Power Losses Estimation in Low Voltage Smart Grids

Electrical Engineering, Electronics and Automation Ph.D. Program

Madrid, June 23th 2022

Author
José Ángel Velasco Rodríguez

Advisor
Hortensia Amarís Duarte

uc3m | Universidad Carlos III de Madrid

Contents

1. Introduction

Context

Overall efforts intended to achieve a sustainable power system.

- *The actual environmental situation led us to the application of sustainable energy policies.*

EU Objectives by 2030 (*"Fit for 55 package"*) [1]

1. Increase **Energy Efficiency** in 32 % *(compared to 1990 levels).*
2. Increase **Renewable-based Generation** in: 32 %.
3. Reduce the **GreenHouse (GHG) Gas Emissions** in: 55 %.
   - *2050: Net zero GHG in the EU.*

[1] *European Commission launches proposals to reach 55% emissions reduction by 2030 [1]*

Motivation

How to meet the EU objectives?

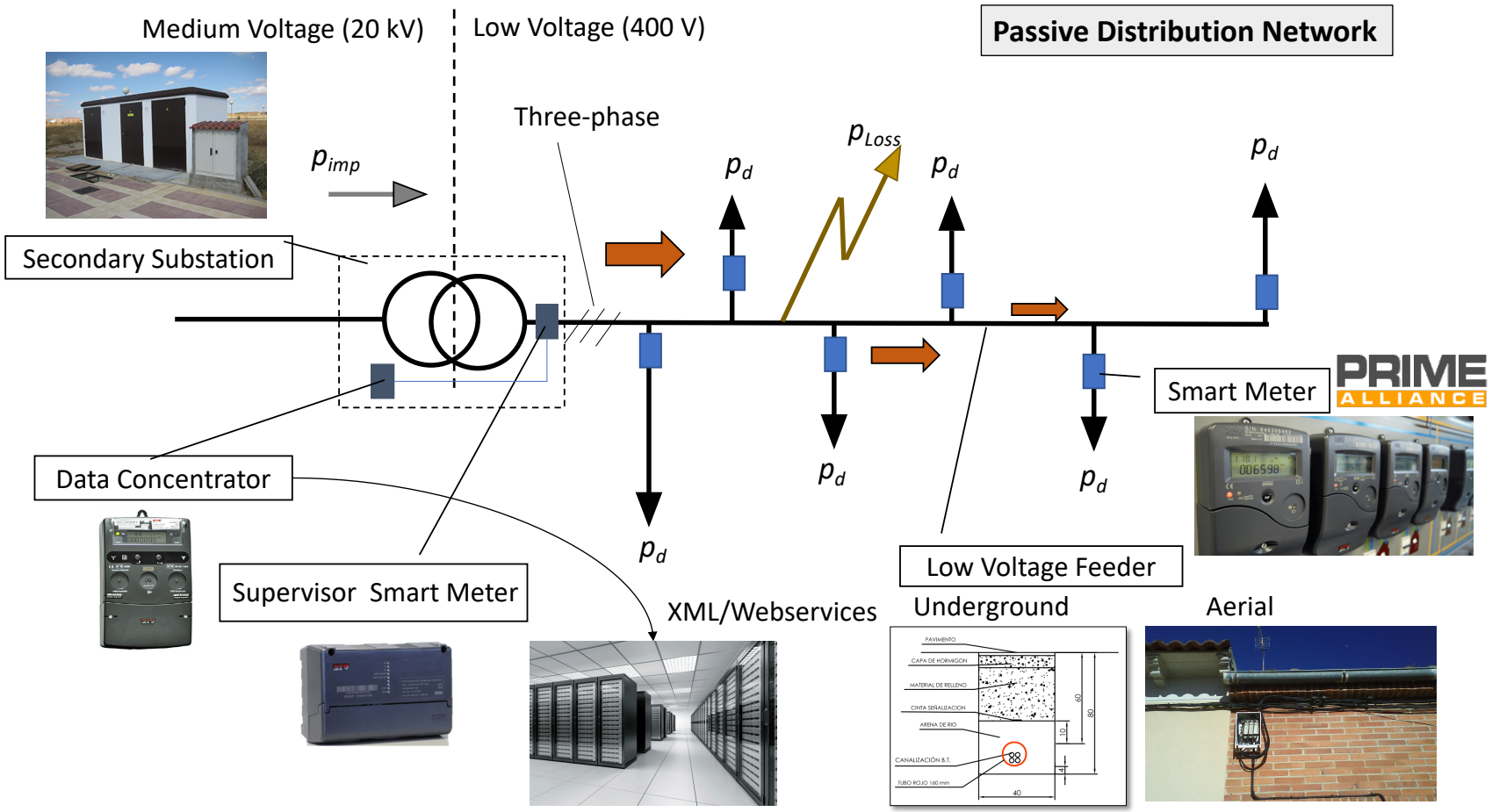
- ✓ **Reduce the energy lost in power supply** ⇒ *Energy Efficiency.*
  - Potential in LV distribution systems ⇒ *Higher power losses*
  - *Smart grids*: Digitalization of the LV network ⇒ Deployment of the smart metering infrastructure (smart meters).
    - Improve observability of the network.
- ✓ **Massive integration of Distributed Energy Resources (DERs)** ⇒ *Renewable-based generation & GHG emissions*:
  - Photovoltaic Panels (PVs)
  - Plug-in Electric Vehicles (PEVs)
  - Battery Energy Systems (BES)

Motivation

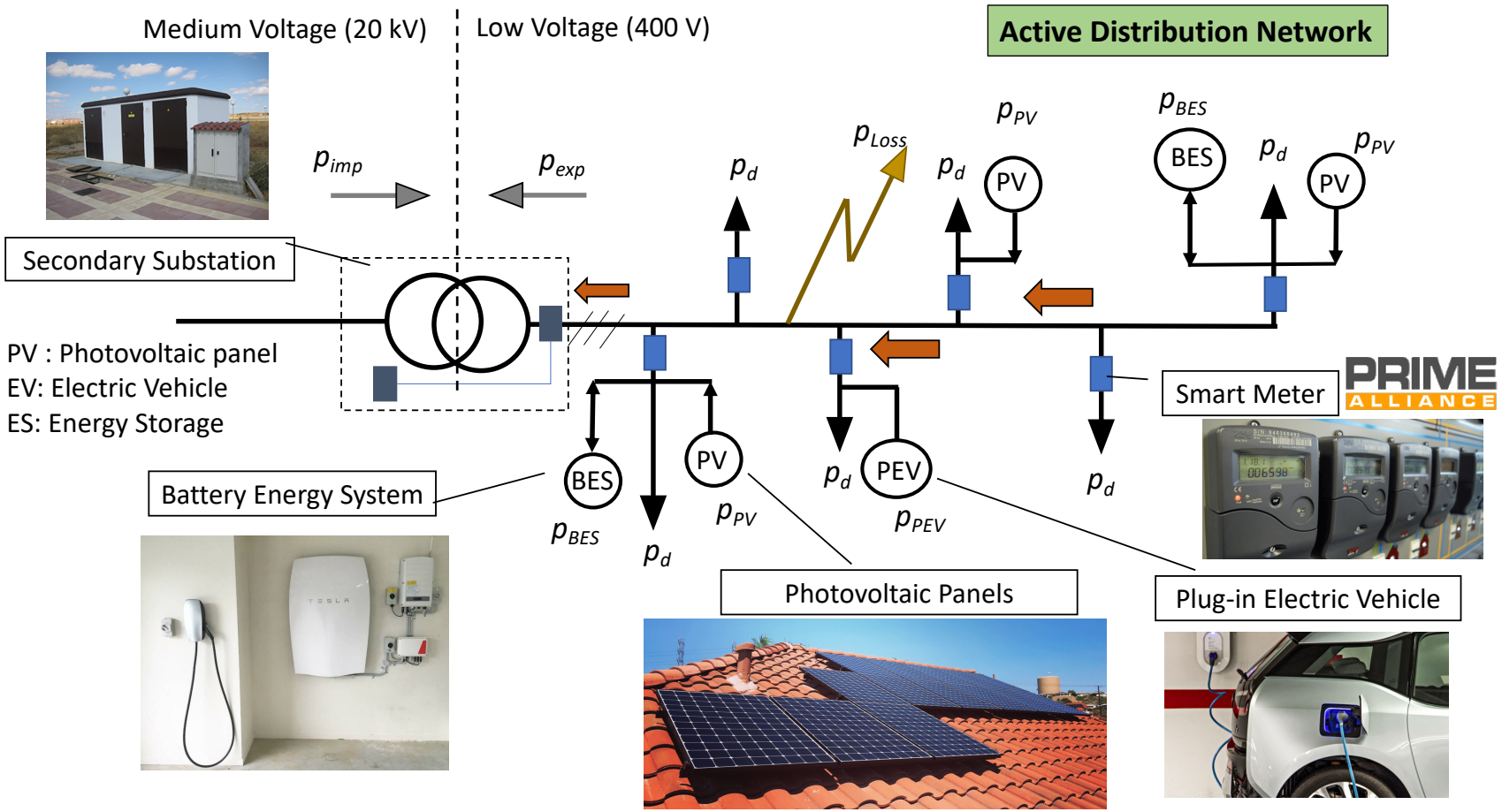
Challenges

- **Smart meter's massive roll-out in progress** *(100 % expected in 2026 in EU)* ⇒ Transition to smart grid not finished.
  - Presence of non-telemetered customers ⇒ Large customers ($P_{ctd} > 15\ kW$) ⇒ Load demand estimation ⇒ Uncertainty.
- **Technical Contingencies from massive DERs integration**
  - LV systems designed as "Passive" networks: Radial topology with telescopic cross sections ⇒ Contingencies.
    - Over-voltages, Overloadings & High values of power losses
    - Uncertainty associated ⇒ weather dependant & stochastic.
- **Unbalance operation of LV smart grids**
  - Presence of single-phase connections (Customers & DERs).
    - Power injections not uniform in the three phases.
    - ⇒ Underestimation of power losses (single-phase equivalent).

LV Smart Grid Feeder without DERs



LV Smart Grid Feeder with DERs



Thesis Objectives

1. **Power losses estimation in unbalanced LV smart grids under uncertainty.**
2. **Power losses estimation in large-scale distribution areas under uncertainty.**
3. **Power losses minimization in unbalanced smart grids under uncertainty**.

2. State of the art

State of the Art

- **Power losses estimation methods**.
  - Analytical equations → *Empirical Loss Factor Method (ELFM).*
  - Based on power flow equations.
  - Based on Machine Learning algorithms (AI).
- **Power losses minimization methods**.
  Mathematical Programming (OR) ⇒ decision-making
  - Deterministic Optimisation.
  - Non-deterministic Optimisation.
    - Probabilistic Optimisation.
    - Stochastic Optimisation (SO).
    - Robust Optimisation (RO).

Machine learning methods for power losses estimation

Machine learning methods for power losses estimation

Power Losses estimation from (supervised) learning:

$$\mathbb{Y} = \mathcal{F}_M(\mathbb{X})$$

where:

- $\mathbb{Y}$: Target (power losses).
- $\mathcal{F}_M$: Mapping function (ML model).
- $\mathbb{X}$: Predictors (smart grids features).

$$P_{LOSS,T} = \overbrace{\mathcal{F}_M(P_D, \widetilde{P}_D, \widetilde{P}_{PV}, \widetilde{P}_{PEV}, \widetilde{P}_{BES}, \mathcal{G})}^{UNBALANCED\ \ POWERFLOW}$$

Where $\mathcal{G}$ is the graph that represents the LV smart grid topology.

Smart Grids Features

Table 2.2: Features based on network operation for power losses estimation

Source	**Feature**
Smart Meters	Historical power consumption from smart meters
Smart Meters	Smart Meters penetration (ratio of non-telemetered customers)
Smart Meters	Customer consumption pattern
Network	Area of the Network (Residential, Industrial, Commercial)
Network	Network Topology characteristics
Network	Energy tariff's (influence in consumption patterns)
Network	Contractual customer power (influence in consumption magnitude)
Network	DG penetration and DERs presence
External	Calendar days
External	Weather data (temperature) and meteorological special events

Calendar days refer to: day of the week, holidays, non-working days, special days and events; Topology characteristics refers to: feeder length, number of nodes and number of ramifications

Supervised Machine Learning Models

- Linear regression models.
- Statistical-based models.
- Support Vector Machines (SVM) Regressor.
- K-Nearest Neighbours (KNN) Regressor.
- Artificial Neural Networks (ANN).
- Decision Trees (DTs).
- Ensembles methods: Random Forest, Boosted Trees.

Model Assessment

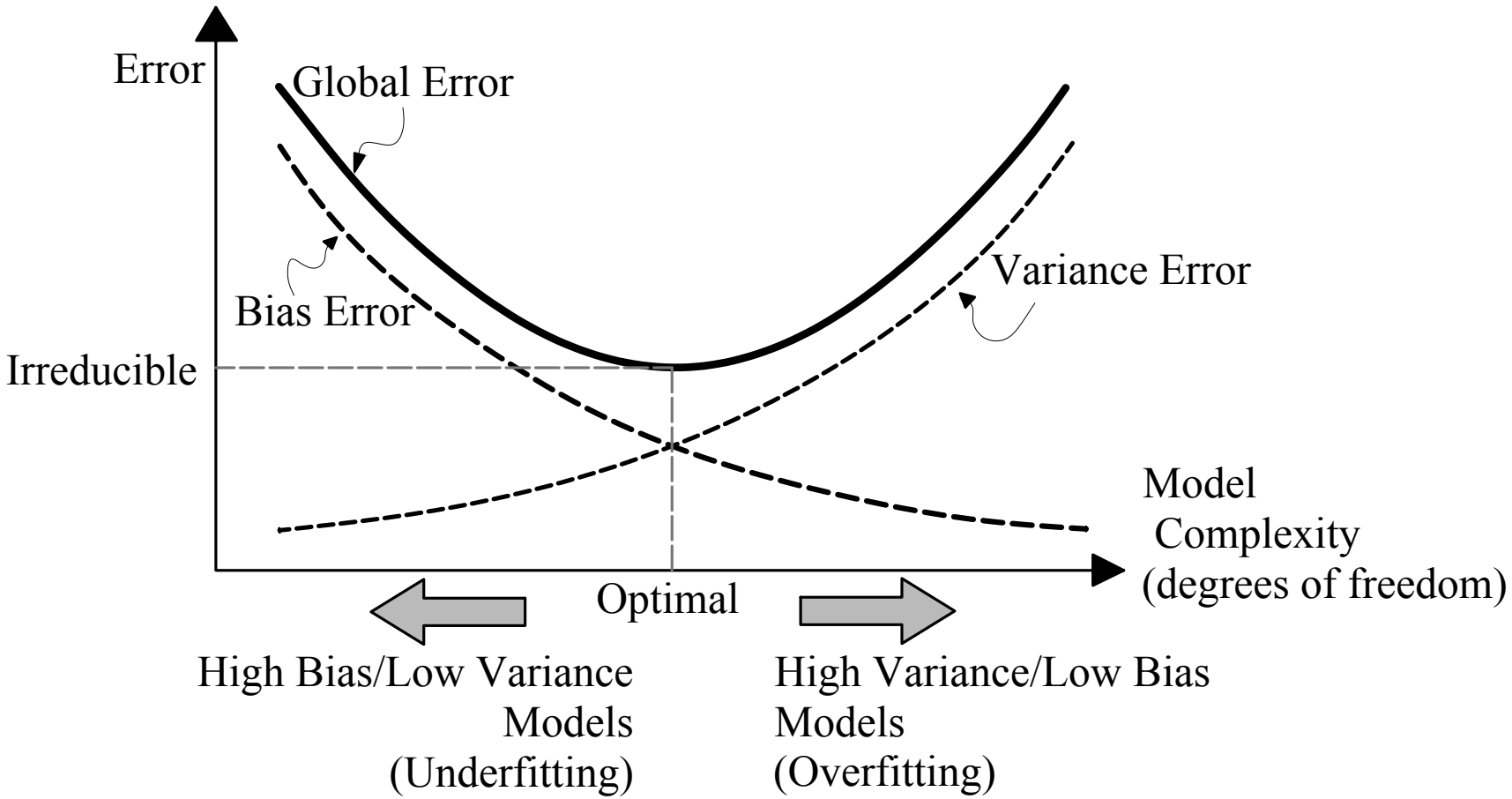
Prediction Error:

$$\epsilon = \widehat{\mathbb{Y}} - \mathbb{Y} = \epsilon_b + \epsilon_v + \epsilon_i$$

where

- $\widehat{\mathbb{Y}}$: Expected prediction (Known power losses value).
- $\mathbb{Y}$: Actual prediction (power losses prediction).
- ϵ_b: Bias error.
- ϵ_v: Variance error.
- ϵ_i: Irreducible error.

Model Assessment



Model Assessment

- **High-Bias / Low-Variance**: Rigid models → Strong assumptions → Linear-based models and statistical-based models.
- **Low-Bias / High-Variance**: Flexible models → Adaptative structure → ANNs, DTs and Ensembles.

Low-bias models provides flexibility to learn the power losses calculation function → **hyper-parameter tunning**.

Power losses minimization methods

Power losses minimization methods

Mathematical programming (OR) approach :
(unbalanced) Optimal Power Flow (MINLP)

$$\min_{\mathbf{x}\in X, \mathbf{y}\in Y} \quad OF(\mathbf{x},\mathbf{y})$$
$$\text{s.t.} \quad F(\mathbf{x},\mathbf{y},\mathcal{X}) = 0$$
$$G(\mathbf{x},\mathbf{y},\mathcal{X}) \leq 0$$

- $OF(\mathbf{x},\mathbf{y})$ Objective function → *Minimize power losses/ Cost*.
- $\mathbf{x} \in (0,1), \mathbf{y} \in R$ Decision variables vector → *Energy policies*.
- $\mathcal{X}$ State variable → *phase voltage*.
- $F(\bullet)$ Equality constraints → *unbalanced power flow equations*.
- $G(\bullet)$ Inequality constraints → *Technical constraints*.
- Uncertain parameters : $\tilde{p}^{p}_{PV,k}$, $\tilde{p}^{p}_{D,k}$, $\tilde{p}^{p}_{PEV,k}$ (Forecast)

Non-Deterministic Optimisation

- **Probabilistic Optimization** based sampling the uncertain parameters from their fitted PDFs and then solving the uOPF within a **Monte Carlo Simulation**.
- **Stochastic Optimization** Based on generate a plausible set of **scenarios** and the solve the uOPF through all of them:

$$\min_{\mathbf{x}_\omega X, \mathbf{y}_\omega} \sum_{\omega=1}^{N_\omega} \pi_\omega OF(\mathbf{x}_\omega, \mathbf{y}_\omega)$$

- **Robust Optimization** Based on define uncertain parameters as **uncertain variables** that takes values within a **robust set** $\Phi_{\mathbf{u}}$, controlled by an **uncertainty budget** $\zeta_{\mathbf{u}}$.

$$\mathbf{u} \in \Phi_{\mathbf{u}} = \left\{ \mathbf{u} \in R | u^L \leq u_i \leq u^U : \sum_{i=1}^{n} \frac{u_i - \bar{u}}{u_i} \leq \zeta_{\mathbf{u}} \right\}$$

Summary of previous works

Power Losses Estimation:

- Underestimation of power losses
  - Assuming balanced operation.
  - Assuming all customers metered.
  - Time-resolution of the measurements.
- No estimation in large-scale areas.

Power Losses Minimization:

- Previous works focused on Medium-High Voltage Networks.
- No consideration of unbalanced operation.
- No robust treatment of uncertainty sources.

3. Power losses estimation in unbalanced smart grids under uncertainty

Introduction

- **Uncertainty source** : Load demand → Focus on non-telemeterd customers.
- Only monthly energy consumption.
- Smart meters measurements available (including supervisor).
- **Goal**: Estimate the hourly power demand for non-telemetered.
- Produce synthetic intra-hour high-resolution load demand profiles for non-telemetered customers (Markov Process).
- Solve the unbalanced power flow and calculate complete power losses within a Monte Carlo Simulation.

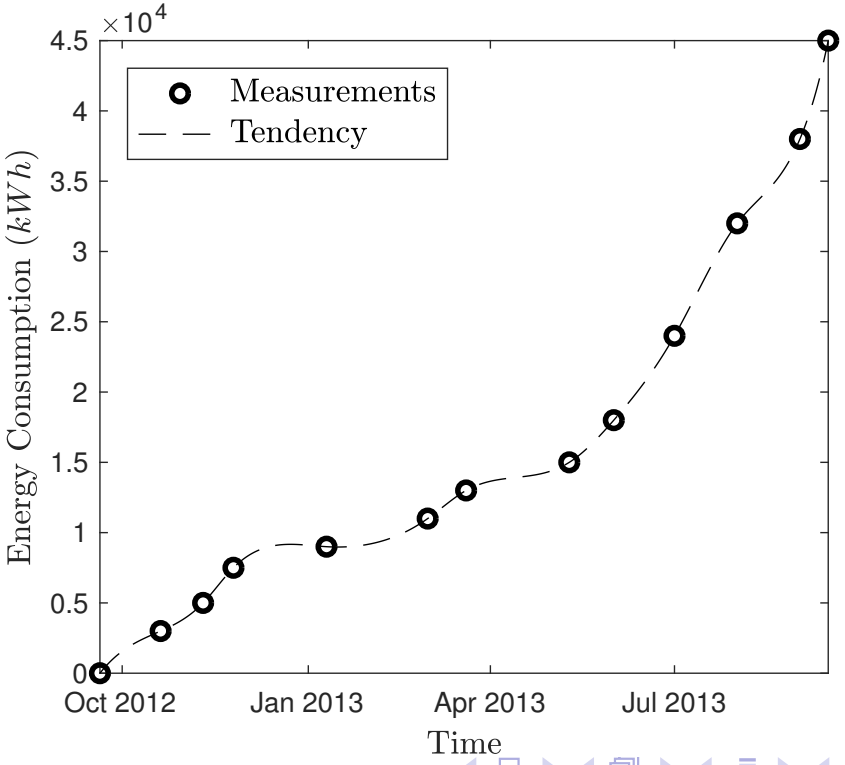
Non-telemetered customer's demand estimation

Top-down approach

1. **Upper Level**: Monthly Energy Consumption.
2. **Middle Level**: Daily Energy Consumption.
3. **Lower Level**: Hourly Power Demand Consumption.
4. **High-resolution**: Intra-hour load demand modelling.

Upper Level: Monthly Energy Consumption

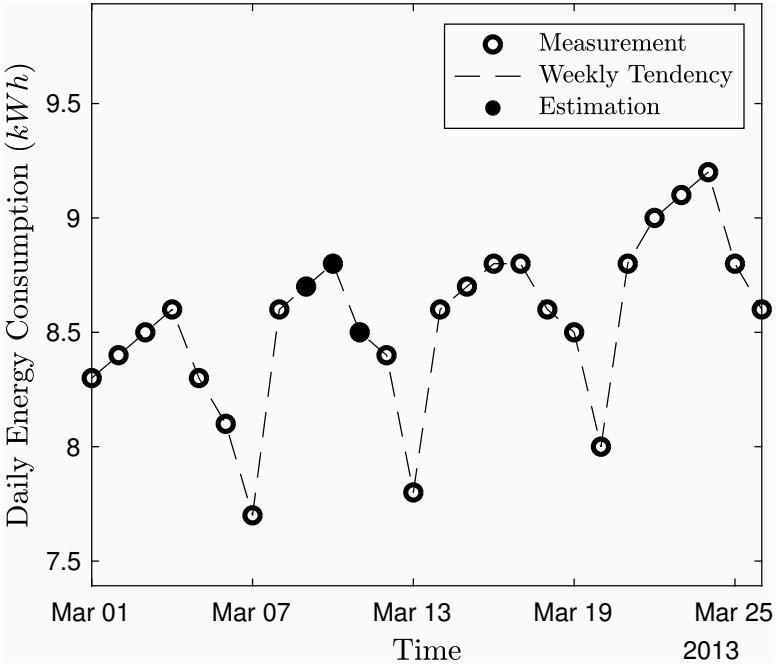
Estimate Energy Consumption Tendency (ECT) : $\widehat{E}_{i,m}$.
→ linear interpolation/Spline.



Middle Level: Daily Energy Consumption

Estimate daily energy consumption based on **pattern** $\widetilde{E}_{i,d}$.

$$\widehat{E}_{i,m} = \sum_{d \in T_h} w_{i,d} \cdot \widetilde{E}_{i,d} \ \ ,_{i,d} \in (0,1), \ \ \forall i \in \Omega_{c''}$$



Lower Level: Hourly Energy Consumption

Minimise hourly load demand estimation error:

$$\min_{P_{i,d,t}} : \; OF = \left\{ \overbrace{\sum_{i \in \Omega_{c'}} \sum_{d \in \Omega_d} \sum_{t \in T_d} \left(P_{i,d,t} - \widehat{P}_{i,d,t} \right)^2}^{Hourly\ Measurements\ Smart\ Meters} \right.$$

$$+ \overbrace{\sum_{i \in \Omega_{c''}} \sum_{d \in \Omega_d} \left(E_{i,d} - \widetilde{E}_{i,d} \right)^2}^{Daily\ Estimations\ Non-telemetered}$$

$$\left. + \overbrace{\sum_{d \in \Omega_d} \sum_{t \in T_d} \left(\sum_{i \in \Omega_c} P_{i,d,t} - \widehat{P}^S_{t,d} \right)^2}^{Total\ Hourly\ Deviation} \right\}$$

Lower Level: Hourly Energy Consumption

Subject to:

Non-telemetered customers energy consumption

$$\sum_{i \in \Omega_{c''}} \sum_{t \in T_d} P_{i,d,t} = \sum_{i \in \Omega_{c''}} E_{i,d}$$

Smart meter supervisor daily energy

$$\left(\sum_{i \in \Omega_{c''}} E_{i,d} + \sum_{i \in \Omega'_c} \widehat{E}_{i,d} \right) \leq \widehat{E}^S_d$$

Smart meter supervisor hourly energy

$$\sum_{i \in \Omega_c} \sum_{t \in T_d} P_{i,d,t} \leq \sum_{t \in T_d} \widehat{P}^S_{t,d}$$

High-resolution: Intra-hour load demand

First-Order **Markov chain** (X_t) of k states (sequence of states).

$$\mathbf{p}_{i,j}^{(t)} = Prob\Big(X_{t+1} = j | X_t = i\Big)$$

Transitional probability matrix $\mathcal{P}$.

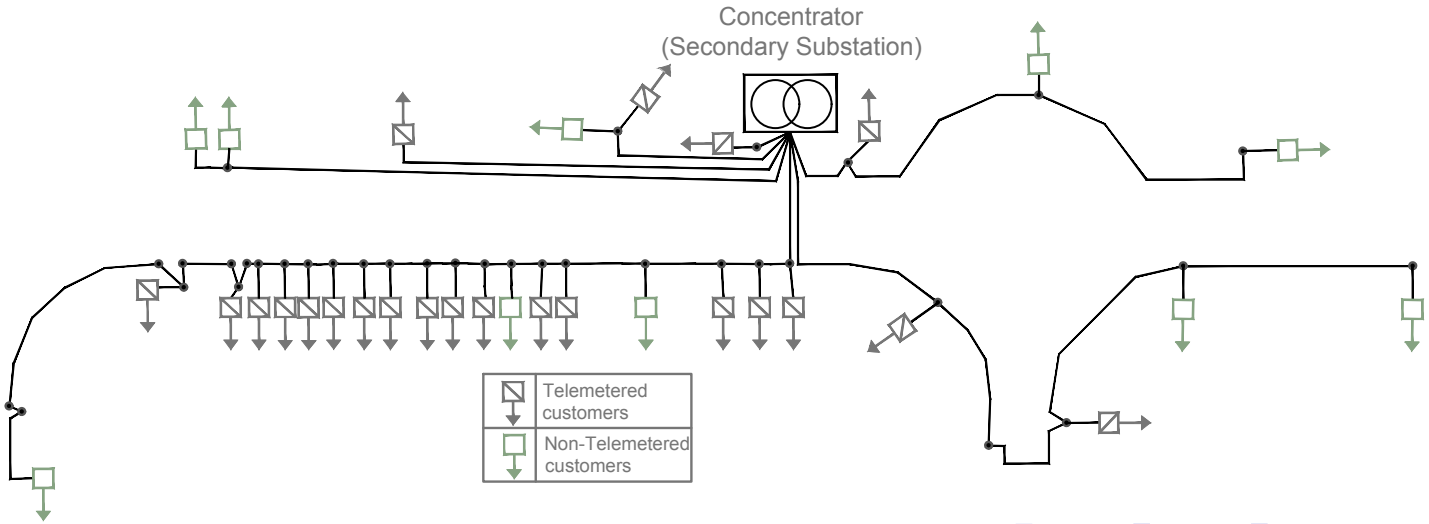
$$\mathcal{P} = \begin{pmatrix} \mathbf{p}_{1,1} & \cdots & \mathbf{p}_{i,k} \\ \vdots & \ddots & \vdots \\ \mathbf{p}_{k,1} & \cdots & \mathbf{p}_{k,k} \end{pmatrix} \in \mathbb{R}^{k \times k}$$

- $k = 60$ states (minutes, time resolution).
- Estimate PDFs by KDE method from historical hourly load demand $(P_{i,d,t})$.

Case Study

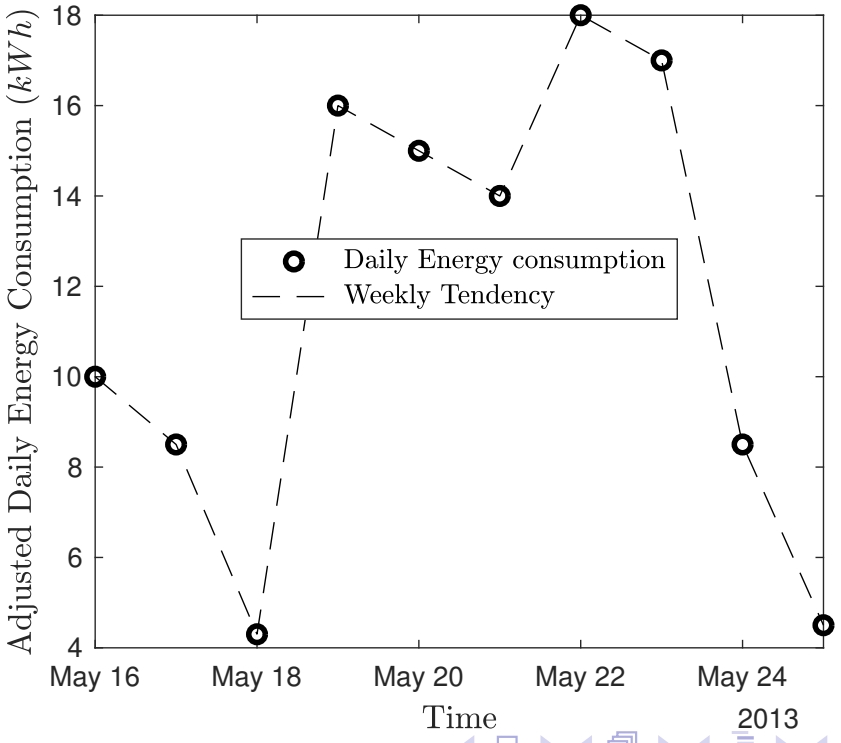
Case Study

- Spanish R&D demonstration project **OSIRIS** (*Optimizacion de la Supervision Inteligente de la Red de DIstirbuticion*).
- Secondary Substation: 680 kVA (Madrid, Spain).
- 8 LV feeders (commercial).
- 32 single-phase customers (10 non-telemetered).



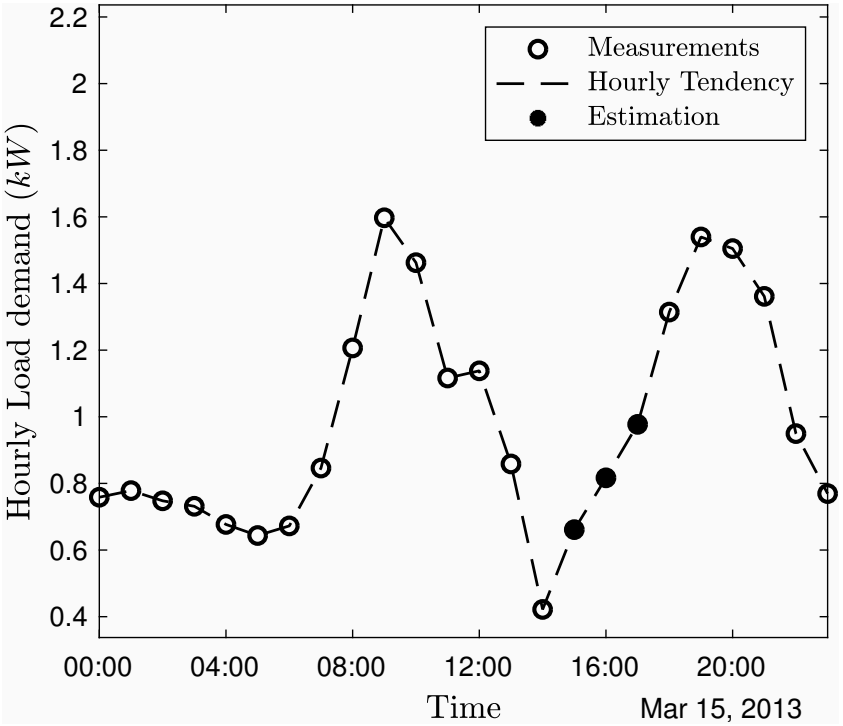
Obtain Daily Energy Consumption

1) Upper level: Build ECTs for non-telemetered customers.
2) Medium level: Adjust daily energy consumption according to the pattern.

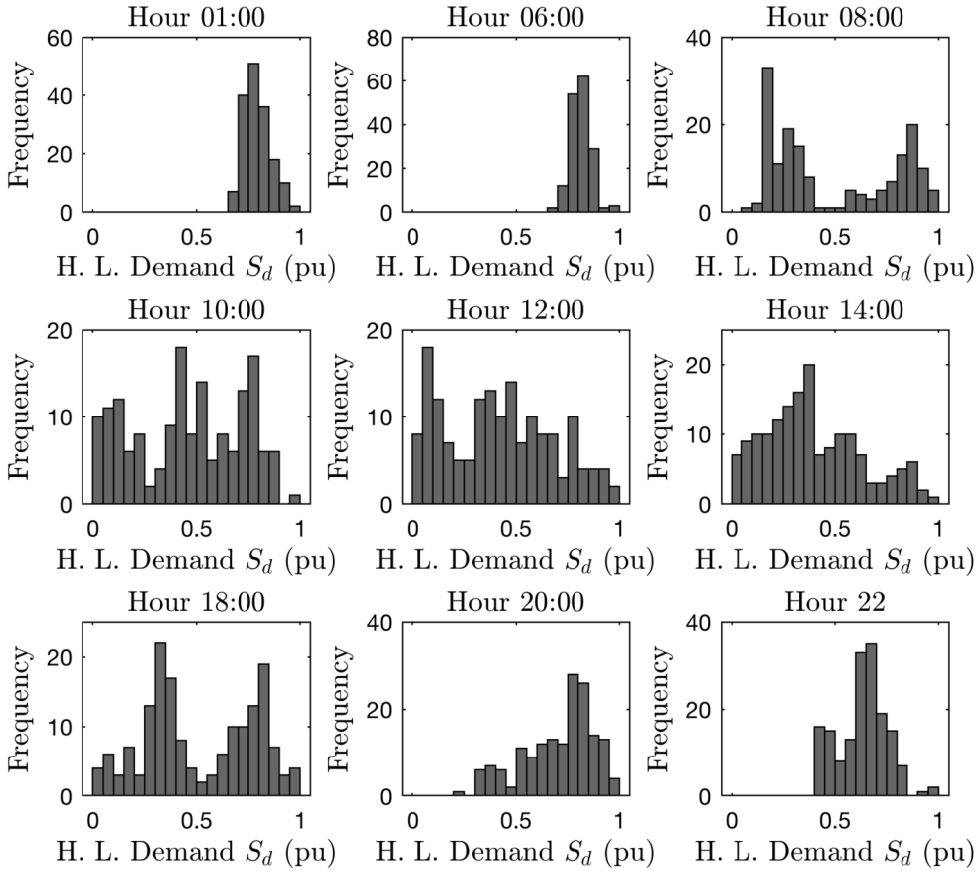


Obtain hourly Load demands

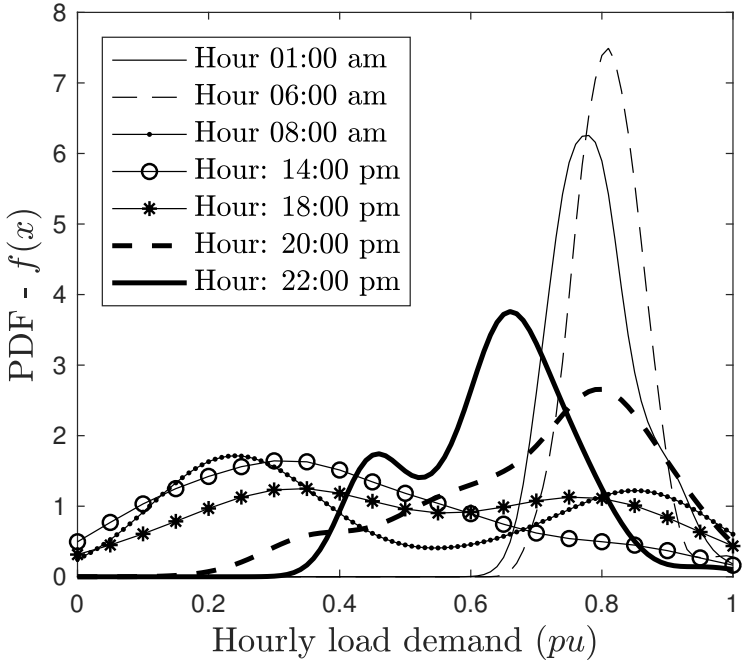
3) Lower level: Solving the NLP problem hourly load demand.

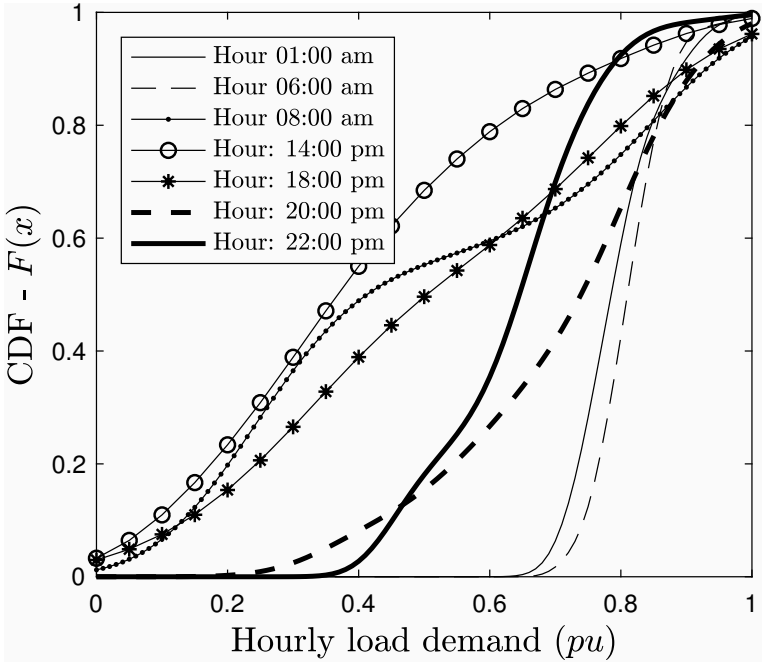


High resolution: Descriptive Statistics



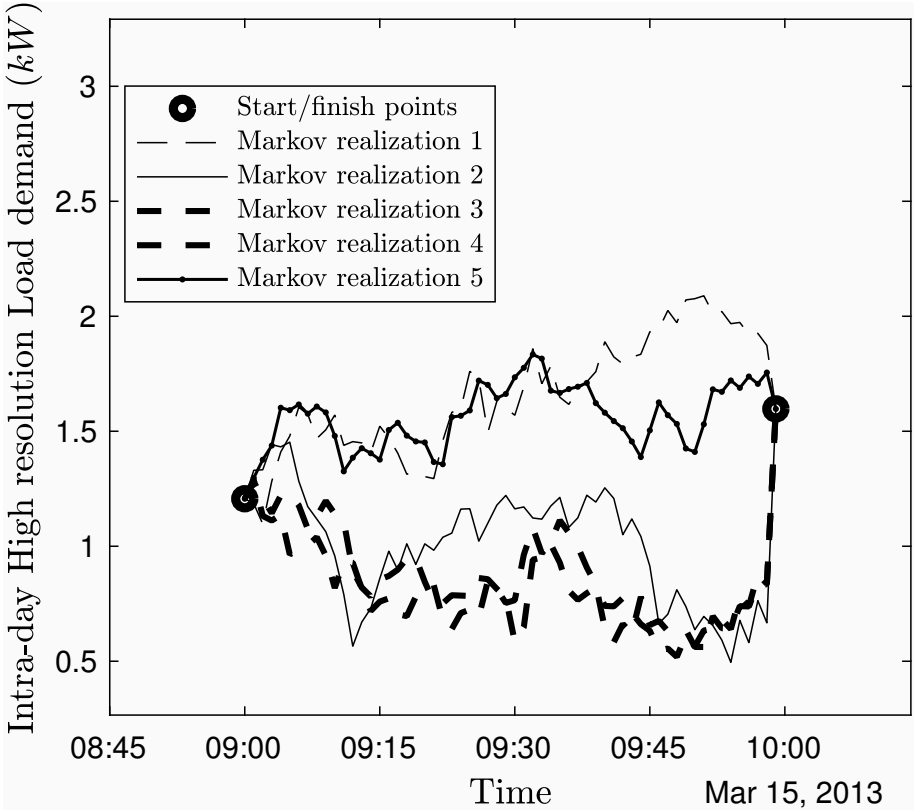
High resolution: Fit the hourly demand PDFs (KDE)





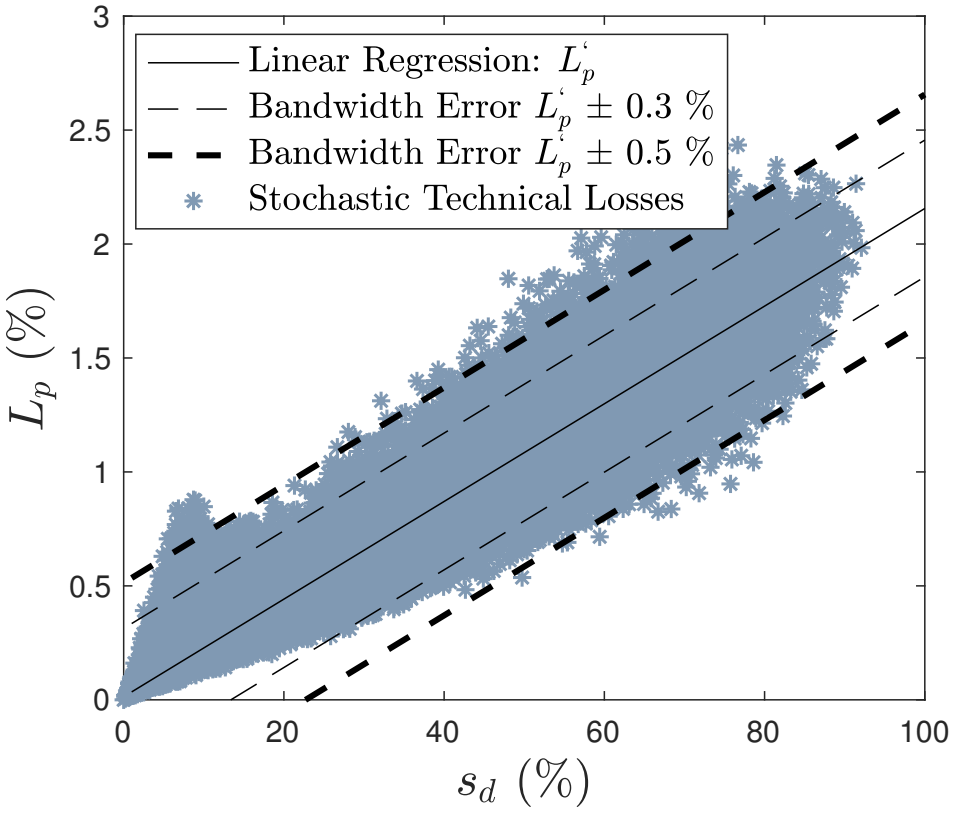
High resolution: Intra-hour load demand consumption

4) Transitional probability matrix $\mathcal{P}$ → Generate a set of realizations.



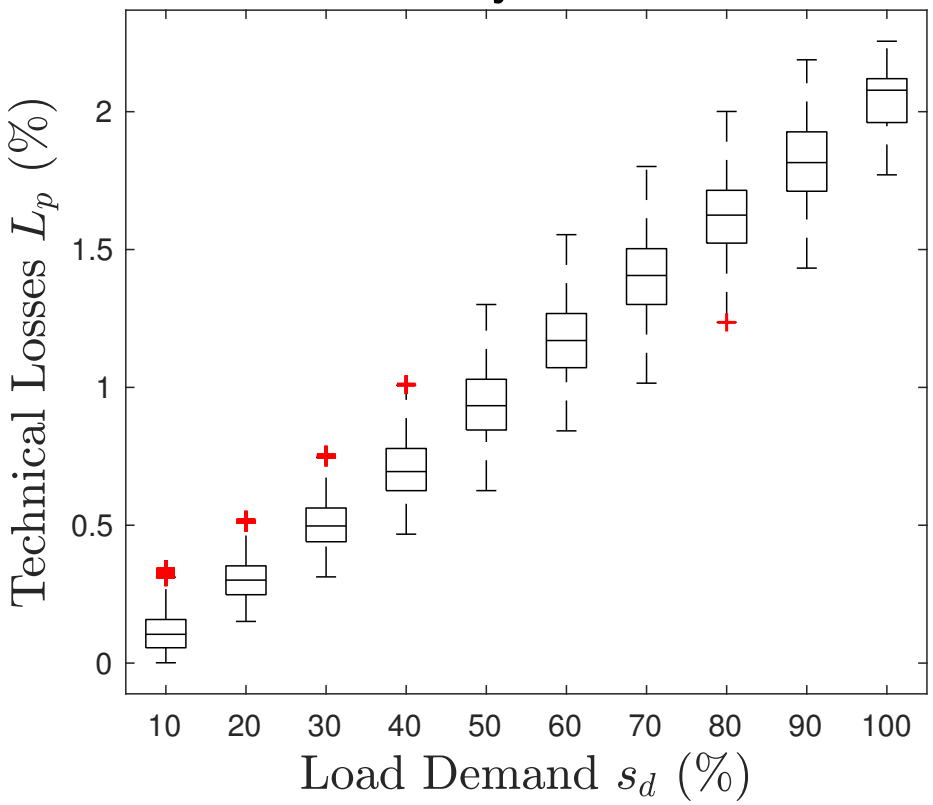
Stochastic Power Losses

Solve the unbalanced power flow for each realization and calculate relative losses $L_p = P_{LOSS}/S_D$.



Stochastic Power Losses

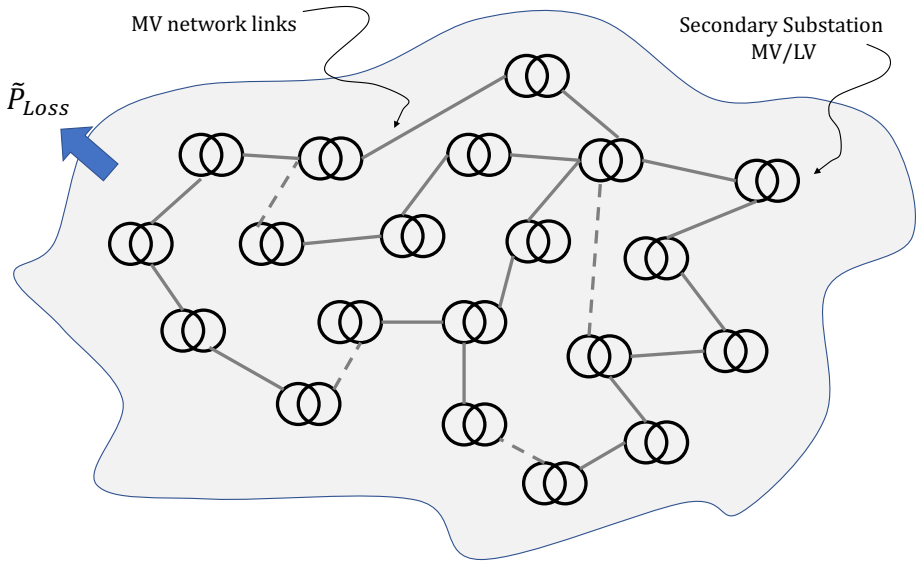
Characterize statistical behaviour of the power losses in the context of load demand uncertainty.



4. Power losses estimation in large-scale smart grids

Introduction

- Power losses in each LV feeder represent between 1-5 % of the power supply.
- Large distribution areas can comprise > 1000 smart grids, each one with 1-25 LV feeders.
- Potential of **energy efficiency** improve in large-scale distribution areas.



Introduction

- It requires flexible power losses estimator models → **scalability**.
- Exploit smart grid infrastructure → Large amount of data (Big data problem).
- Considering uncertainty (both in demand and generation sides).
- Considering unbalance operation.

Methodology

1. Data Collection.
2. Data Normalisation.
3. Features Extraction.
4. Feeder's Clustering & Representative feeders selection.
5. Deep Learning-based Power Losses Estimator model.

1) Data Collection

Data Collection

Table 4.1: Low-voltage network properties for representative feeder's clustering

Property	Definition	Data source	Range
$X^f_{1,p}$	Unbalance power level phase p	Smart Meters	$(0,1)$
X^f_2	Feeder loading level	Network	$(0,1)$
X^f_3	Smart meters deployment	Network	$(0,1)$
X^f_4	DG penetration level	Network	$(0,1)$
X^f_5	DG spatial location	Network	$(-0.5,0.5)$
X^f_6	Customers spatial location	Network	$(-0.5,0.5)$
X^f_7	Self-consumption ratio	Network	$(0,1)$
X^f_8	Impedance feeder path	Network	$(0,\infty)$
$X^f_{9,p}$	Demand MV WD phase p	Smart Meters	$(0,\infty)$
$X^f_{10,p}$	Demand SD WD phase p	Smart Meters	$(0,\infty)$
$X^f_{11,p}$	Demand MV NWD phase p	Smart Meters	$(0,\infty)$
$X^f_{12,p}$	Demand SD NWD phase p	Smart Meters	$(0,\infty)$
X^f_{13}	Ratio customers per feeder	Network	$(0,1)$
$X^f_{14,p}$	Ratio customers per phase	Network	$(0,1)$

MV: Mean Value, SD: Standard Deviation, WD: Working days, NWD: Non Working days

Data Collection

Feeder's characteristics data matrix X:

$$X = \begin{pmatrix} \mathcal{X}_{1,1} & \mathcal{X}_{1,2} & \dots & \mathcal{X}_{1,\tilde{p}} \\ \mathcal{X}_{2,1} & \mathcal{X}_{2,2} & \dots & \mathcal{X}_{2,\tilde{p}} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{X}_{n,1} & \mathcal{X}_{n,2} & \dots & \mathcal{X}_{n,\tilde{p}} \end{pmatrix} \in \mathbb{R}^{n \times \tilde{p}}$$

- n no. feeders.
- $\tilde{p}$ no. feeder's characteristics.

2) Data Normalization

Data Normalization

Normalised Feeder's characteristics data matrix M:

$$M = \begin{pmatrix} x_{1,1} & x_{1,i} & \cdots & x_{1,\tilde{p}} \\ x_{2,1} & x_{2,i} & \cdots & x_{2,\tilde{p}} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,i} & \cdots & x_{n,\tilde{p}} \end{pmatrix} \in \mathbb{R}^{n \times \tilde{p}}$$

$$x_{i,j} = \frac{\mathcal{X}_{i,j} - \mathtt{min}\{\mathcal{X}_j\}}{\mathtt{max}\{\mathcal{X}_j\} - \mathtt{min}\{\mathcal{X}_j\}} \in (0,1), \ \forall i \in (1,\ldots,n), \forall j \in (1,\ldots,\tilde{p})$$

3) Feature Extraction

Feature Extraction: Principal Component Analysis (PCA)

- Reduced set of relevant features $\widetilde{M}$, $s < \tilde{p}$ (i.e the number of components).

$$\widetilde{M} = M \cdot \widetilde{U} \in \mathbb{R}^{n \times s}$$

- Projection matrix $\widetilde{U} \in \mathbb{R}^{\tilde{p} \times s}$ composed by $\widetilde{U}_k \in \mathbb{R}^{\tilde{p} \times 1}, k = (1, \ldots, s)$ projections vectors (columns).
- (s) chosen by defining the total variability to be captured: $\eta(\%)$.

$$\underset{s}{\operatorname{argmin}} \left\{ \frac{\sum_{k=1}^{s} \lambda_k}{\sum_{k=1}^{\beta_\lambda} \lambda_k} > \eta(\%) \right\}$$

λ_k: k^{th} eigenvalue associated of covariance matrix $S \in \mathbb{R}^{\tilde{p} \times \tilde{p}}$ of M.
β_λ: No. of eigenvalues of S.

Feature Extraction: Principal Component Analysis (PCA)

Obtain the projections solving maximizing the variability captured :

$$\begin{aligned} \underset{\widetilde{U}_k}{\text{argmax}}: \quad & \left\{ \widetilde{U}_k^T \cdot S \cdot \widetilde{U}_k \right\}, \ \forall k \in (1, \ldots, \beta_k) \\ \text{s.t.} \quad & \widetilde{U}_k^T \cdot \widetilde{U}_k = 1, \ \forall k \in (1, \ldots, \beta_k) \\ & \widetilde{U}_k^T \cdot \widetilde{U}_{k'-1} = 0, \ \forall k' \in (2, \ldots, \beta_k) \end{aligned}$$

Applying Lagrangian multipliers procedure is equivalent to find **eigenvector** associated to the kth large **eigenvalue** λ_k of covariance matrix S.

$$(S - \lambda_k \cdot I) \cdot \widetilde{U}_k = 0$$

4) Feeder's Clustering & Representative Selection

Feeder's Clustering

- Split the reduced feeder's data matrix $\widetilde{M} \in \mathbb{R}^{n \times s}$ into $\hat{k}$ separated subsets (i.e. feeders' clusters).
- **Clustering problem** → `K-means++` (unsupervised learning).

$$C_r = \{c_1, \ldots, c_{n_r}\} \subseteq \widetilde{M}, \ \forall r \in (1, \ldots, \hat{k})$$

- Composed by three iterative steps:
  1. Centroids initialization.
  2. Assignment.
  3. Evaluation.

Feeder's Clustering: Initialization

A) Centroid Initialization:

A.1) Centroids inizialization (random):

$$c_r^{*(0)} = c_i$$

A.2) Compute Euclidean distance $\mathcal{D}$ between each feeder data sample and the centroid:

$$\mathcal{D}(c_i, c_r^{*(0)}) = \left|\left|c_i - c_r^{*(0)}\right|\right|^2$$

A.3) Updating of the centroid $c_{r+1}^{*(0)}$ as the one that offers the highest probability according to:

$$Prob\left(c_{r+1}^{*(0)} = c_j\right) = \mathcal{D}(c_j, c_r^{*(0)})^2 / \sum_{k,j\neq k} \mathcal{D}(c_k, c_r^{*(0)})$$

- Repeat A.2 and A.3 until $\hat{k}$ centroids has been chosen.

Feeder's Clustering: Assignment

B) Assignment Step

B.1) Assign each feeder data sample to the closest centroid cluster by minimising the Mean Square Error (MSE) according to:

$$MSE\left(C_1, \ldots, C_{\hat{k}}\right) = \sum_{r=1}^{\hat{k}} \frac{1}{n_r} \sum_{i=1}^{n_r} \left\| c_i - c_r^{*(0)} \right\|^2$$

B.2) Recalculate centroids as the average value:

$$c_r^{*(0)} = \frac{1}{n_r} \sum_{c_i=1}^{n_r} c_i, \ \ \forall r \in (1, \ldots, \hat{k})$$

B.3) Repeat B.1)-B.2) until convergence:

$$c_r^{*(q)} - c_r^{*(q-1)} \leq \varepsilon, \ \ \forall r \in (1, \ldots, \hat{k})$$

Feeder's Clustering: Evaluation

C) Clustering Evaluation

Metric for cluster quality: **Global Silhouette (GS)** coefficient:

$$GS = \frac{1}{\hat{k}} \sum_{r=1}^{\hat{k}} \frac{1}{n_r} \sum_{i=1}^{n_r} \mathbb{S}(c_i)$$

where $\mathbb{S}(c_i) \in (0,1)$ is the **Silhouette** of each feeder sample c_i:

$$\mathbb{S}(c_i) \approx 1 - \frac{a(c_i)}{b(c_i)}, \ \forall i \in (1, \ldots, n)$$

- $a(c_i)$ average distance between the feeder sample c_i and all the feeder's cluster.
- $b(c_i)$ smallest average distance between the feeder sample c_i and all feeders.

Feeder's Clustering: Representative Feeders Selection

Representative feeders $\tilde{c}_r$ are chosen as the feeder sample closest to the centroid within each feeders' cluster by means of the minimum Euclidean Distance to the centroid c_r^*:

$$\tilde{c}_r = \arg\min_{c_i \in C_r} \left\{ \|c_i - c_r^*\|^2 \right\}, \ \forall r \in \left(1, \ldots, \hat{k}\right)$$

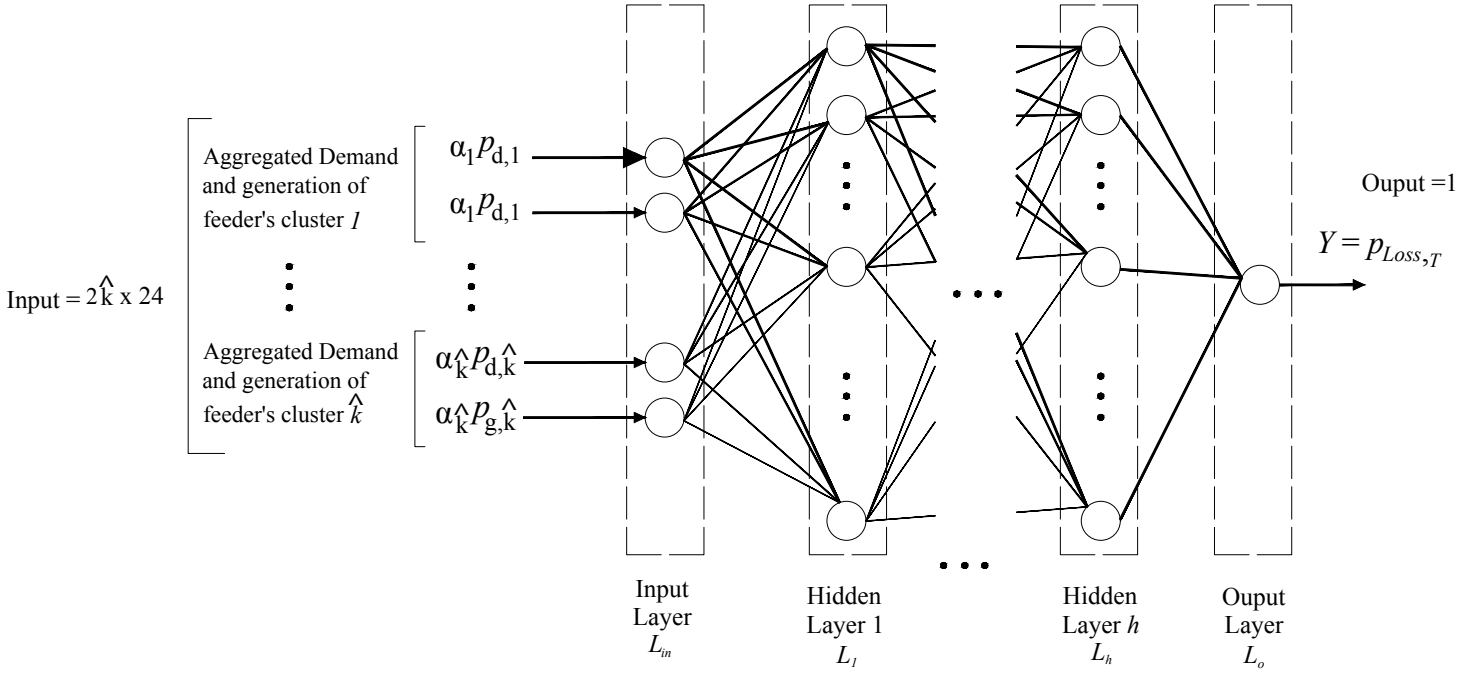
Extrapolating factor α_r: from the representative feeder to the rest of the feeders inside the cluster C_r:

$$\alpha_r = \frac{1}{n_r} \sum_{i=1}^{n_r} \|c_i - \tilde{c}_r\|^2, \ r \in \left(1, \ldots, \hat{k}\right)$$

5) Model Formulation

Model Overview

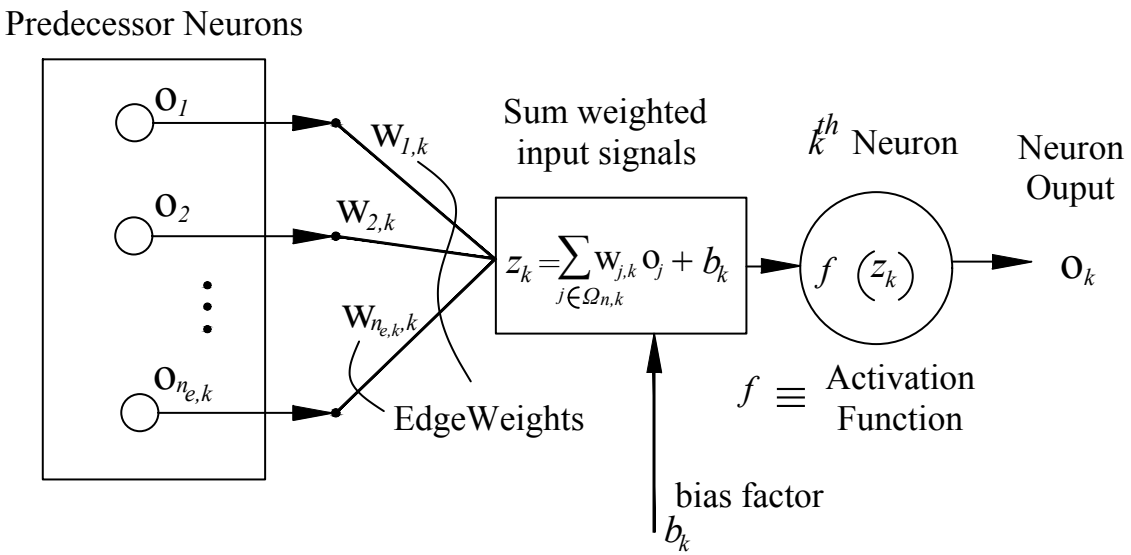
- Feed-Forward Deep Neural Network → Regression problem (Supervised Learning).
- Underlying complex structure → Flexible and adaptive.



Model Structure

$$\mathbb{W} = (\mathcal{N}, \mathcal{E})$$

Set of neurones: $\mathcal{N} = \{\mathcal{G}_{k,i} | \forall k \in (in, 1, \ldots, h, o), \forall i \in (1, \ldots, n_k)\}$
Set of directed edges $\mathcal{E} : \mathcal{E} \subseteq \mathcal{N} \times \mathcal{N} = \{e_{i,k} | \forall i, j \in \mathcal{N}\}$



Model Input/output

Mapping function:

$$\mathbb{Y} = \mathcal{F}_N(\mathbb{W}, \mathbb{X}) = p_{loss,T}$$

Representative Patterns (PR) of demand and generation:

$$\mathbb{X} = \begin{bmatrix} \mathbb{X}_{d,1} & \mathbb{X}_{g,1} \ldots \mathbb{X}_{d,\hat{k}} & \mathbb{X}_{g,\hat{k}} \end{bmatrix}^T$$

$$\mathbb{X}_{d,r} = \alpha_r \cdot p_{d,r,t}, \ \forall r \in (1, \ldots, \hat{k})$$

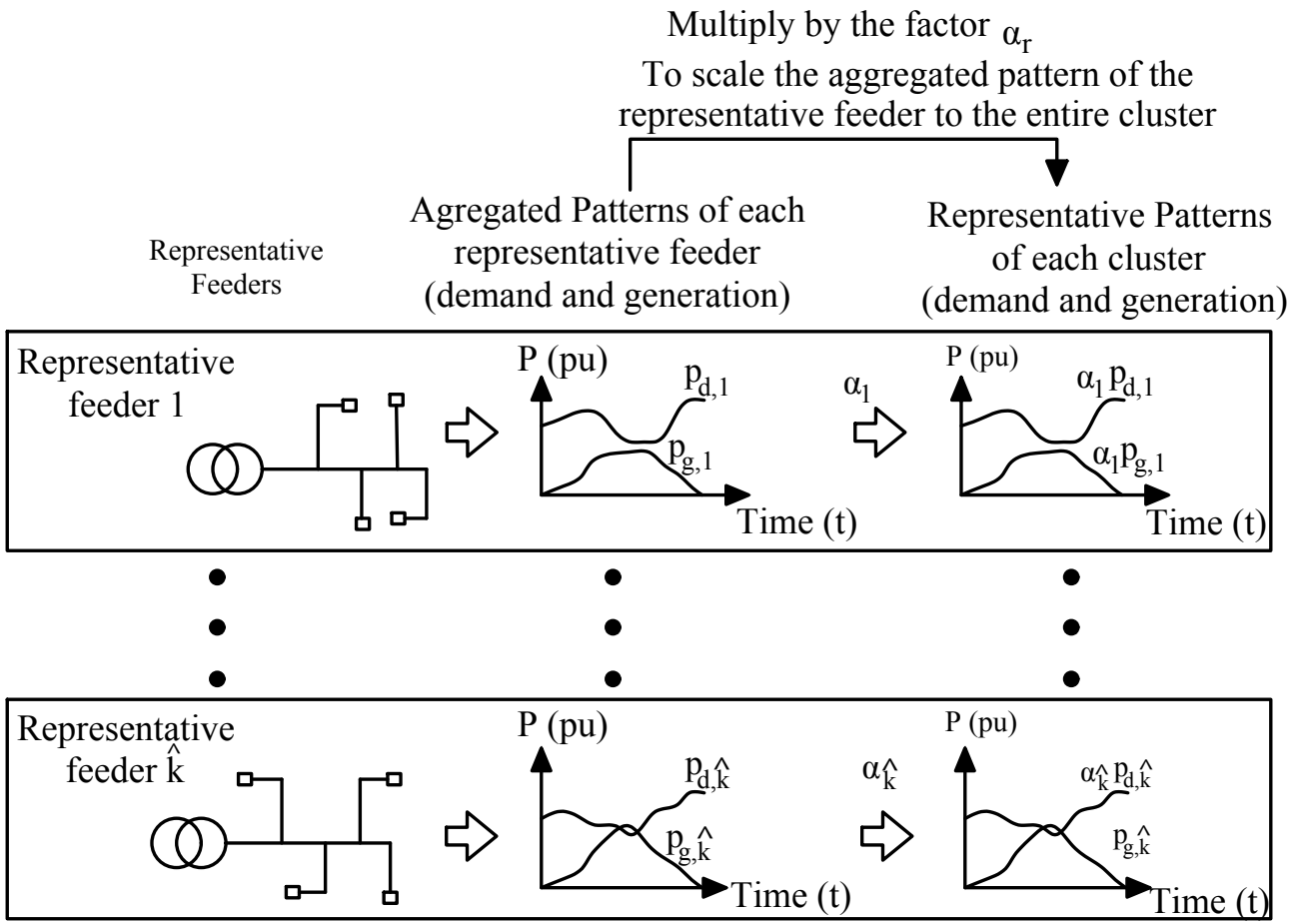
$$\mathbb{X}_{g,r} = \alpha_r \cdot p_{g,r,t}, \ \forall r \in (1, \ldots, \hat{k})$$

Aggregated Patterns (APs) for demand and generation:

$$p_{d,r,t} = \sum_{i \in \Omega_{c,r}} p^{(i)}_{d,r,t}$$

$$p_{g,r,t} = \sum_{i \in \Omega} p^{(i)}_{g,r,t}$$

Model Input/output



Training Data

Demand APs:

1. Demand AP for Winter & Working days (D1):

$$D_1 : p_{d,r}^{wi+wd(\omega)}(t) = \mu_{p_{d,r}^{wi+wd}(t)} \pm \sigma_{p_{d,r}^{wi+wd}(t)}$$

2. Demand AP for Winter & Non-Working days (D2):

$$D_2 : p_{d,r}^{wi+nwd(\omega)}(t) = \mu_{p_{d,r}^{wi+nwd}(t)} \pm \sigma_{p_{d,r}^{wi+nwd}(t)}$$

3. Demand AP for Summer & Working days (D3):

$$D_3 : p_{d,r}^{su+wd(\omega)}(t) = \mu_{p_{d,r}^{su+wd}(t)} \pm \sigma_{p_{d,r}^{su+wd}(t)}$$

4. Demand AP for Summer & Non-Working days (D4):

$$D_4 : p_{d,r}^{su+nwd(\omega)}(t) = \mu_{p_{d,r}^{su+nwd}(t)} \pm \sigma_{p_{d,r}^{su+nwd}(t)}$$

Training Data

Generation APs:

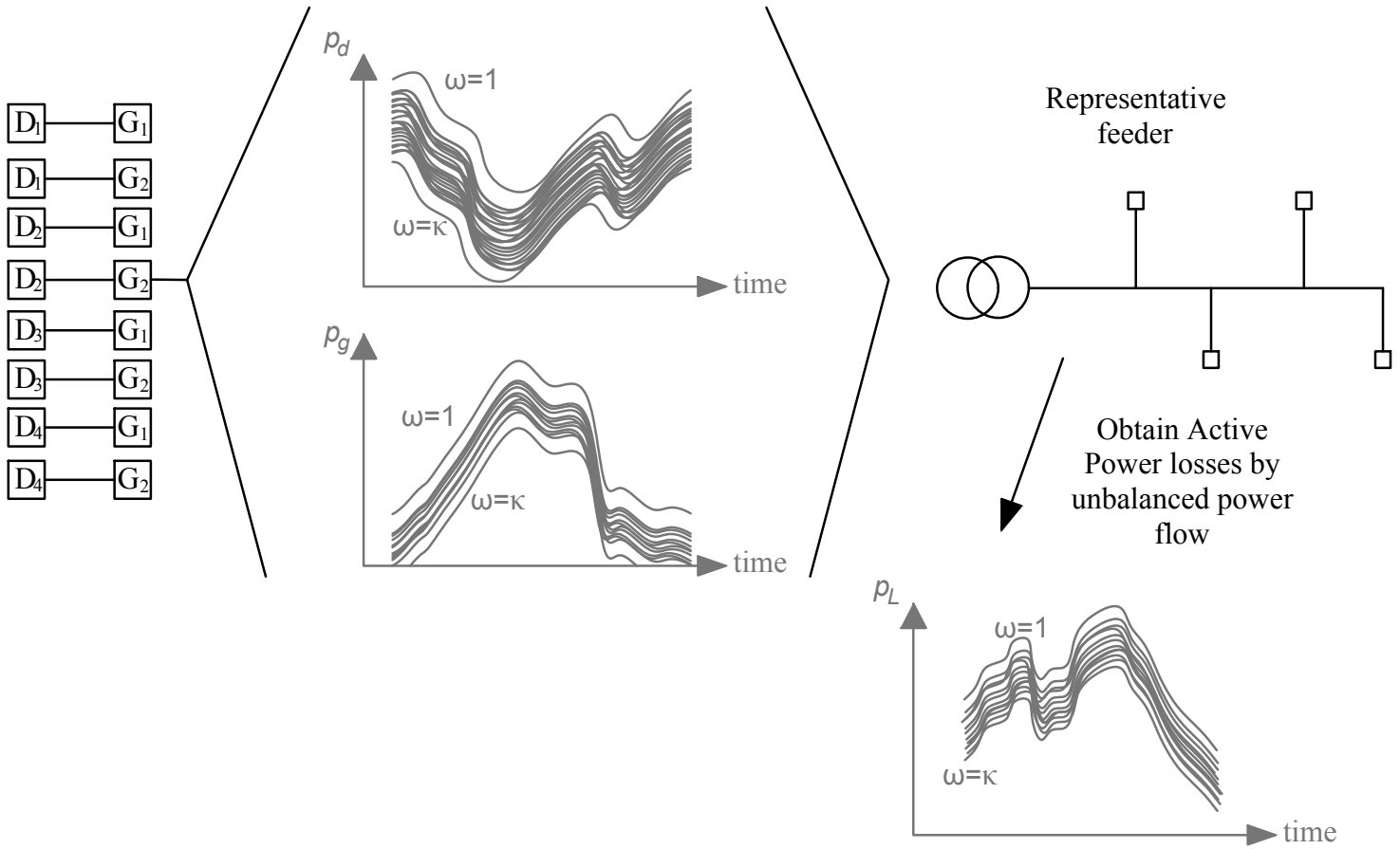
1. Generation AP for Winter (G1).

$$G_1 : p_{g,r}^{wi(\omega)}(t) = \mu_{p_{g,r}^{wi}(t)} \pm \sigma_{p_{g,r}^{wi}(t)}$$

2. Generation AP for Summer (G2).

$$G_2 : p_{g,r}^{su(\omega)}(t) = \mu_{p_{g,r}^{su}(t)} \pm \sigma_{p_{g,r}^{su}(t)}$$

Training Data



Model Training

Objective: Minimize **loss function** (MSE):

$$\mathfrak{L} = \frac{1}{\nu} \sum_{\pi=1}^{\nu} \left| \hat{p}^{(\pi)}_{loss,T} - p^{(\pi)}_{loss,T} \right|^2 = \mathfrak{L}(w_{i,j})$$

Subject to weight updating through **Back-Propagation**:

$$\Delta w_{i,j} = w_{i,j}(\hat{t}+1) - w_{i,j}(\hat{t}) = \eta \underbrace{\frac{\partial \mathfrak{L}}{\partial w_{i,j}}}_{Gradient}, \quad \forall w_{i,j} \in \mathbb{W}$$

ν batch size

η leaning rate

$\hat{t}$ time discrete parameter for iteration training step.

Model Training

0) **Weight matrix initialisation**:

$$\mathbb{W}(\hat{t}) \sim U(\underline{w_{i,j}^{min}(0)}, \overline{w_{i,j}^{max}(0)})$$

1) **Propagation of errors**:

$$\epsilon_k^{(\pi,\hat{t})} = W_k(\hat{t}) \cdot \underbrace{\epsilon_o^{(\pi,\hat{t})}}_{\mathfrak{L}}, \quad \forall k \in (1, \ldots, h)$$

2) **Gradient Descent**:

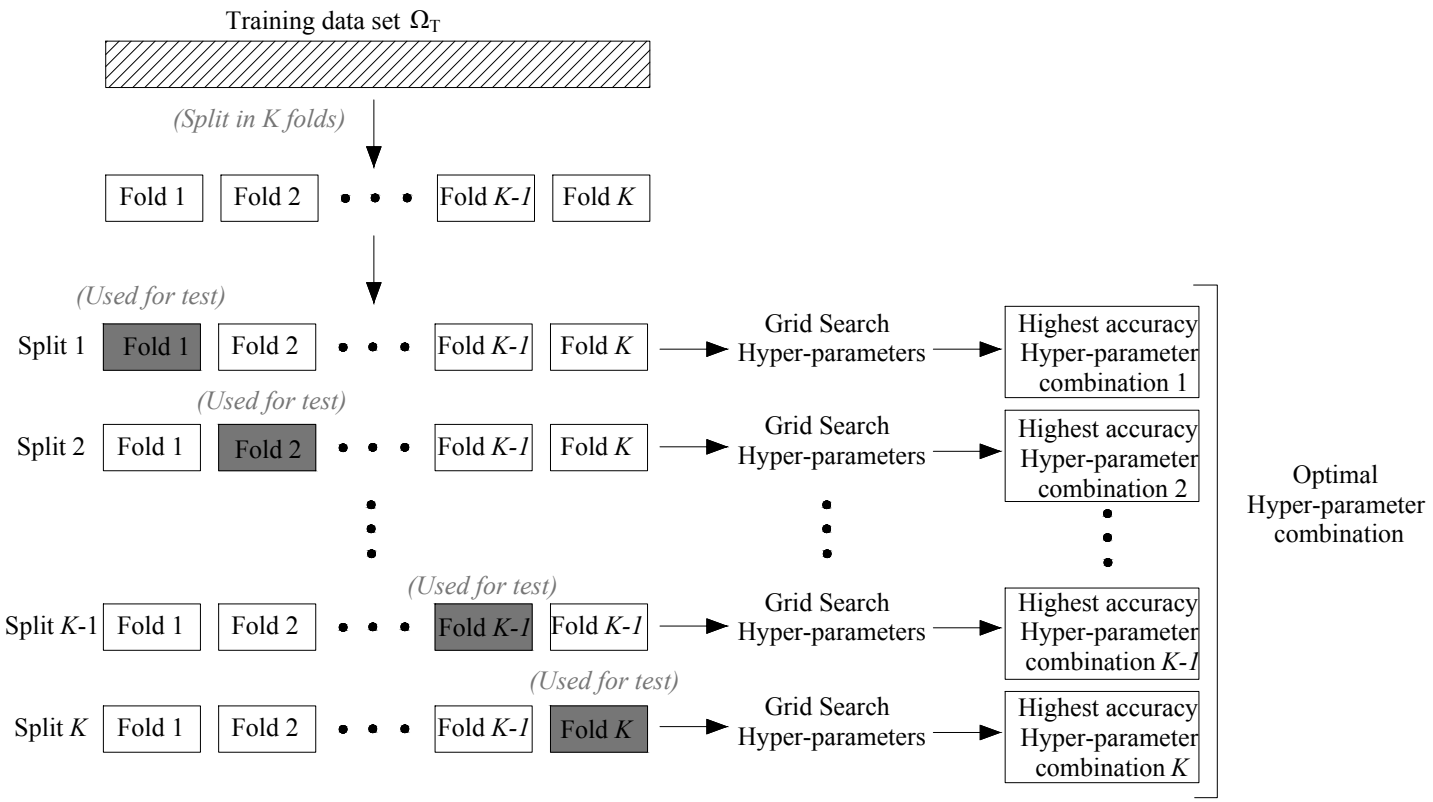
$$\frac{\partial \mathfrak{L}}{\partial W_k(\hat{t})} = -\epsilon_k^{(\pi,\hat{t})} \cdot O_k(I_{n_k} - O_k) \cdot O_{k-1}^T, \quad \forall k \in (1, \ldots, h)$$

3) **Weight updating**:

$$W_k(\hat{t}+1) = W_k(\hat{t}) - \eta \frac{\partial \mathfrak{L}}{\partial W_k(\hat{t})}, \quad \forall k \in (1, \ldots, h)$$

Model Validation

Validation strategy: $\mathcal{K}$-Fold validation and Grid-Search



Model Validation

Table 4.3: Candidate values of the hyper-parameters' model for $\mathcal{K}$-fold process.

Hyper-parameters	**Candidate values**				
No. Hidden Layers (h)	2	3	5	10	20
No. Hidden neurons (n_h)	4	6	8	12	16
Dropout (ζ)	0.05	0.10	0.20	0.25	0.5
Learning rate (η)	1e-6	1e-4	1e-3	1e-2	1e-1
Batch size (ν)	0.1ξ	0.25ξ	0.5ξ	1.0ξ	2ξ
Upper edge-weight $\overline{w_{i,j}(0)}$	0.6	0.7	0.8	0.9	1.0
Lower edge-weight $\underline{w_{i,j}(0)}$	0.1	0.2	0.3	0.4	0.5

Case Study

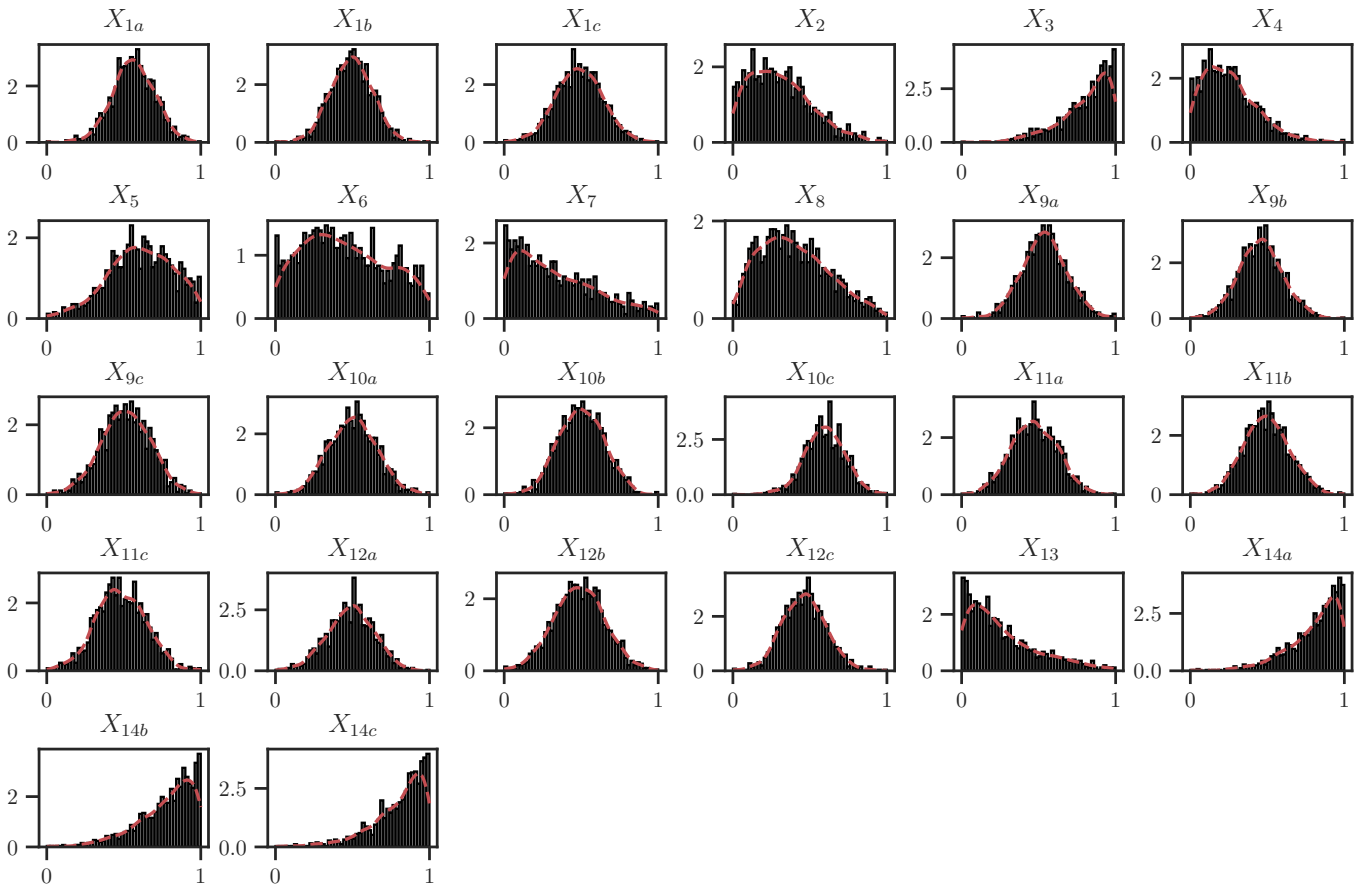
Case Study

Table 4.4: Case study data for large-scale power losses estimation

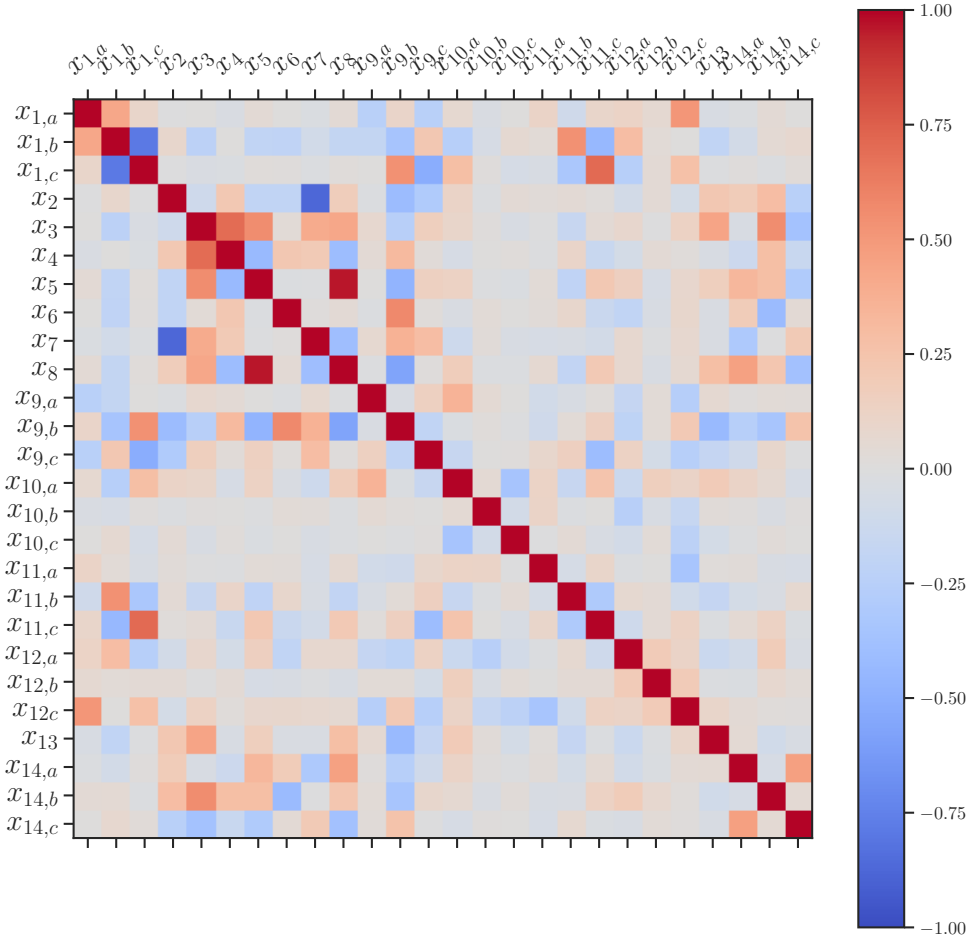
Case study properties	**Value**
Geographical Area Covered (ha.)	605
No. Secondary Substations (No. SS)	147
No. Customers (three phase, single-phase) †	30,429
No. Feeders (n)	1,256
Power Contracted (MW)	546
Accumulated feeder Length (km)	273
Max. (PV-based) DG-presence level (%)	55
Avg. Smart Meter penetration level (%)	88
Avg. Power Phase Unbalance (%)	13.8
Aver. Ratio Customers/SS	207
Aver. Ratio Customers/Feeder	24
Network Type	100% Urban
Cables Material	100% Aluminium
Cables impedance	(0.1 , 3) Ω/km
Network Configuration	80% Underground
	20% Overhead

† residential, commercial and industrial

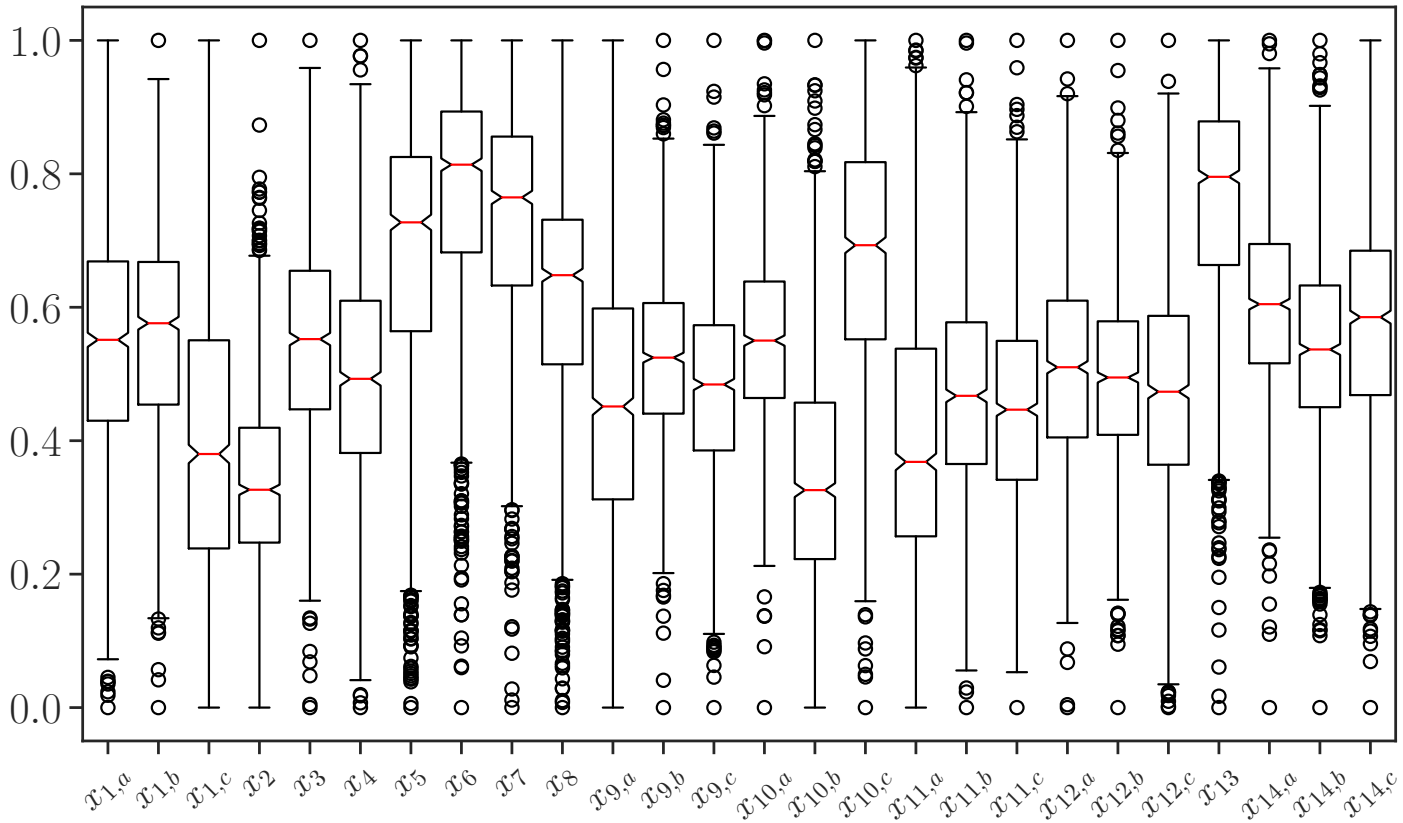
PDFs and histograms of the feeders' characteristics



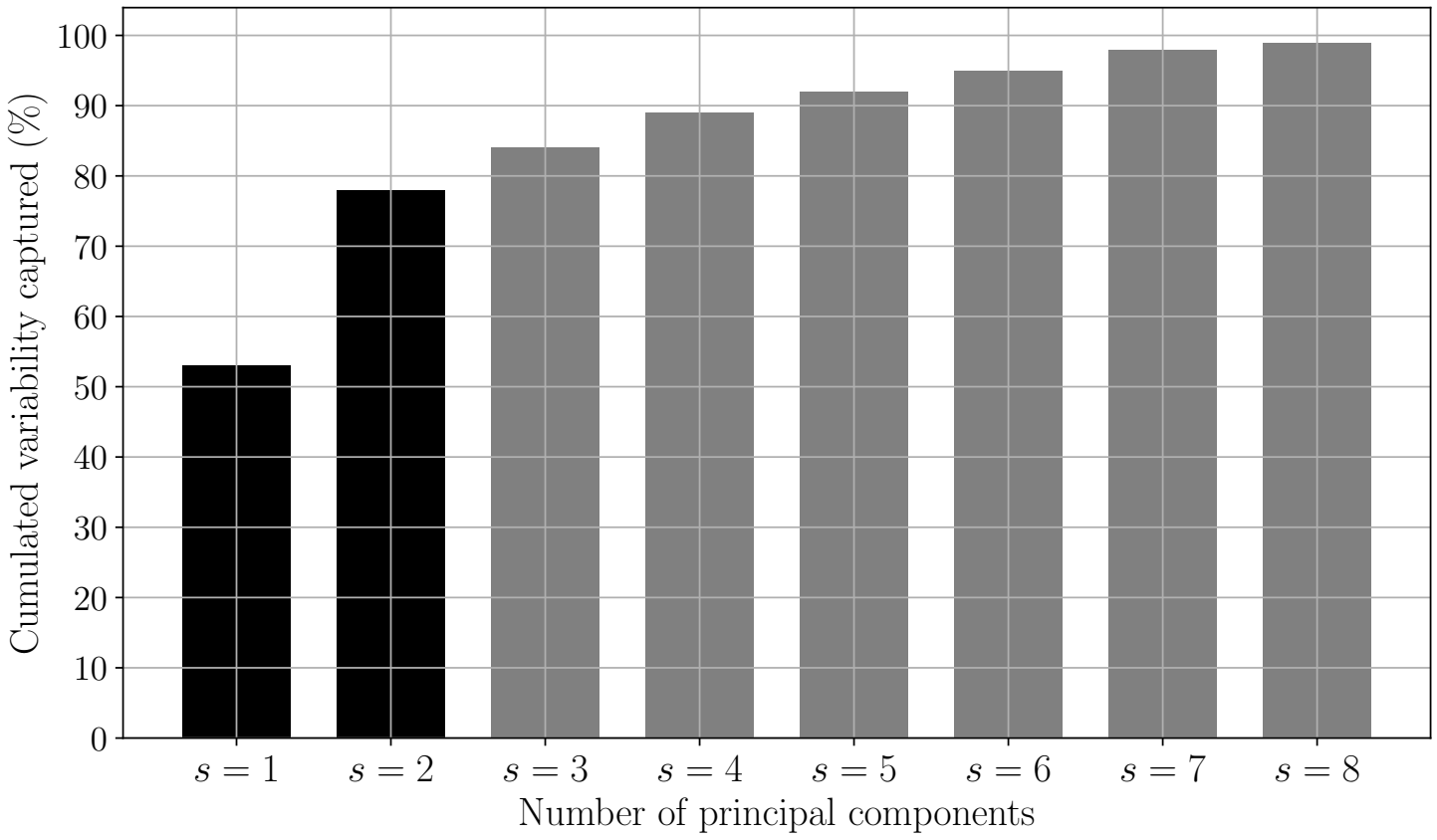
Correlation matrix of the feeder's characteristics



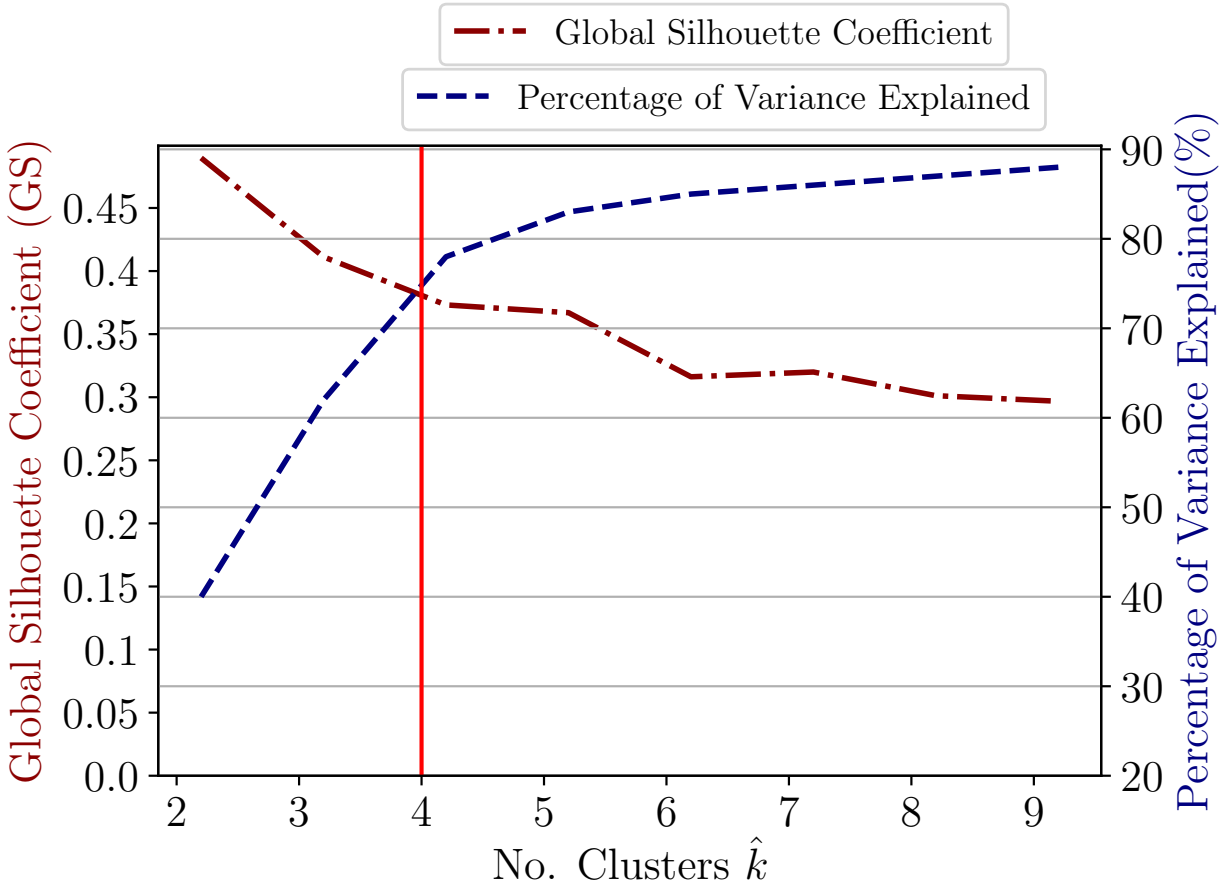
Normalize feeder's characteristics



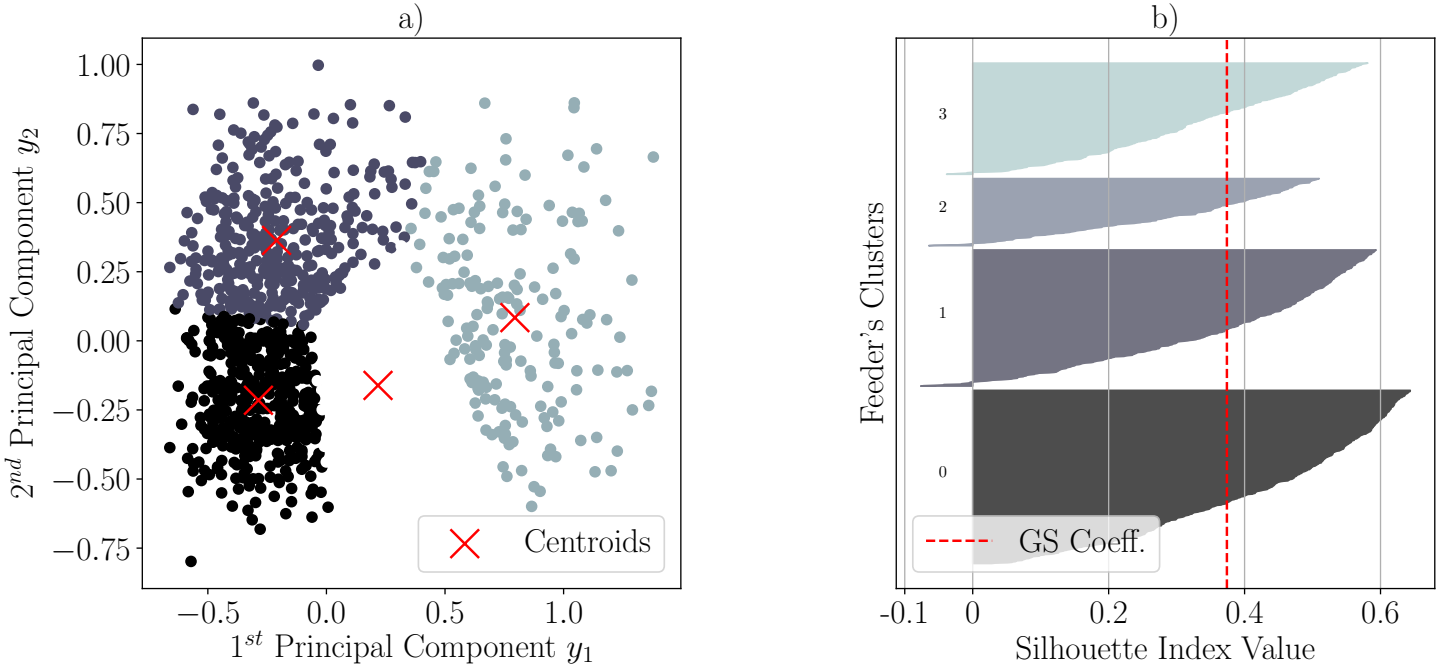
Features Selection



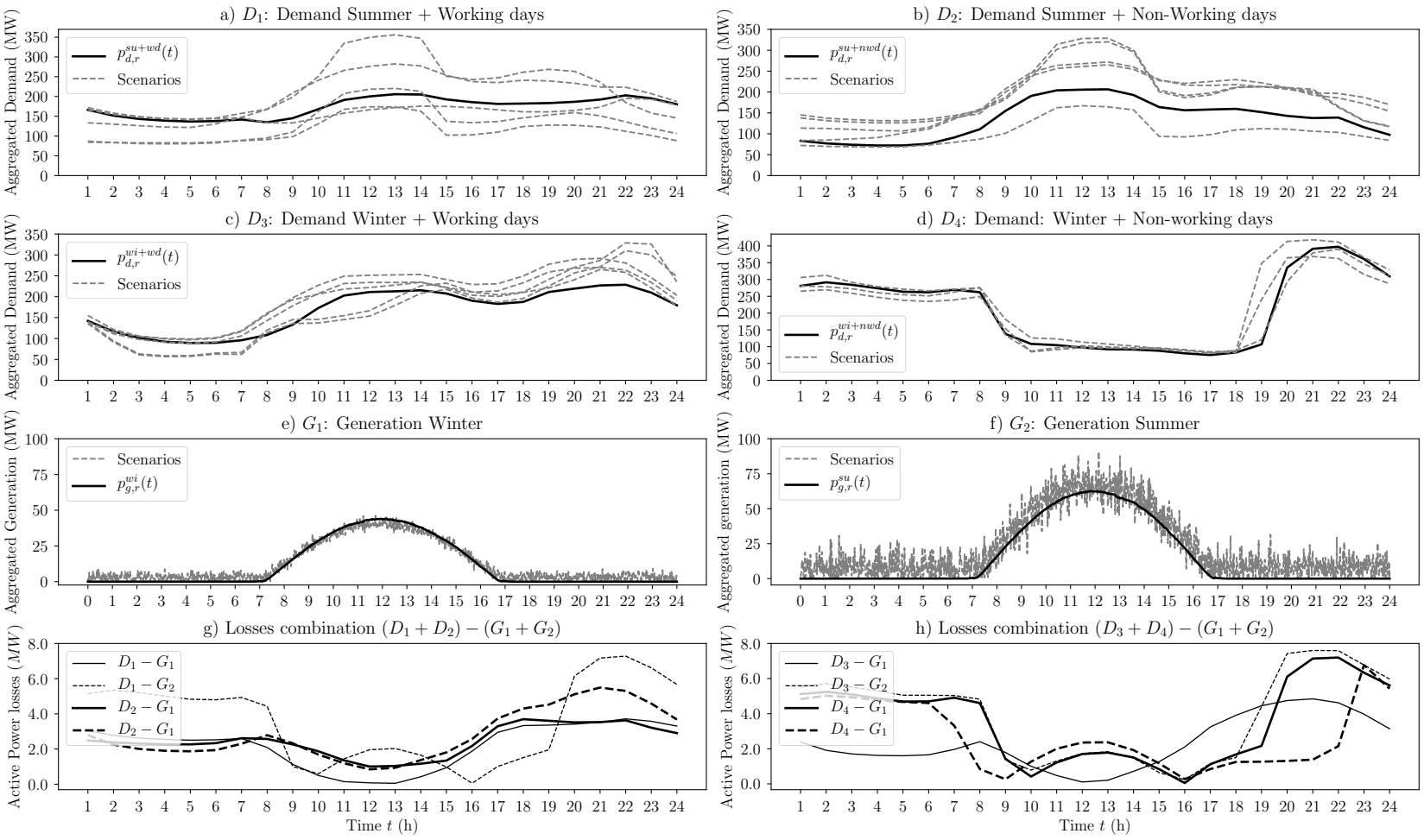
Feeder's Clustering



Feeder's Clustering



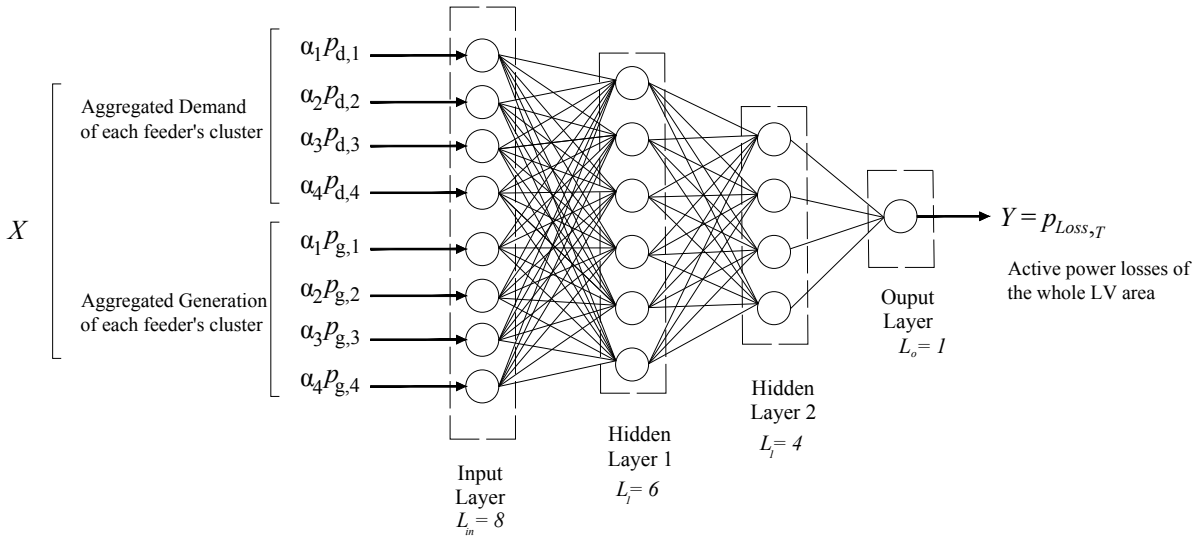
Training data



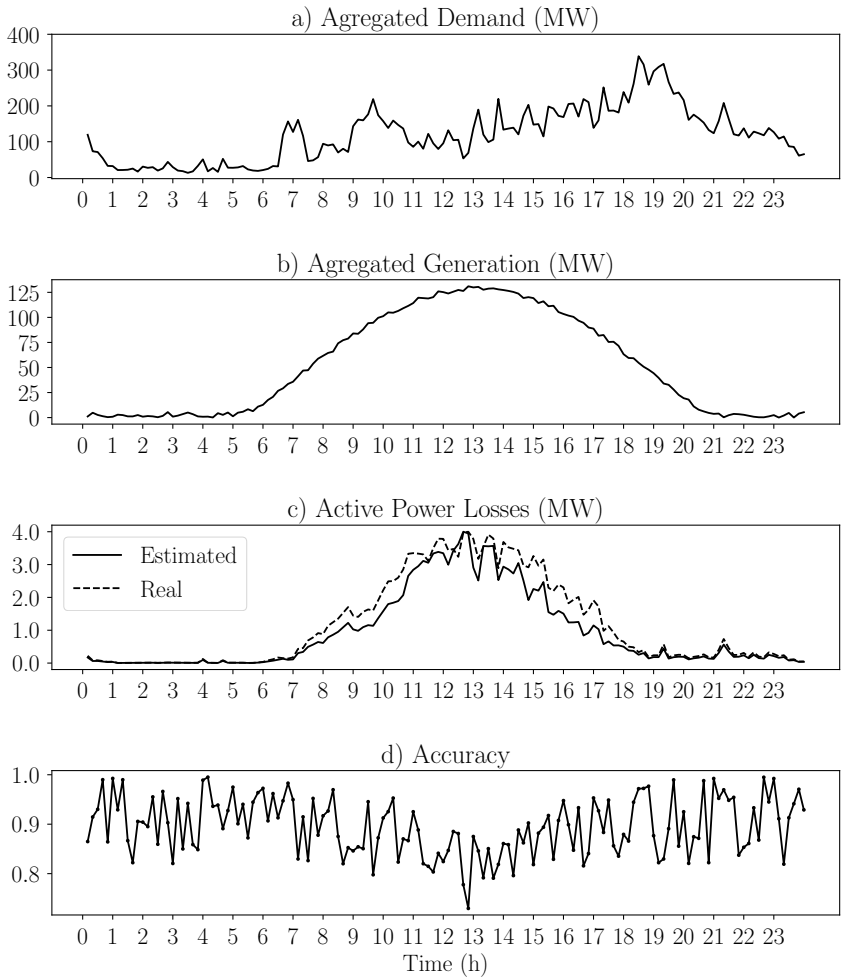
Model validation: K-fold and Grid-Search

Table 4.7: Hyper-parameter combinations with the highest accuracy for each split

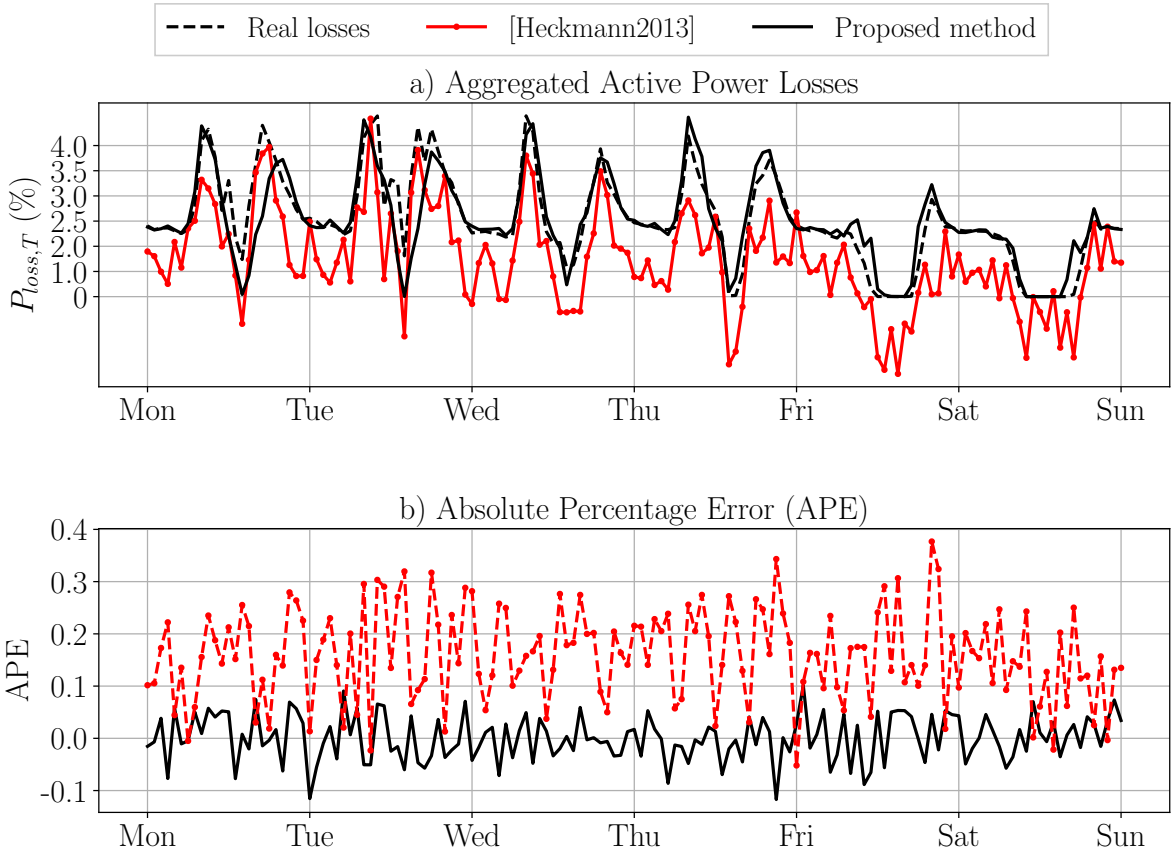
Accuracy Λ	$\mathcal{K}$ Fold	Architecture $(n_1 : \ldots : n_h)$	Dropout ζ	Learning rate η	Weight initialisation $(w_{i,j}^{min}, w_{i,j}^{max})$	Batch size ν
0.9365	1	(8:6:4:1)	0.24	1e-2	(0.06,0.75)	240
0.9271	2	(8:6:4:1)	0.32	1e-3	(0.27,0.52)	240
0.9205	3	(8:4:8:1)	0.41	1e-2	(0.35,0.60)	240
0.9200	4	(8:12:12:1)	0.38	1e-4	(0.02,0.72)	48



Power losses estimation: Time horizon : 1 day



Power losses estimation: Time Horizon: 1 week

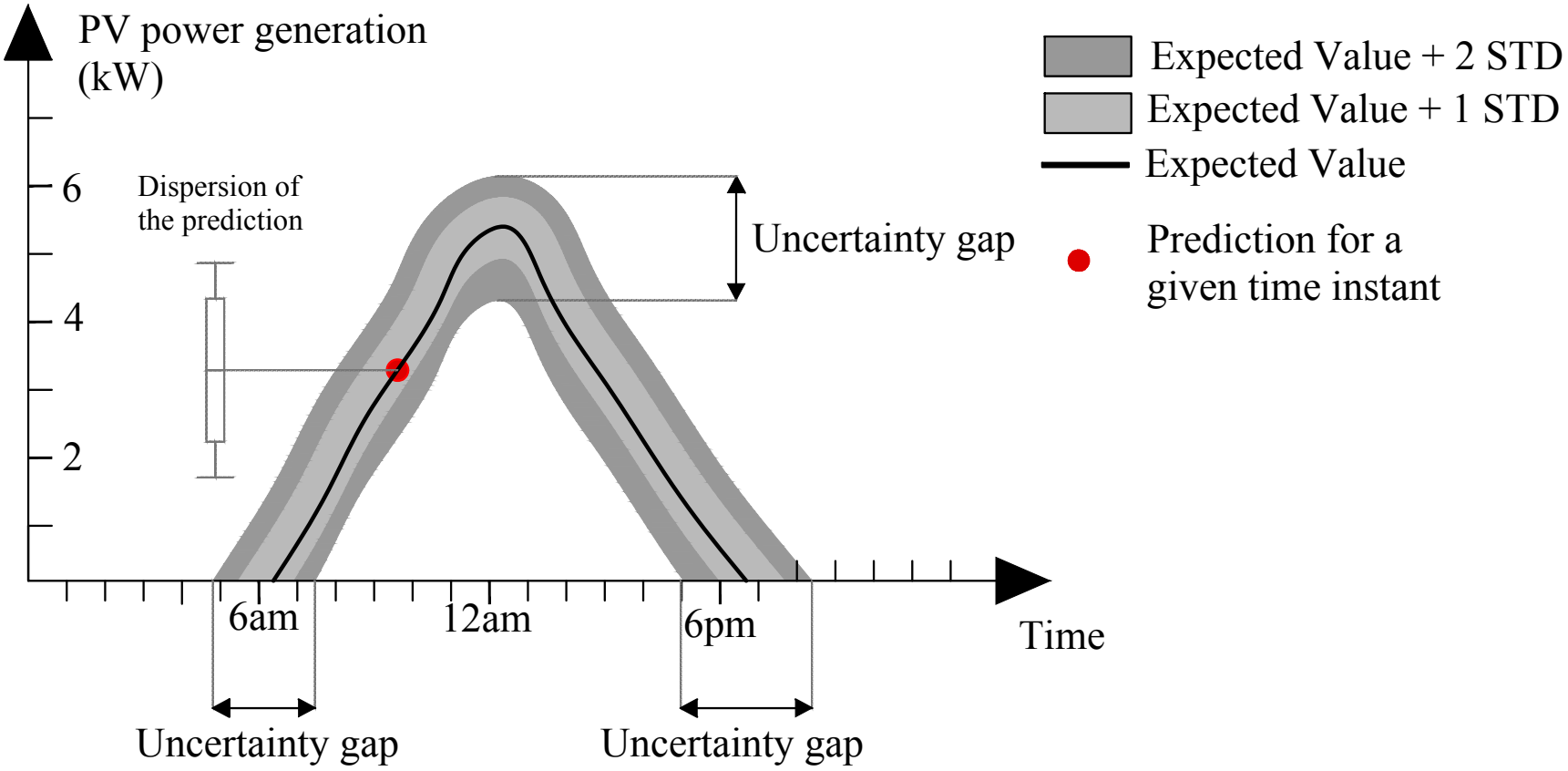


5. Power losses minimization in smart grids under uncertainty

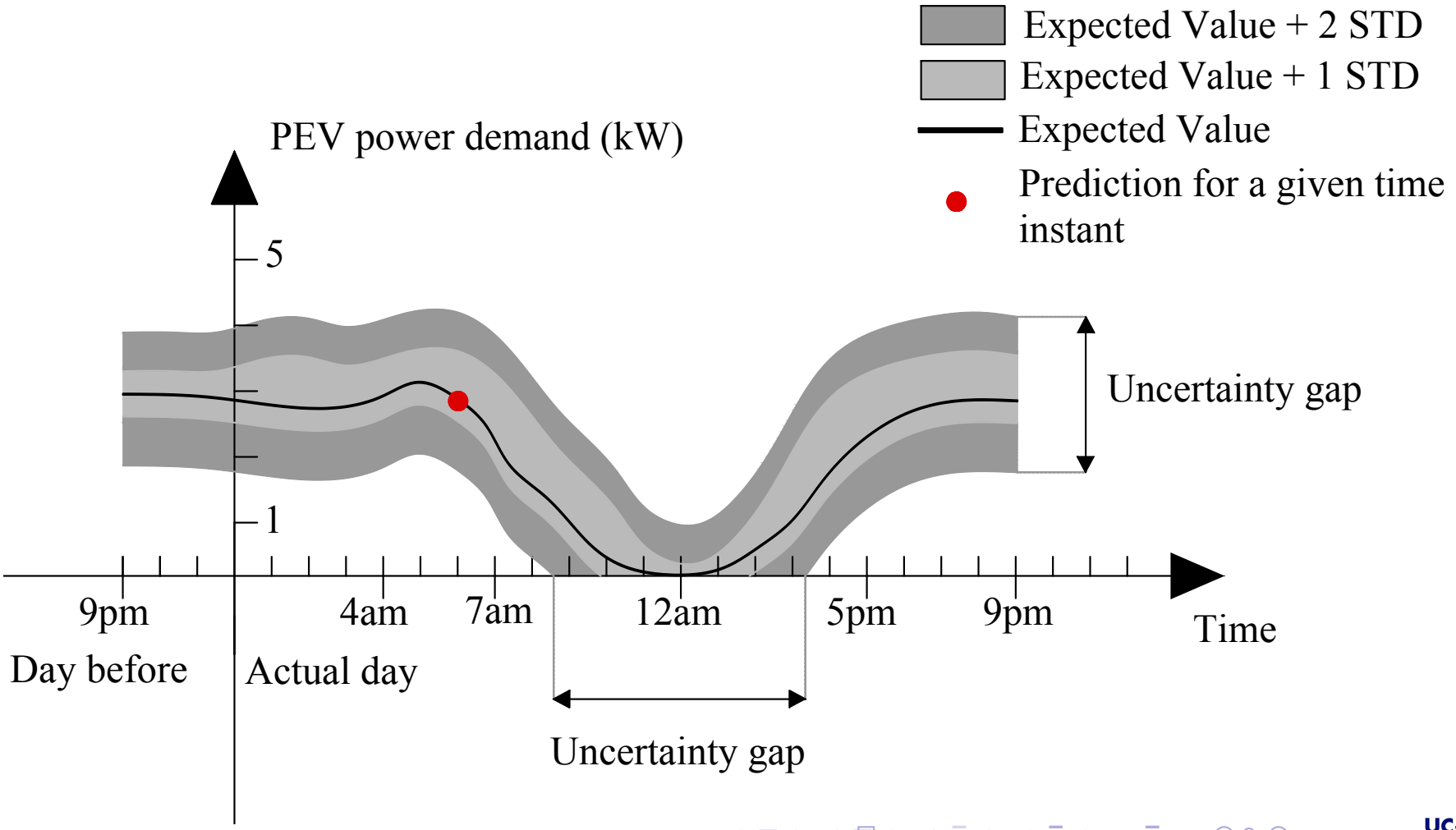
Introduction

- **Goal**: Determine the flexibility actions required to keep the system within normal operational conditions and minimizing the network power losses.
- Required three-phase model due to the unbalanced operation of LV smart grids → **Unbalanced Optimal Power Flow** (uOPF).
- Explicit **Uncertainty** treatment:
  - Intermittent behaviour of renewable-based resources.
  - E.g. Photovoltaic power injections or Unexpected PEV charging sessions.
- Flexibility mechanism: **Prosumer**

PV Uncertainty



PEV Uncertainty



Robust uOPF: Objective Function and variables (1/2)

$$\min_{\mathbf{x},\mathbf{y}^+,\mathbf{y}^-} \max_{\mathbf{u},\mathbf{z}} \quad \mathcal{OF}: \overbrace{\sum_{t\in T}\sum_{k\in V}\sum_{p\in P} x_{k,t}^{p}\cdot\xi_t + y_{k,t}^{p,+}\cdot\lambda_t + y_{k,t}^{p,-}\cdot\lambda_t}^{Flexibility\ \ Cost}$$

where:

$$\overbrace{\mathbf{x} = \left[x_{k,t}^{p}\right]^T}^{Flexibility\ \ Commitment} \quad \in (0,1),\ \ \forall k \in V^* \subseteq V$$

$$\overbrace{\mathbf{y}^+ = \left[y_{k,t}^{p,+}\right]^T}^{Flexibility\ \ Dispatch} \quad \in (y_k^{min}, y_k^{max}),\ \ \forall k \in V^* \subseteq V$$

$$\mathbf{y}^- = \left[y_{k,t}^{p,-}\right]^T \in (y_k^{min}, y_k^{max}),\ \ \forall k \in V^* \subseteq V$$

Robust uOPF: Objective Function and variables (2/2)

$$\mathbf{u} = \left[u^{p}_{k,t}\right]^{T} \in \Phi_{u,t}$$

$$\mathbf{z} = \left[z^{p}_{k,t}\right]^{T} \in \Phi_{z,t}$$

$$\Phi_{u,t}(\mu_{\mathbf{u},\mathbf{t}}, \sigma_{\mathbf{u},\mathbf{t}}) = \left\{ u^{p}_{k,t} \in \mathbb{R} : \sum_{\forall k \in N_u} \frac{\left|u^{p}_{k,t} - \mu_{u,t,k}\right|}{\sigma_{u,t,k}} \leq \varphi_{u,t} \right\}$$

$$\Phi_{z,t}(\mu_{\mathbf{z},\mathbf{t}}, \sigma_{\mathbf{z},\mathbf{t}}) = \left\{ z^{p}_{k,t} \in \mathbb{R} : \sum_{\forall k \in N_z} \frac{\left|z^{p}_{k,t} - \mu_{z,t,k}\right|}{\sigma_{z,t,k}} \leq \varphi_{z,t} \right\}$$

$$\vartheta = \left[v^{p,re}_{k,t} \quad v^{p,im}_{k,t}\right]^{T} \in \mathbb{R}, \forall k \in V, \forall t \in T, \forall p \in \Omega_p$$

Robust uOPF: Constraints (1/2)

s.t.

$$\mathbf{y}^{+} + \mathbf{y}^{-} = 0$$

$$x^{p}_{k,t} \cdot y^{min}_{k} \leq y^{p}_{k,t} \leq x^{p}_{k,t} \cdot y^{max}_{k}$$

$$p^{p,sp}_{i,k,t}(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{z}) + p^{p,cal}_{i,k,t}(\vartheta) = 0$$

$$p^{p,sp}_{i,k,t}(\mathbf{x},\mathbf{y},\mathbf{u},\mathbf{z}) = \overbrace{u^{p}_{k,t}}^{Generation} - \overbrace{(p^{p,sp}_{d,k,t} + y^{p}_{k,t} + z^{p}_{k,t})}^{Demand}$$

$$p^{p,cal}_{i,k,t}(\vartheta) = u^{p,re}_{k} \cdot i^{p,re}_{i,k} + u^{p,im}_{k} \cdot i^{p,im}_{i,k}$$

$$i^{p,re}_{i,k,t} = \sum_{k' \in \Omega_k} \sum_{q \in \Omega_p} g^{pq}_{kk'} \; v^{q,re}_{k',t} - b^{pq}_{kk'} \; v^{q,im}_{k',t}$$

$$i^{p,im}_{i,k} = \sum_{k' \in \Omega_k} \sum_{q \in \Omega_p} b^{pq}_{kk'} \; u^{q,im}_{k'} + g^{pq}_{kk'} \; u^{q,im}_{k'}$$

Robust uOPF (4/4): : Constraints (2/2)

$$\overbrace{\sum_{p\in\Omega_p}\sum_{k,j\in V,\ i\neq j}\left[\left(i^{p,re}_{kj,t}\right)^2+\left(i^{p,im}_{kj,t}\right)^2\right]\cdot r^p_{k,j}\cdot e^p_{k,j}}^{\mathcal{L}_p}\leq L^M_t$$

$$i^{p,re}_{kj,t} = g^p_{k,j}\left(v^{p,re}_k - v^{p,re}_j\right) + b^p_{k,j}\left(v^{p,im}_j - v^{p,im}_k\right)$$

$$i^{p,im}_{kj,t} = g^p_{k,j}\left(v^{p,im}_k - v^{p,im}_j\right) + b^p_{k,j}\left(v^{p,re}_k - v^{p,im}_j\right)$$

$$\left(i^{p,re}_{kj,t}\right)^2 + \left(i^{p,im}_{kj,t}\right)^2 \leq \left(i^{max}_{kj}\right)^2$$

$$\left(v^{min}\right)^2 \leq \left(v^{p,re}_{k,t}\right)^2 + \left(v^{p,im}_{k,t}\right)^2 \leq \left(v^{max}\right)^2$$

Method Comparison

- **Deterministic** version: → Forecast with SARIMA model for uncertain variables.
- **Stochastic** version:

$$\min_{\mathbf{x}_\omega, \mathbf{y}_\omega} \quad \sum_{\omega \in \Omega_s} \pi_\omega \cdot \mathcal{OF}_\omega$$

$\omega = 5$ Scenarios.

Case Study

Case Study

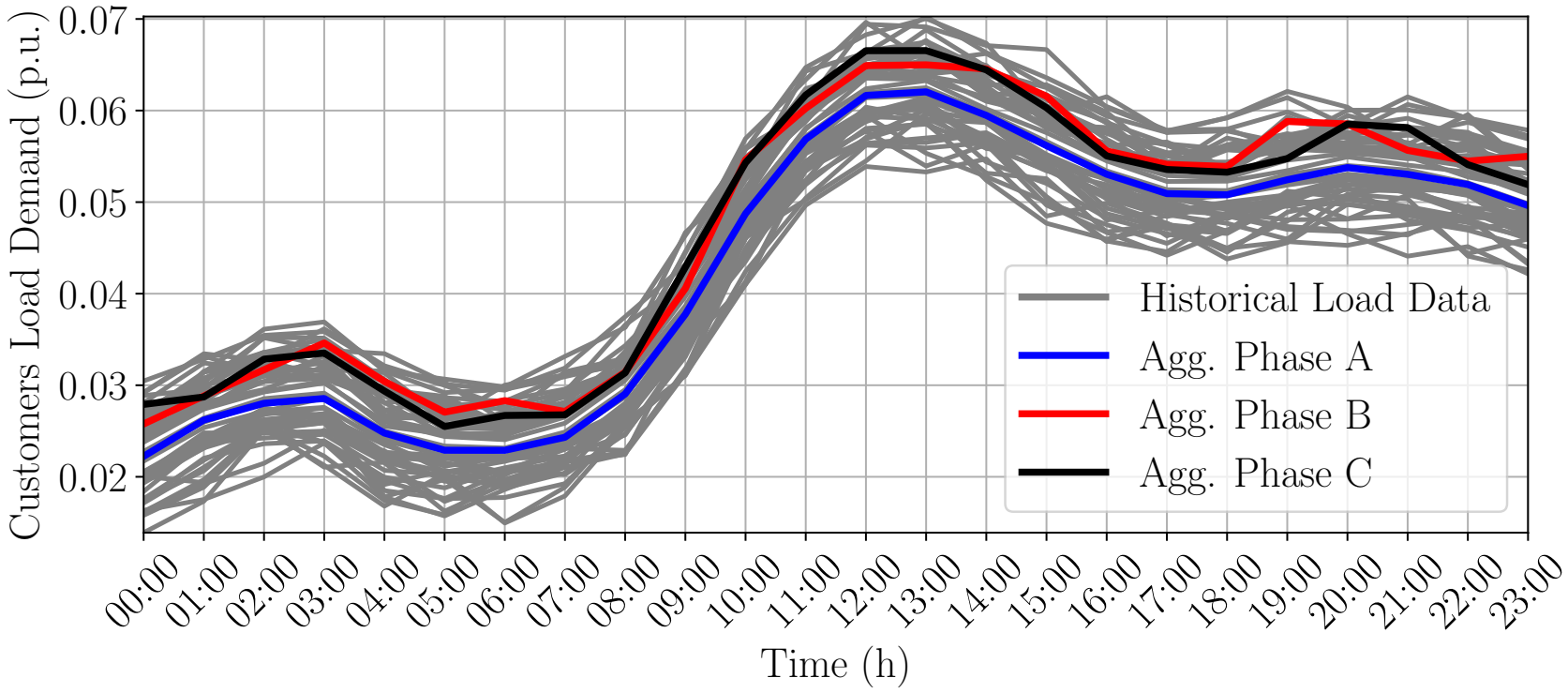
- IEEE European LV Test feeder.
- 1 SS (630 kVA).
- LV customers: 53 *(residential)*.
- Feeder length: 1.4 km.
- Contractual power: (3.4, 9.2) kW.
- PV panels arrays: (1, 4) kW.
- PEV charging: (2.3, 3.7) kW.
- Initial situation: expected values in demand and generation.
  $\Rightarrow$ Contingencies.

Case Study

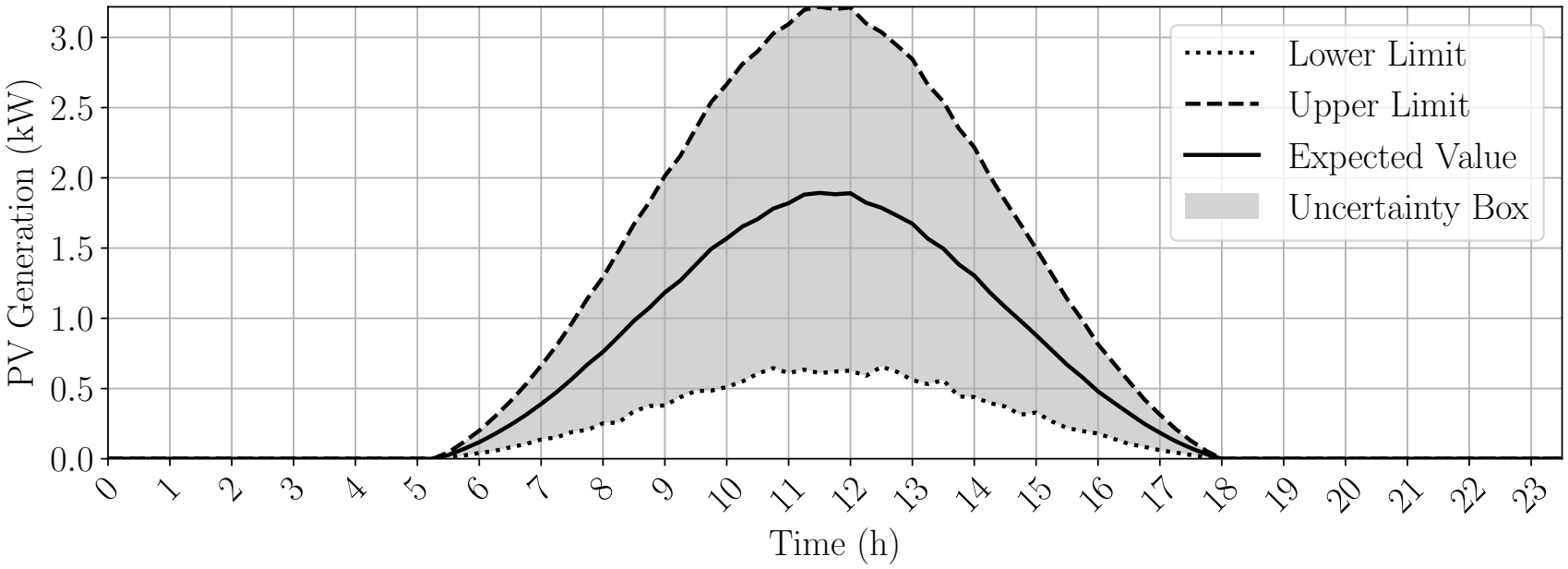
Table 5.1: Principal unbalanced network metrics without flexibility

Metric	**Value**
Load Unbalance ratio phase A	29 %
Load Unbalance ratio phase B	35 %
Load Unbalance ratio phase C	36 %
Unbalance Degree CIGRE-2 DQV	3.45 %
Max. phase voltage mag. (p.u.)	1.14
Max. phase current mag. (p.u.)	0.16

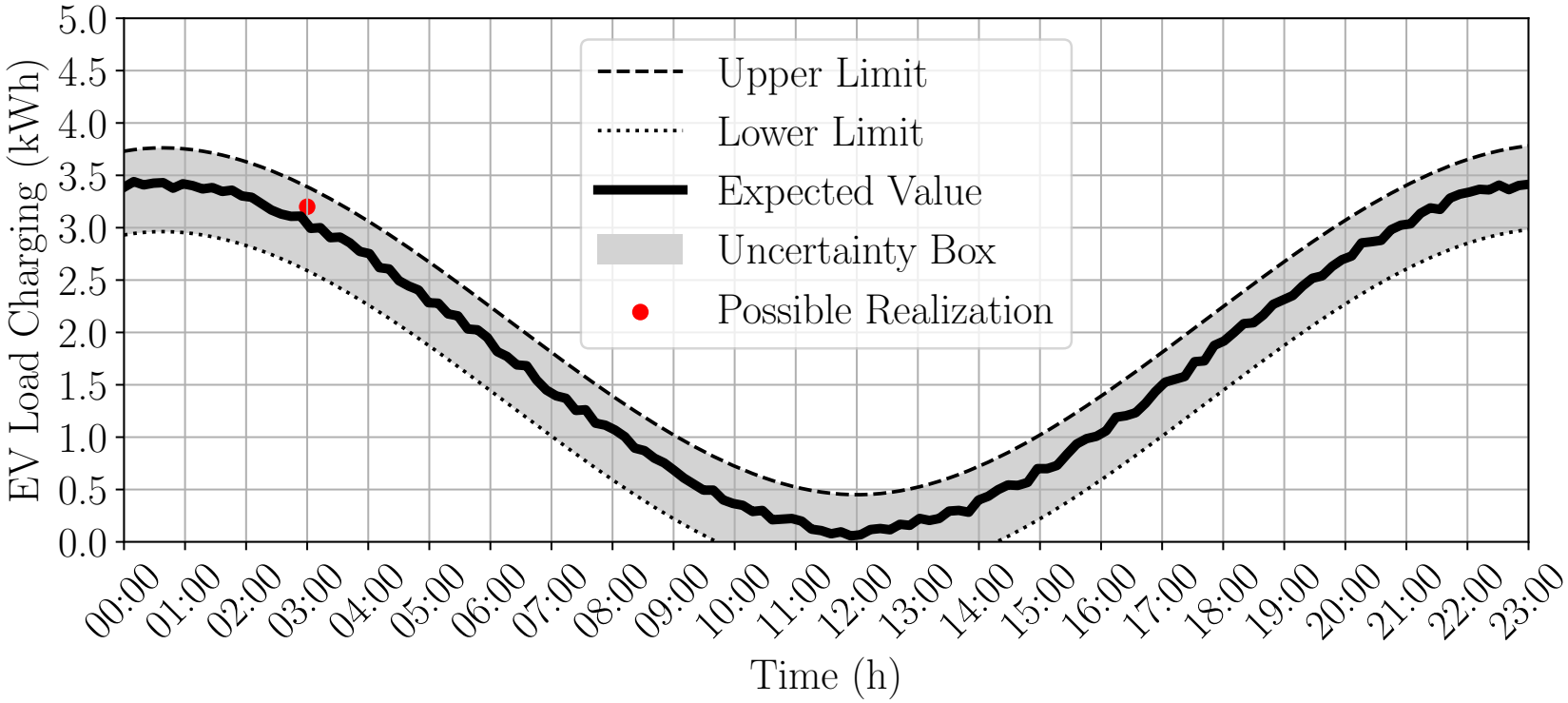
Initial Situation: Aggregated load demand per phase



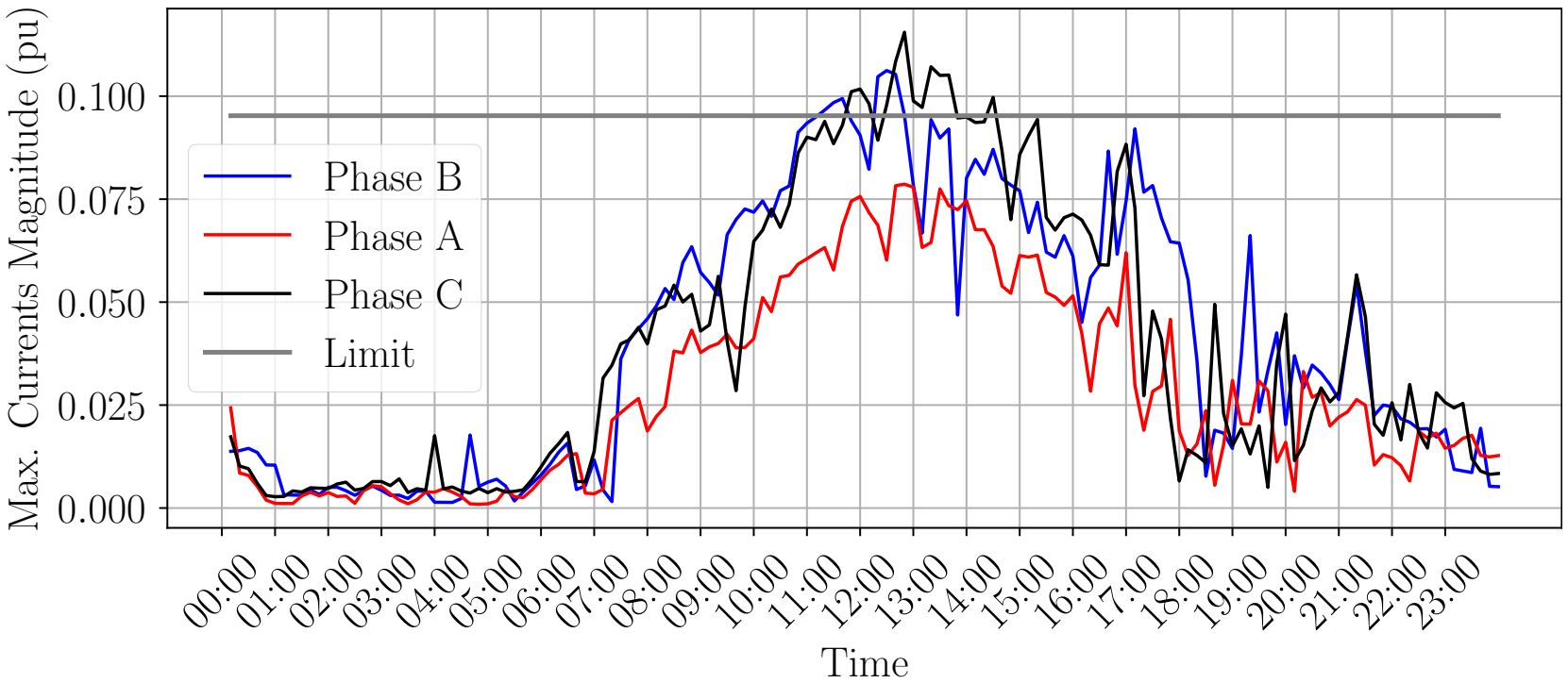
Initial Situation: PV Generation (node 27, phase C)



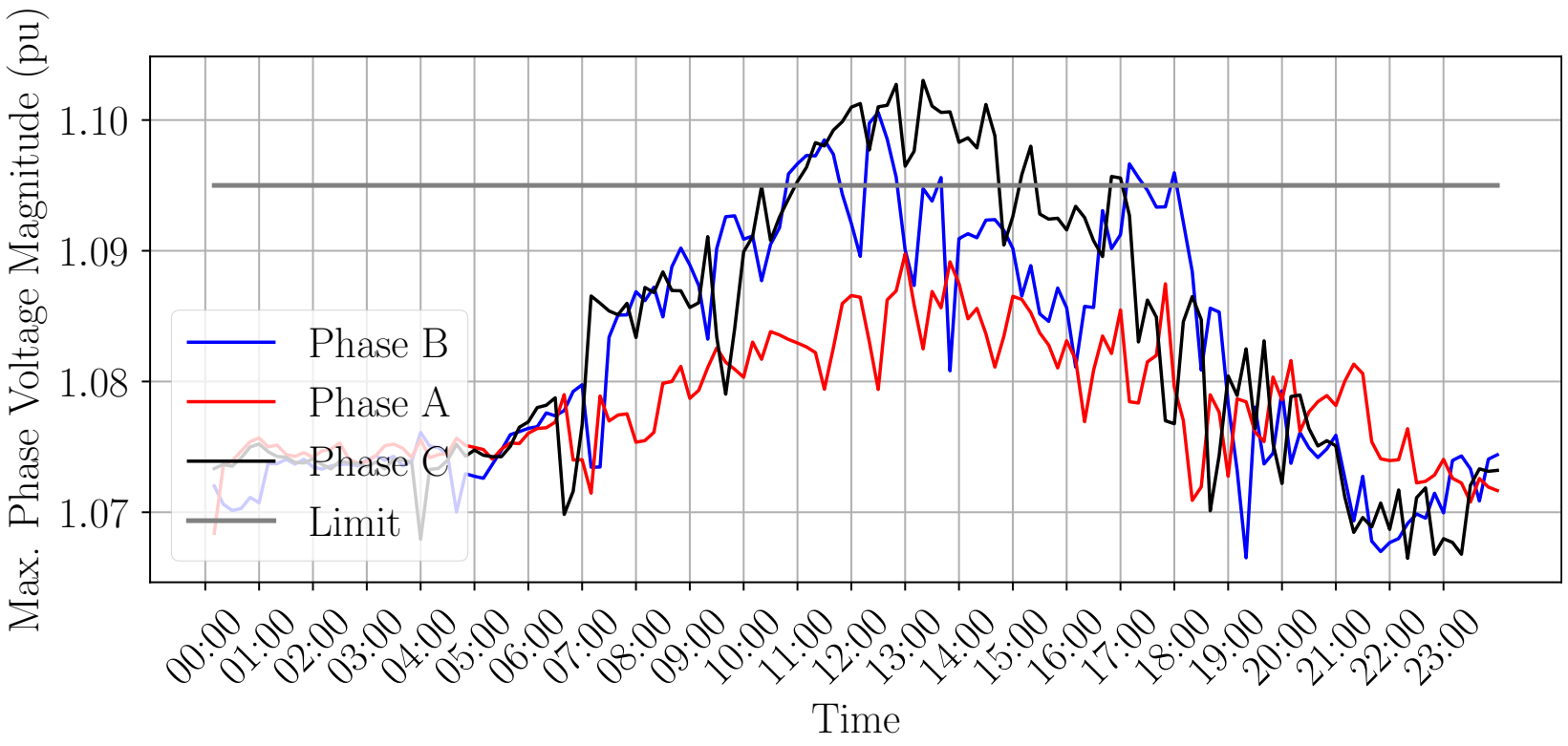
Initial Situation: PEV Load demand (node 27, phase C)



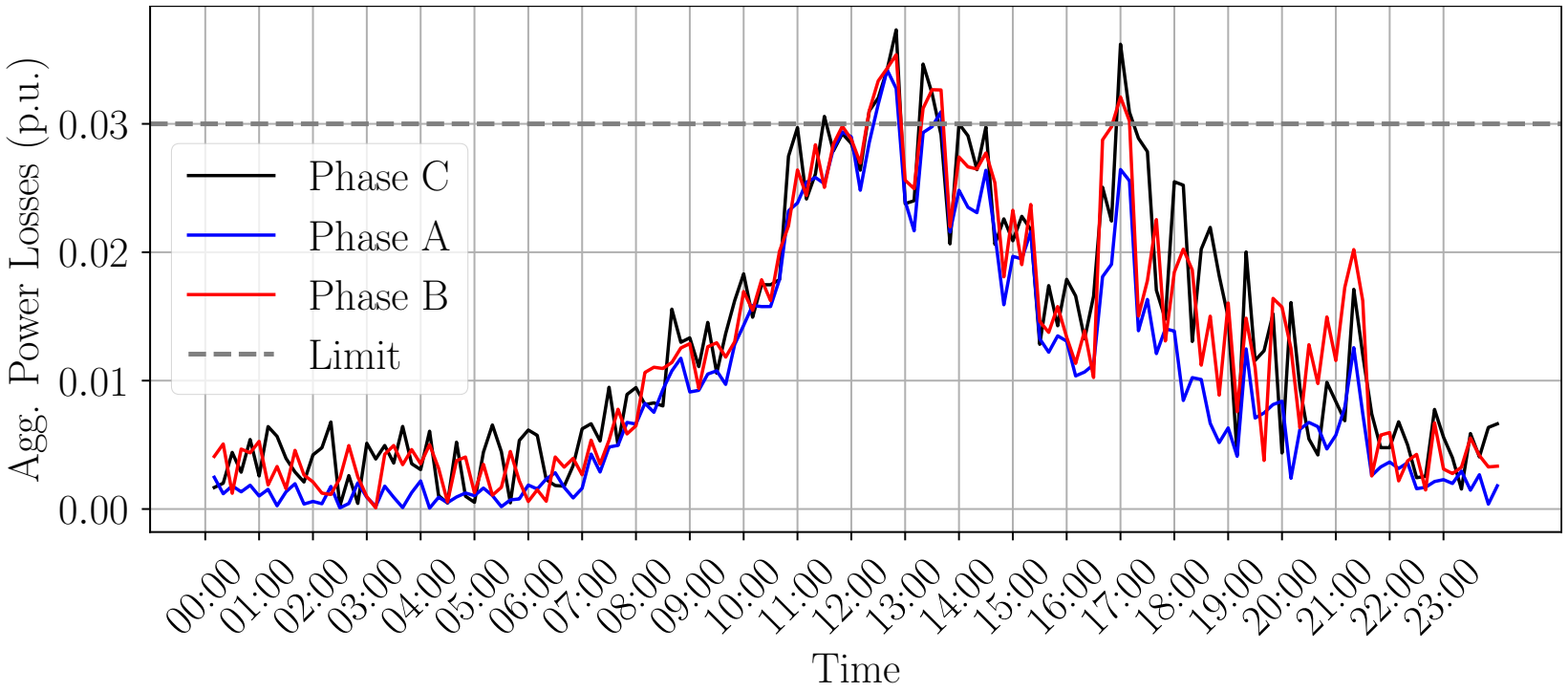
Initial situation: Currents (node 34)



Initial situation: Voltage (node 34)

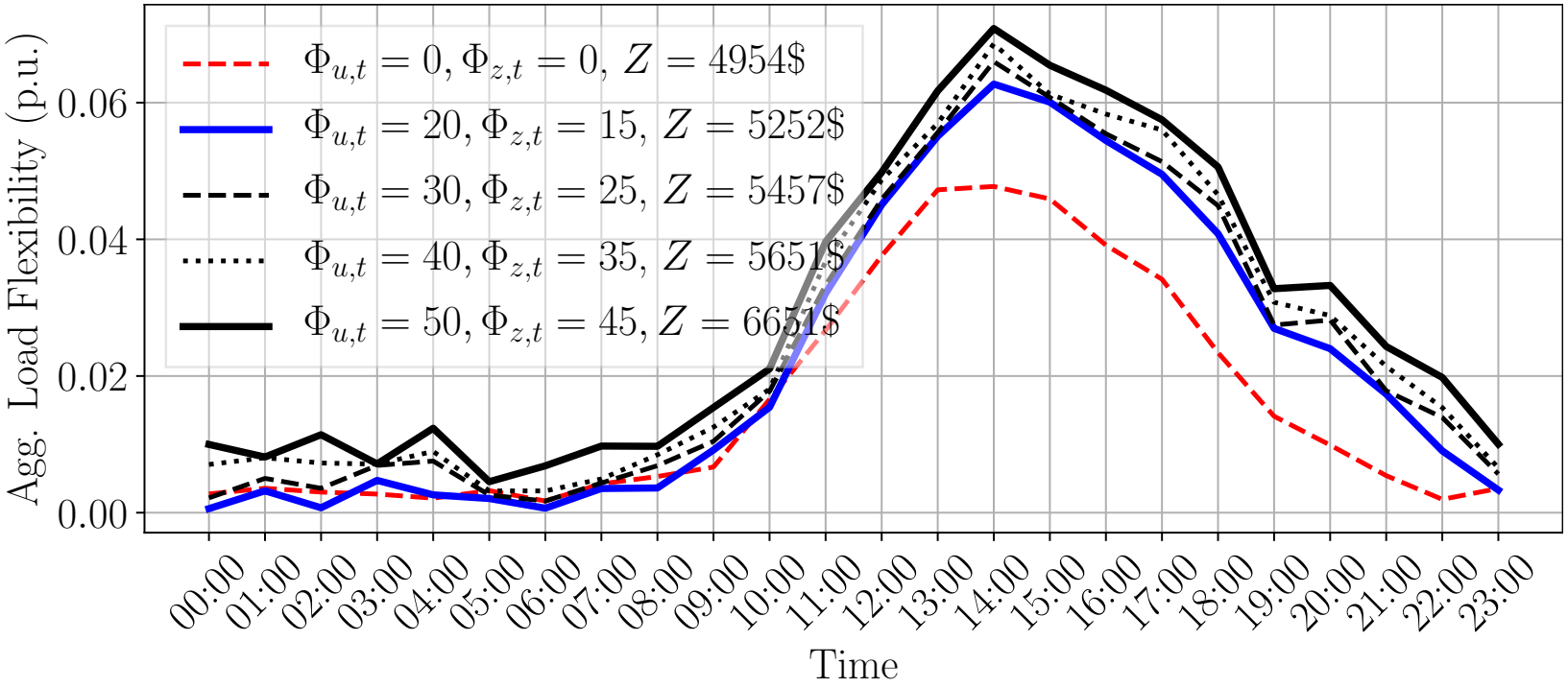


Initial situation: Losses (node 34)

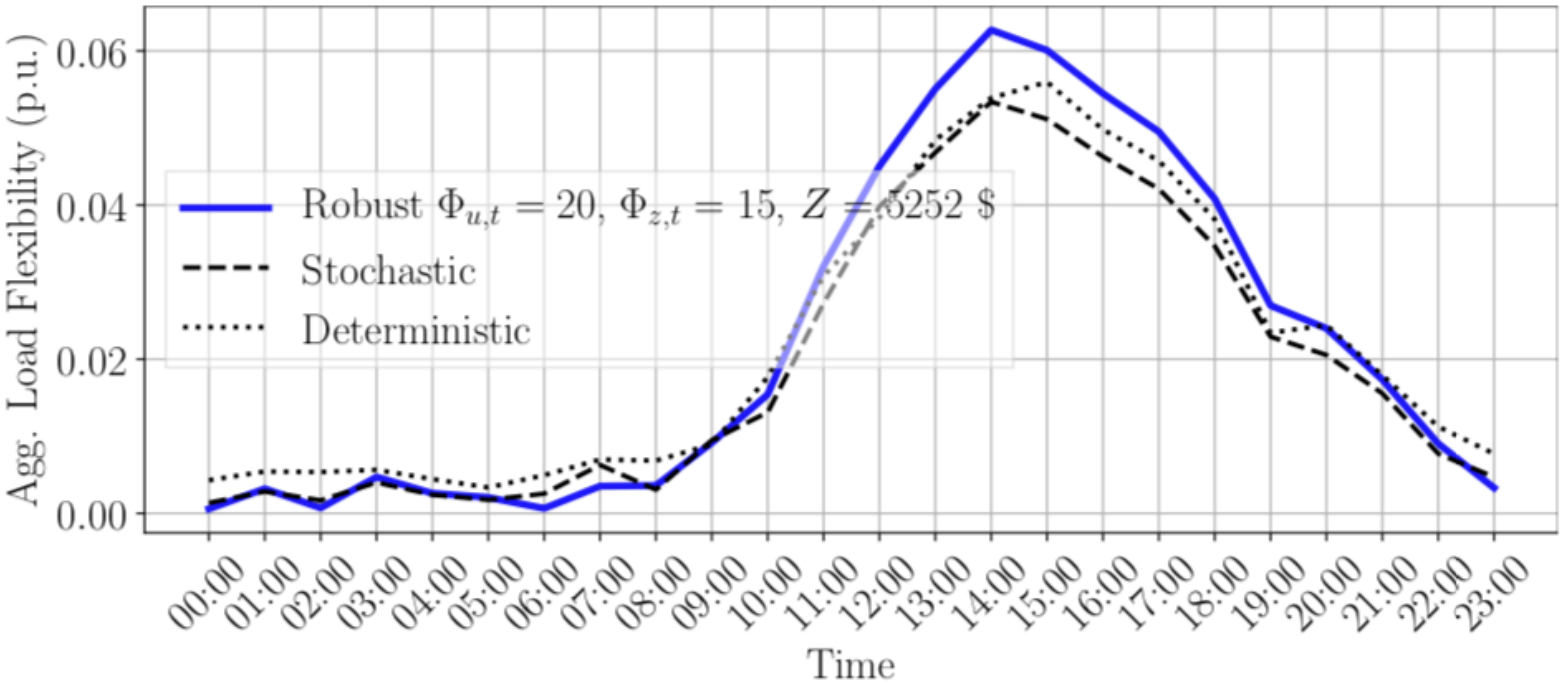


Flexibility Results

- Formulated with Pyomo (python)
- Robust counterpart → MINLP Couenne + RoModel [2]



Flexibility Results: Comparison



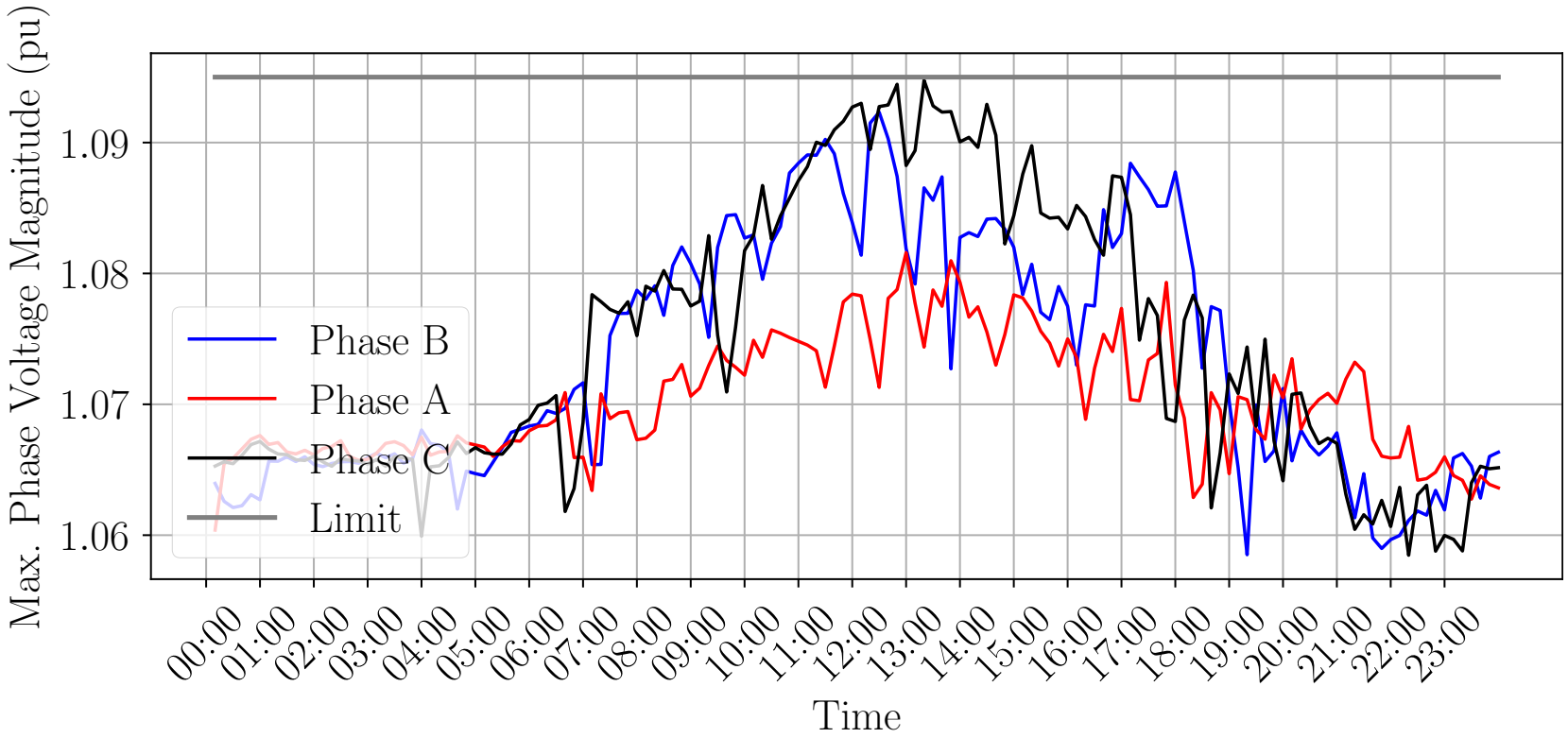
Flexibility results: Comparison

Table 5.3: Results comparison between robust flexibility scheduling and deterministic power flow

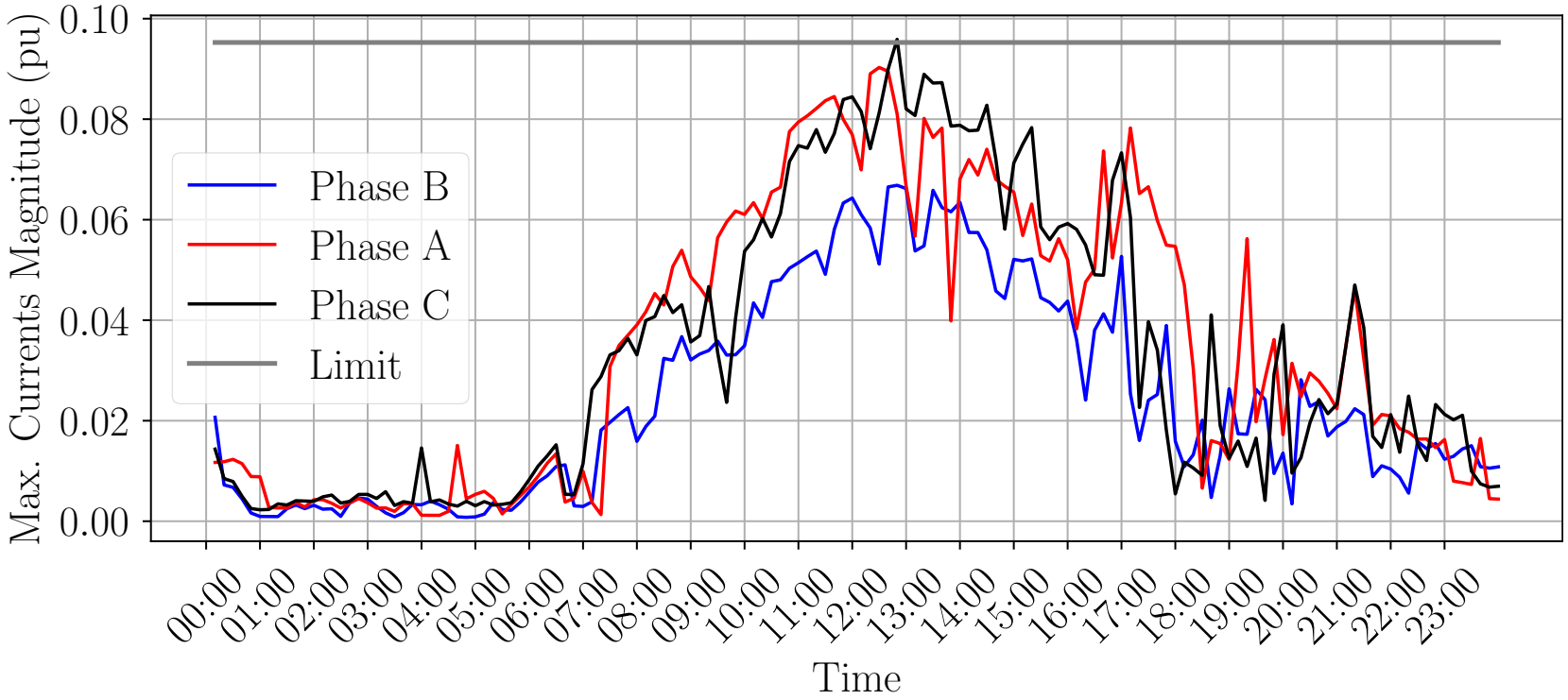
	Robust FS	**Stochastic FS**	**Deterministic FS**
Total Power Losses (kWh)	197.52	229.40	238.67
Total Energy Imp. (kWh)	264.87	332.32	316.45
Total Load Demand (kWh)	1443.78	1314.21	1352.45
Total PV Production (kWh)	1617.20	1439.75	1464.21
Total Flexibility (kWh)	240.77	228.45	189.45
Computation time (min)	674	267	6
Max. phase voltage mag. (p.u.)	1.06	1.09	1.14
Max. phase current mag. (p.u.)	0.07	0.09	0.16

FS: Flexibility Scheduling

Flexibility results: Voltages



Flexibility results: Currents



6. Thesis conclusions: contributions, framework and future research

Thesis Contributions

- **Objective 1**: Power losses estimation in unbalanced LV smart grids under uncertainty conditions.
  - **Contribution 1.1**: An optimization-based procedure to estimate hourly load consumption of non-telemetered customers.
  - **Contribution 1.2**: A Markov chain-based process to estimate the intra-hour high-resolution load demand.

Thesis Contributions

- **Objective 2**: Power losses estimation in unbalanced LV smart grids in large-scale areas with a presence of DERs.
  - **Contribution 2.1** A data mining approach to reduce a high-dimensionality in smart grids feeders dataset, obtaining a reduced set of relevant features.
  - **Contribution 2.2**: A clustering process to obtain representative feeders within a large-scale distribution area of smart grids.
  - **Contribution 2.3**: A deep learning-based power losses estimation model for large-scale LV smart grids.

Thesis Contributions

- **Objective 3**: Power losses minimization in unbalanced smart grids under uncertainty.
  - **Contribution 3.1** A robust optimization model for flexibility scheduling in unbalanced smart grids with distributed resources.

Publications

Publications

Journal Papers

- J-A. Velasco, H. Amaris and M. Alonso, **"Deep learning-Based Power Losses Estimation Model for Large-Scale Distribution Areas,"** *Electric Power System Research* , vol. 121, Oct. 2020, no. 106054.

Conference Papers

- J-A. Velasco, V. Rigoni, A. Soroudi, A. Keane and H. Amaris, **"Optimising Load Flexibility for the Day Ahead in Distribution Networks with Photovoltaics"** in IEEE PES PowerTech 2019, Milano (Italy), July 2019.
- J-A. Velasco, H. Amaris, M. Alonso and M. Casas, **"Energy Losses Estimation Tool for Low Voltage Smart Grids"** in International Conference and Exhibition on Electricity Distribution (CIRED 2019), Madrid (Spain), June 2019.

Publications

- J-A. Velasco, H. Amaris and M. Alonso and M. Miguelez,**"Stochastic Technical Losses Analysis of Smart Grids under Uncertain Demand"** in International Universities Power Engineering Conference (UPEC2018), Glasgow (Scotland), September 2018.
- J-A. Velasco, H. Amaris and M. Alonso and M. Miguelez, **"Energy Losses Estimation in Low Voltage Smart Grids by using Loss Maps,"** in International Conference on Energy, Environment and Economics (ICEEE2018), Edinburgh (Scotland), August 2018.

Thesis Framework

Thesis Framework

National R&D Project: **OSIRIS** (*Optimización de la Supervisión Inteligente de la red de distribución*). 2017. MINECO - Retos Colaboración.

International R&D Project: **IDE4L** (*Ideal Grid for All*). 2017. Seventh Framework Programme (FP7-ENERGY).

Thesis Framework

- **Research Stay** at University Colleage Dublin (UCD).
- Sept.-Dic. 2018.
- Working in flexibility scheduling with the **UCD Energy Institute**:
  - Professor Andrew Keane (Head).
  - Professor Alireza Soroudi.
  - Valentin Rigoni.
- Publication in IEEE PES Powertech 2019.

Future Research

Future Research

- **Non-technical losses** modeling → deep-learning **anomaly detection** techniques.
- **BES systems** modeling → **flexibility providers**.
- **Customer appliances** modeling (such us air conditioners, heaters and hot water devices) → comfort temperatures and the dwelling's consumption.

Thank you very much for your attention

Power Losses Estimation in Low Voltage Smart Grids

Electrical Engineering, Electronics and Automation Ph.D. Program

uc3m | Universidad Carlos III de Madrid

Author
José Ángel Velasco

Advisor
Hortensia Amarís

References

V. H. Sam, "European commission launches proposals to reach 55% emissions reduction by 2030," 2021.

J. Wiebe and R. Misener, "Romodel: Modeling robust optimization problems in pyomo," 2021.