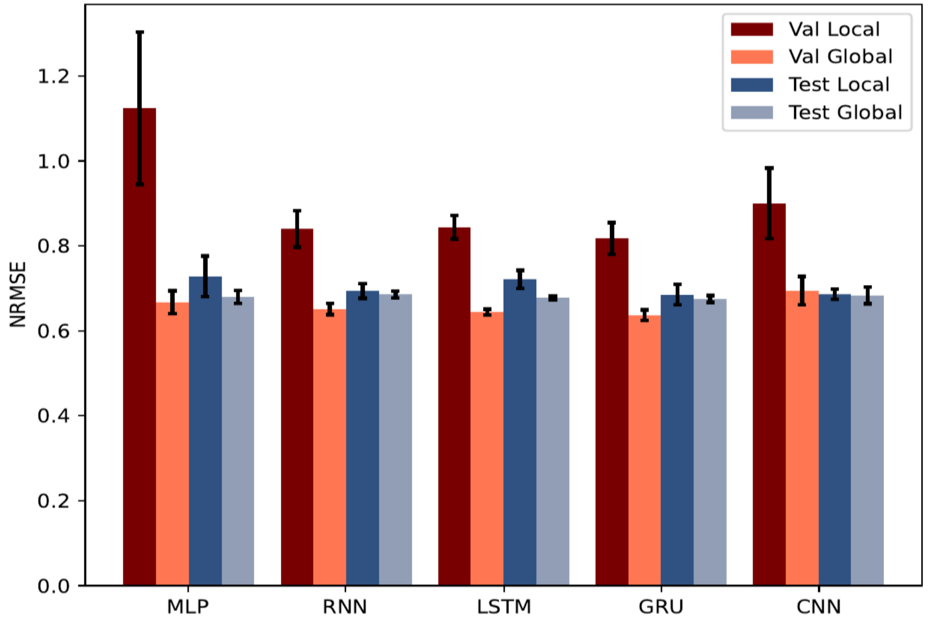


Averaged validation and test NRMSE using the global and local scaling.

Results – Influence of Pre-processing

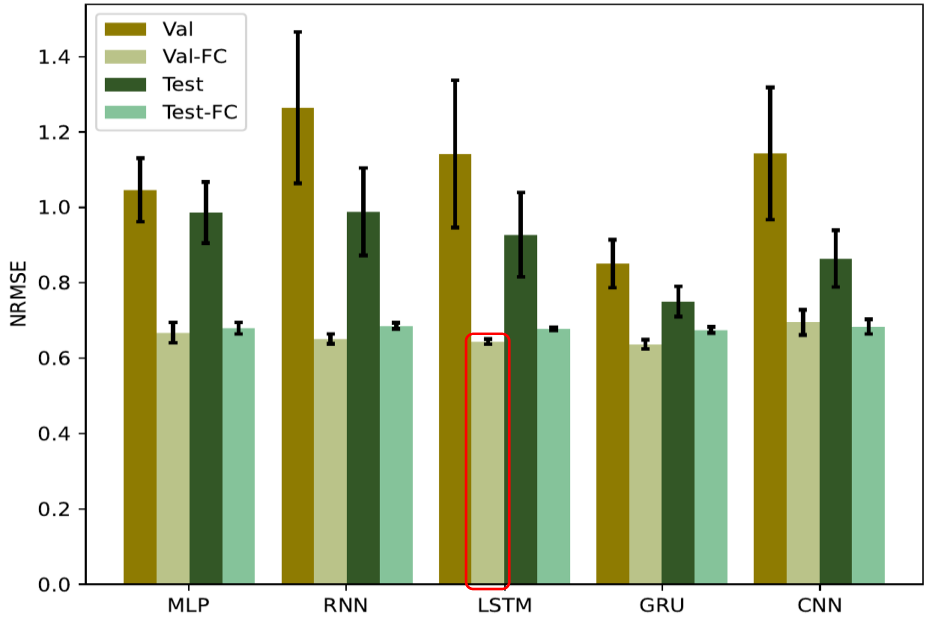


Averaged validation and test NRMSE using the raw data and with outliers handling.

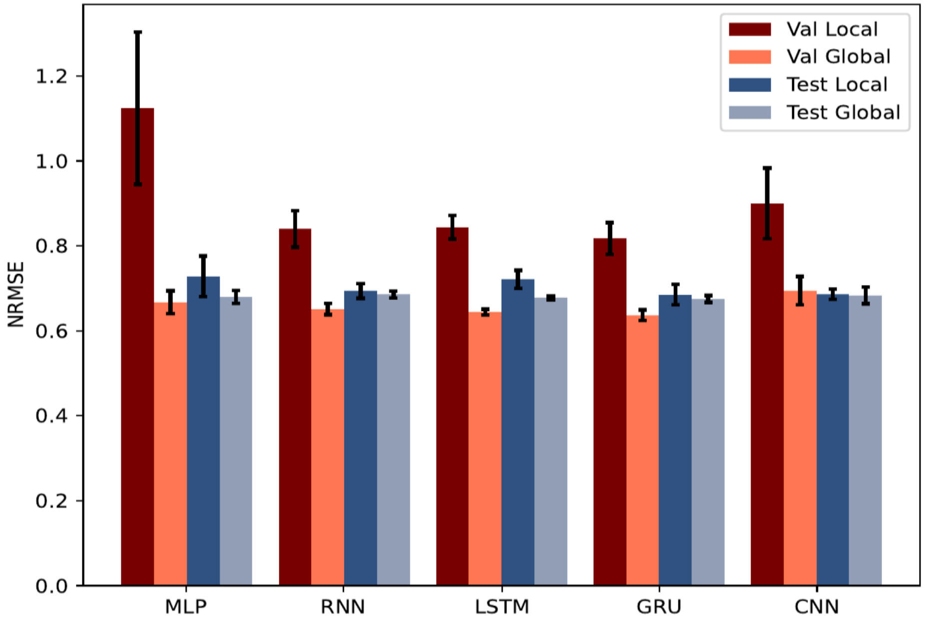


Averaged validation and test NRMSE using the global and local scaling.

Results – Influence of Pre-processing

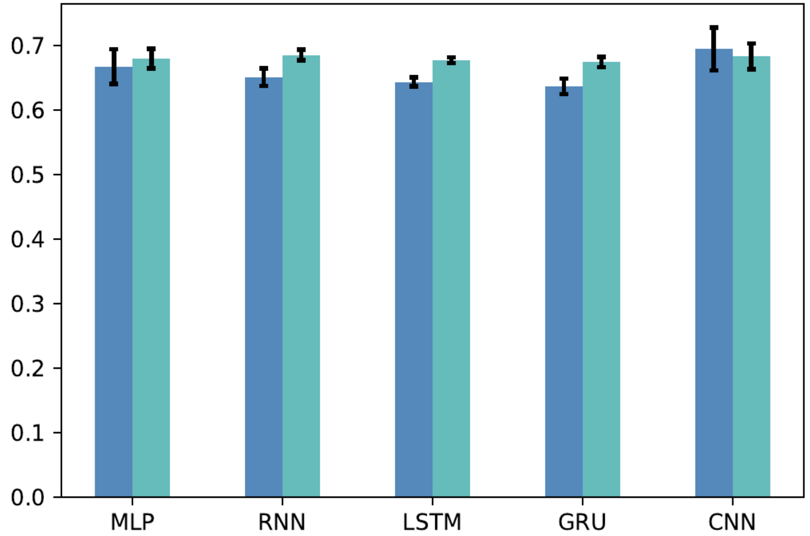


Averaged validation and test NRMSE using the raw data and with outliers handling.



Averaged validation and test NRMSE using the global and local scaling.

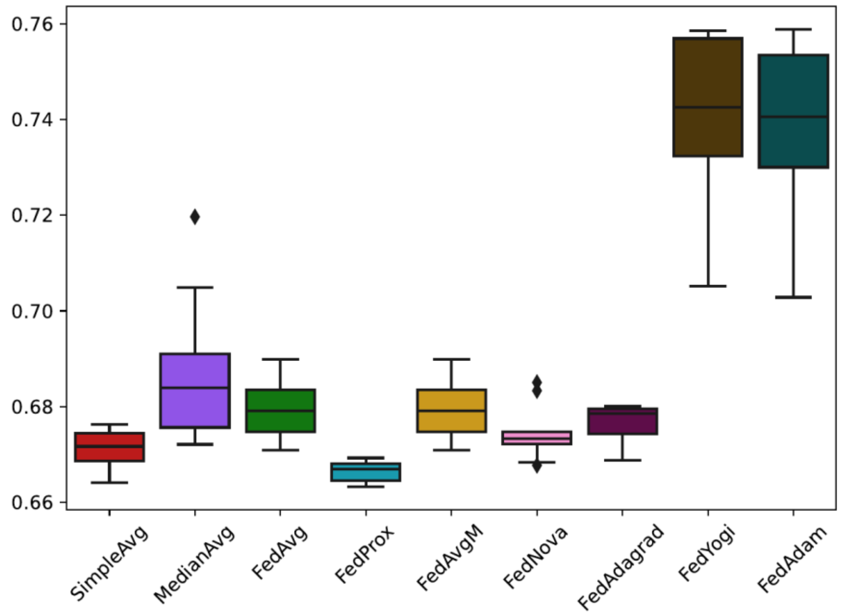
Results – Federated Learning



Averaged validation and test NRMSE per federated model.

Model	NRMSE (Cen)	NRMSE (FL)
MLP	<u>0.6705</u>±0.013	0.6797±0.015
RNN	0.6852±0.0085	0.6856±0.0081
LSTM	0.6927±0.0058	<u>0.6776</u>±0.0049
GRU	0.6734±0.0057	**0.6747**±0.0079
CNN	**0.6701**±0.0219	0.6836±0.0198

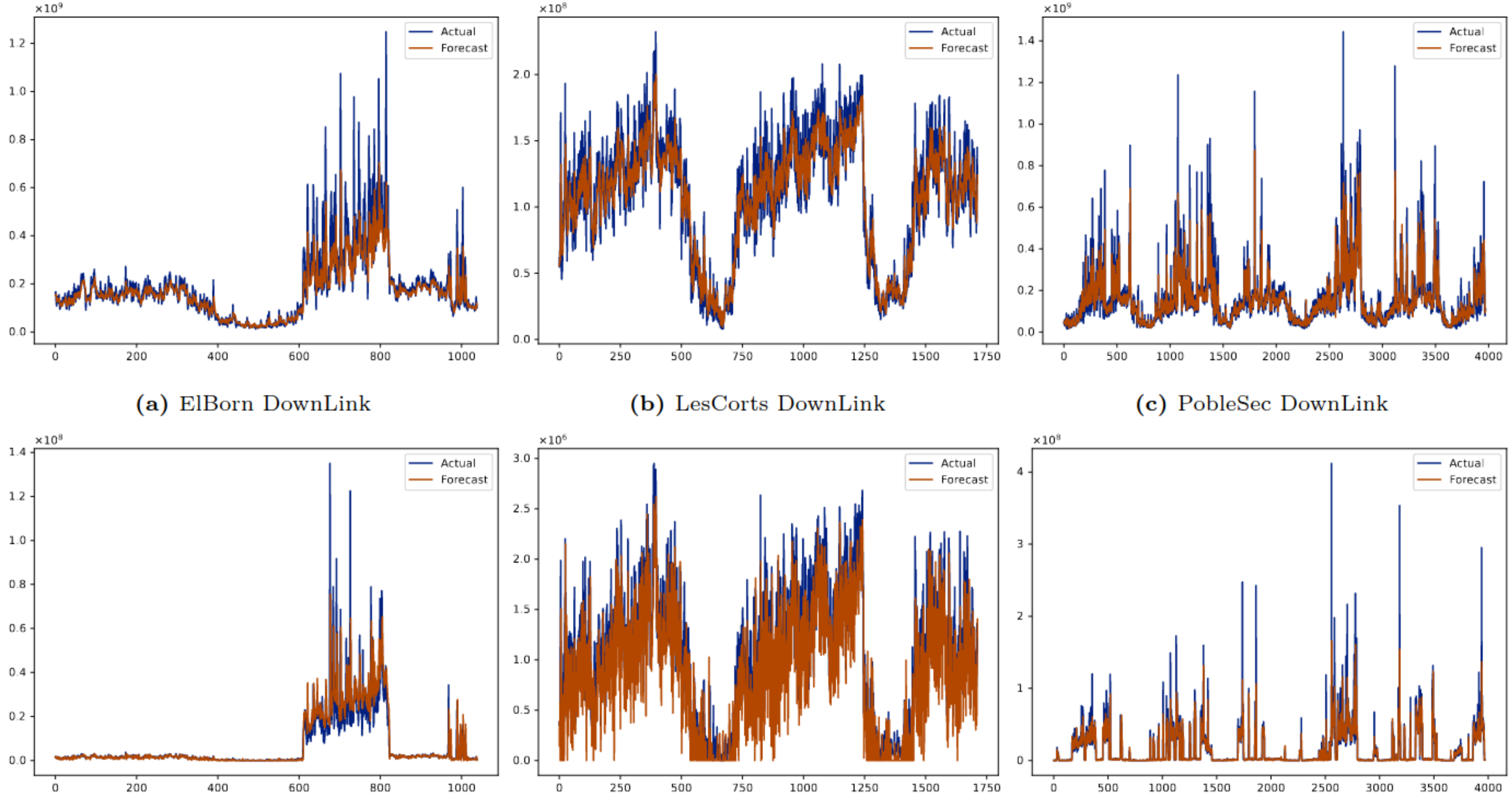
Results – Aggregation & Energy Consumption



Averaged test NRMSE per aggregation algorithm using the federated LSTM model.

Setting	Measure	MLP	RNN	LSTM	GRU	CNN
Centralized-1	Energy	0.0029	0.0032	0.0033	0.0033	0.0051
	CO_{2eq}	0.8414	0.9415	0.9788	0.9803	1.4975
Federated	Energy	0.0011	0.0012	0.00135	0.0012	0.0019
	CO_{2eq}	0.3216	0.3429	0.3975	0.3645	0.5640

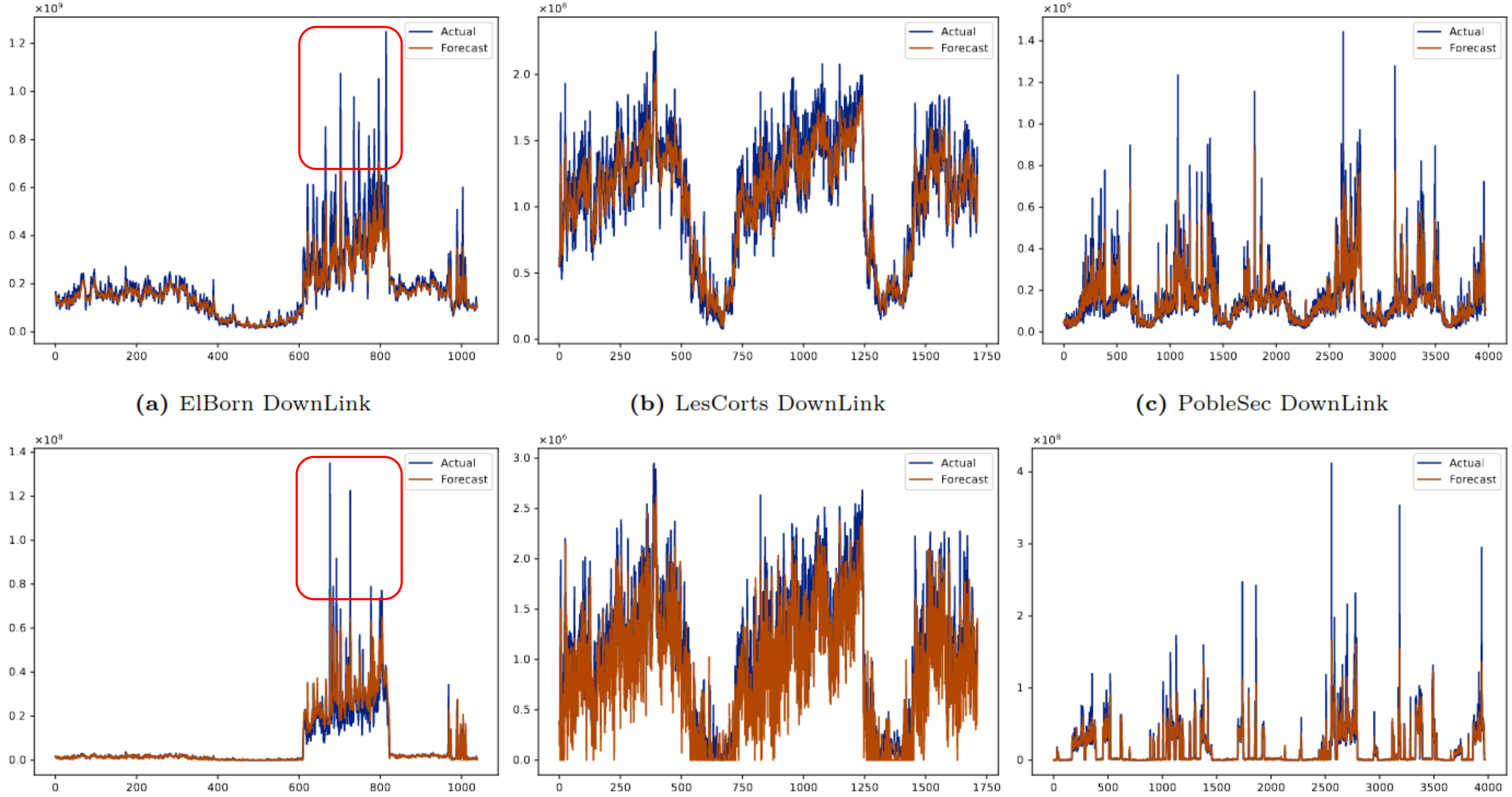
Results – Forecasting



(a) ElBorn DownLink (b) LesCorts DownLink (c) PobleSec DownLink

(d) ElBorn UpLink (e) LesCorts UpLink (f) PobleSec UpLink

Results – Forecasting



(a) ElBorn DownLink **(b)** LesCorts DownLink **(c)** PobleSec DownLink

(d) ElBorn UpLink **(e)** LesCorts UpLink **(f)** PobleSec UpLink

Conclusion

- Pre-processing hugely affects the forecasting error
  - Outlier handling leads to lower error
  - Global scaling results in higher accuracy than local scaling
- Federated Learning is on par with the centralized setting
  - Privacy
  - Dynamic execution
  - Fine-Tuning
- Aggregation algorithms do not heavily influence the learning performance
- Future Direction: Temporal Models & Aggregators

Thank you

Any questions?