POLYTECHNIC OF TURIN

MASTER's Degree in DATA SCIENCE AND ENGINEERING

MASTER's Degree Thesis

TITLE

Supervisors

Prof. LUCA CAGLIERO

Prof. DAMIEN ERNST

Candidate

MAURIZIO VASSALLO

July 2022

Table of Contents

Abbreviations III

Symbols IV

1 Changes 1
1.1 11/03 . . . 1
1.2 18/03 . . . 1
1.3 25/03 . . . 1
1.4 01/04 . . . 2
1.4.1 Questions . . . 2

2 Introduction 3
2.1 Aim of the thesis . . . 4

3 Background 5
3.1 Power system . . . 6
3.1.1 Description of a power system . . . 6
3.1.2 Power flow . . . 8
3.1.3 Power system reliability . . . 11
3.2 Pandapower . . . 13
3.2.1 Data structure . . . 13
3.2.2 Network models . . . 14
3.2.3 Optimal power flow . . . 15
3.2.4 Time series . . . 16
3.2.5 Other functionality . . . 17

4 Problem analysis 19
4.1 Network topology . . . 20
4.2 Problem statement . . . 20
4.3 Solving methodology . . . 21

5 Project implementation 23
5.1 GYM-ANM 23
5.1.1 Generate dataset 25
5.1.2 Voltage lines forecasting 26
5.2 MV Oberrhein 29
5.2.1 Simbench database 29
5.2.2 Time series 30

References 38

Abbreviations

ANN	Artificial neural network
ANM	Active network management
API	Application programming interface
DER	Distributed energy resources
DES	Distributed energy storage
DSO	Distribution system operator
FACTS	Flexible alternating current transmission system
GARPUR	Generally accepted reliability principle with uncertainty modelling and through probabilistic risk assessment
GHG	Greenhouse gas
MAE	Mean absolute error
OPF	Optimal power flow
PV	Photovoltaic
SLP	Standard load profile
WP	Wind park

Symbols

CO_2	Carbon dioxide
I	Electric current
V	Voltage magnitude
ϕ	Voltage phase angle
pu	Per unit
P	Active power
W	Active power unit: Watts
Q	Reactive power
VAR	Reactive power unit: Volt-Ampere reactive
$\mathcal{G}$	Directed graph
$\mathcal{N}$	Set of positive integers representing the bus (or node) in the network
$\mathcal{E}$	Set of directed edges linking buses together
e_{ij}	Directed edge with sending bus i and receiving bus j
$\mathcal{D}$	Set of all devices
$\mathcal{L}$	Set of all loads

Chapter 1

Changes

1.1 11/03

- Added ’Changes’ chapter (to be removed in the final version)
- Added ’Problem statement’ section **??**
- Added ’MV Oberrhein’ section 5.2
- Minor changes

1.2 18/03

- Added ’Aim of the thesis’ section 2.1
- Added ’Power system reliability’ section 3.1.3
- Minor changes

1.3 25/03

- Added ’N-1 reliability criterion’ sub section 3.1.3
- Modified network elements’ description part 3.1.1
- Added ’Power Flow’ sub section 3.1.2
- Modified ’Problem formulation’ chapter 4
- Added ’Simbench dataset’ 5.2.1 and ’Time series’ 5.2.2 sub sections
- Minor changes

1.4 01/04

- Changed chapter 4 title to 'Problem analysis'
- Added 'Network topology' section 4.1
- Modified 'Problem statement' section 4.2
- Added 'Solving methodology' section 4.3
- Added 'Pandapower' section 3.2
- Modified 'MV Oberrhein' section 5.2
- Minor changes

1.4.1 Questions

- Evaluation of the model issue 5.2.2

Chapter 2

Introduction

The rapid growth of the world population and the limited ability to supply non-renewable energy lead to a rise of power demand, especially in developing countries. The current situation requires an intensive usage of fossil energy causing environmental pollution, as well as climate change.

Indeed, the increase of global temperature and the worsening of the air quality are posing a real problem for the environment. The changes observed in Earth's climate are primarily driven by human activities, particularly fossil fuel burning. The biggest disadvantage of fossil fuels is that during the process of combustion in addition to produce energy, used for everyday purposes, greenhouse gases (GHG) are emitted [1].

In order to improve the situation, the 2015 Paris Agreement set an ambition to limit global warming to well below 2 °C above pre-industrial levels and pursue efforts to limit it to 1.5 °C - in part by pursuing net carbon neutrality by 2050. The substantial reduction of global greenhouse gas emissions (including CO_2) will limit the increase of global temperature [2].
Countries were asked to go through a process of decarbonization: the reduction of carbon dioxide emissions through the use of low carbon power sources.
These sources convert the energy coming through naturals elements (sun, wind, geothermal heat) and convert it in another form of energy, electricity for example, with low or no waste products such as CO_2 or other chemical pollutants.

Thanks to the emerging trend of decarbonization, more and more renewable power energy devices are introduced in the distribution networks [3]. These devices bring in some complications for the distribution of power and voltage in the grids. The networks, that have been designed around the conventional centralized energy production, have to adapt to the new generators in the system. It is possible to say that the distribution networks are moving from unidirectional power flow (from the distribution system to the consumers) to a bidirectional power flow (in this case the consumers are also producers and the exceed energy can be transported from the consumers to the distribution system). This switch from unidirectional to bidirectional power flow requires a smarter system that can handle in an efficient way the production and distribution of voltage.

In the literature, this smarter way to control a distribution system is known as active network management (ANM) and it refers to the design of control schemes that modulate the generators, the loads, and the distributed energy storage (DES), as well as other elements like switches, connected to the grid.

2.1 Aim of the thesis

The aim of this thesis is to exploit machine learning techniques to forecast the future distribution of voltages based on historical measurements in a medium-voltage distribution system in order to predict under or overvoltages of the lines and the possible consequences related to these issues.

Chapter 3

Background

The voltage control problem has been studied for years, but it only comes under the spotlight in recent years for the increasing number of distributed resources introduced in the networks. It is important to control the voltage in an electrical power system for a regular operation of the electrical equipment, to prevent damage such as overheating of generators and motors, to reduce transmission losses and to maintain the ability of the system to last and prevent voltage collapse.

In particular, it is useful and needed to reduce the output of renewable generators from what they could otherwise have produced given the available resources, often referred to as the process of **curtailment**. Such generation curtailment, along with storage and transmission losses, constitute the principal sources of energy loss that could be minimised with active network management ANM [4].

Controlling the voltage in an active way has many interesting properties:

- It is a combination of local and global problem: the voltage at each node is influenced by the powers of all other nodes, but the impact depends on the distance between them.
- It is a constrained optimization problem where the constraint is to keep the voltage in a given range and the objective is to minimise the total power loss.
- Voltage control has a relatively large tolerance, and there are no severe consequences if the control fails to meet the requirements for short periods of time. [5]
- It is a hierarchical problem where much information is available at the top of the pyramid (distribution stations and substations) and they decrease at the base of the pyramid (houses, factories) mainly due to the absence of many sensors.

3.1 Power system

3.1.1 Description of a power system

Power transmission systems consist of an interconnected set of overhead lines, cables and related equipment that are used for the transfer of electricity at high voltage levels between supply points and load points, such as customers and other electric systems.

In the traditional power system, electricity is produced in large, centralised power plants. The electricity is then transferred to the loads using the transmission and distribution networks. High and extra-high voltages are associated with supply transmission from the power plant. The reason for transmitting power at high and extra-high voltage levels is to increase efficiency. The lower current accompanying the high voltage transmission allows for the use of thinner, lighter-weight cables. This reduces the cost in the tower and electrical line construction. In Belgium, High and extra-high voltages refer to voltage magnitudes $30kV \leq V < 380kV$ for the high voltage and $V \geq 380kV$ for the extra high voltage [6]. (*Symbols problem: V for voltage magnitude and V for volts?*)
Large industrial complexes and factories that require a substantial amount of power often utilise medium supply voltages. The high voltage coming from the power plant is sent to the primary substation, this can supply step-down power to secondary substations or to single buildings. Secondary substations can have transformers to further step down the power, and they are generally located in areas that can serve one or more buildings. Medium voltages refer to voltage magnitude $0.4kV \leq V < 30kV$.
Then the medium supply power is step down again to a low voltage and sent to the domestic household or home appliances power supply. Low voltages refer to voltage magnitude $V < 0.4kV$.

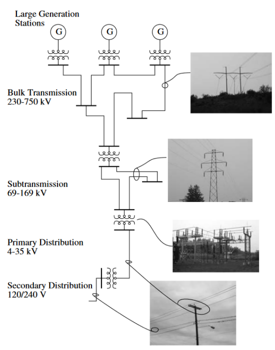

Figure 3.1: Power network distribution [7]

A power system is usually made up of the following main elements:

- **Generator**. These produce energy, converting a form of energy into electricity. Many generators produce energy using turbines: a mass of air or water spin the generator's blades, producing electricity. This mass can be natural: hydroelectric, wind or geothermal turbines, or generated by combustion of some fuel, for example natural gas or nuclear source.
  There are other generators that do not need a turbine to generate electricity, for example solar panels.

- **Transformers**. Transformers are used to interlink systems operating at different voltages. These can increase the voltage magnitude near a generator power plant or decrease it near the consumptions facilities.
  Changing the voltage magnitude allows reducing the power loss due to transportation: the power lost is given by $P = VI$, where V is the voltage and I is the current, so decreasing the voltage can reduce the energy loss.

- **Lines**. These transport the power from where energy is generated to where it is consumed. One of the main issues about transportation lines is insulation. There are different types of lines: overhead cables, they use air to insulate the bare conductors or underground cables, for these cables particular attention must be taken to insulate them from other conductors and from the earth (ground). Also, the material used must be resistant to damages, corrosion, and it must avoid that the water is being absorbed.

- **Switchgear**. In an electricity supply, it is necessary to disconnect equipment from the network quickly if a fault occurs to avoid damage on the elements of the network, or to disconnect some points of the network to avoid excessive losses or too high or low voltages.
  Switchgear is a broad term that describes a wide variety of switching devices that fulfil the need of controlling, protecting, and isolating power systems. Among these switching devices the most common are: *circuit breaker*, during an electrical fault, a circuit breaker will detect the anomaly and interrupt the power flow, effectively limiting damage to the system; *switch* is an electrical component that can disconnect or connect the conducting path in an electrical circuit, interrupting the electric current or diverting it from one conductor to another; *recloser* similar to the circuit breaker but used in high voltage networks, these devices handle trouble temporary occurrences such as lightning, windblown tree branches or wires, birds, or rodents damaging the wires.

- **Loads** are electric components that consume the electric power generated by the generators. The type of loads can be divided base on the consumption in:

Domestic loads, the domestic loads mainly consist of lights, fan, refrigerator, air conditioners, mixer, grinder, heater, ovens, small pumping, motor, etc. The domestic loads consume very little power.

Commercial loads, the commercial loads mainly consist of lightning, fans, heating, air conditioning and many other electrical appliances used in establishments such as markets, restaurants, shops. This type of load occurs for more hours during the day as compared to the domestic load.

Industrial loads, the domestic loads refer from a small-scale industry, to a heavy industry. It includes all electrical loads used in industries along with the employed machinery. Industrial loads may be connected during the whole day [8].

3.1.2 Power flow

An important procedure in power system networks is to perform a numerical analysis to determine the electrical state of the network, starting from some parameters that are known. This analysis is called optimal power flow (OPF).
The OPF is an optimization problem essential for planning purposes: besides the calculation of electrical characteristics of the power system, power flow analysis can also help to optimize the system operating conditions, minimize the power losses and determine control actions to satisfy the demand while meeting operational constraints.

Buses

The power flow gives information about the steady state of the entire system such as voltage, real and reactive power, line loadings and power injection.
Each bus is associated with four quantities: voltage magnitude V, phase angle ϕ, real power P and reactive power Q. Depending on the quantity that have been specified, buses in the power system are classified into the following three different types:

- **Slack bus**. It is taken as reference where the magnitude and phase angle of the voltage are specified. Slack bus magnitude considers 1 pu and phase angle 0 degrees. This bus provides the additional real and reactive power to supply the transmission losses, since there are unknown until the final solution is obtained.

- **Load buses or PQ bus**. At these buses, the real and reactive powers are specified. The magnitude and phase angle of the bus voltage are unknown until the final solution is obtained.

- **Voltage controlled buses or PV bus**. At these buses, the real power and voltage magnitude are specified. The phase angles of the voltages and the reactive power are unknown until the final solution is obtained. The limits on the value of reactive power are also specified.

Solution techniques

Defining and solving the power flow equations are the main tasks in load flow analysis.
The definition of the power flow equations is based on Ohm's Law, which is the relationship between voltages and currents. For a network, it can be expressed in matrix notation as follows:

$$\mathbf{Y} \times \mathbf{V} = \mathbf{I}$$

$$\begin{bmatrix} Y_{1,1} & \cdots & Y_{1,N} \\ Y_{2,1} & \cdots & Y_{2,N} \\ \vdots & \ddots & \vdots \\ Y_{N,1} & \cdots & Y_{N,N} \end{bmatrix} \times \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_N \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix}$$

Where:

- **Y** is the bus admittance matrix
- **V** is an array of bus voltages
- **I** is an array of bus current injections (positive value when generation, and negative value when load)

The power flow formulation is based on the application of Kirchhoff's laws to meshed electric networks. The basic concept is that the sum of all flows into each and every node should be equal to zero.
The flows are in complex form, they consist of real and reactive components, or Ws and VARs. If there are n nodes, then there are n complex equations. The resulting system of equations involves non-linear relationships, making the calculations not easy. Solution methods are primarily iterative with the objective of reducing the sum of flows in all nodes to some acceptably small value known as the mismatch tolerance.

All these iterative methods follow the same basic concept: they assume starting values for the dependent variables, primarily voltage at nominal voltage magnitude (i.e. 1 pu) and zero phase angle; compute new values for those voltages using the nodal network equation or a numerical approximation and repeat until the convergence criteria are met.

The solution has to satisfy some network constrains, in particular:

- Active and reactive power balance: the sum of the power injections (that can be positive or negative) at each bus must be equal to 0. This, as said, results from the Kirchhoff's laws.
- Voltage limits: the voltage magnitude at each bus and the voltage phase difference between two directly connected buses are bounded by some specific values to maintain safe operation.
- Thermal limits on transmission lines: the flow in each transmission line is limited due to the thermal limit of the conductors.
- Generator active and reactive power limits: the generating units have generally a minimum and maximum level of output power.
- Generator ramping limits: the output power of a generating unit can not be instantaneously increased or decreased. The operator must take into account the ramping limits of the generators.

Convergence

The OPF is a non-linear and non-convex optimization problem, with a large number of constraints (both equality and inequality constraints) and variables (that can be both continuous and discrete). It is therefore a hard problem, whose solution cost increase exponentially, particularly with the increasing size of the network. Moreover, there is no guarantee to find the global optimum.
When a solution exists and it is reached, it is said that the network has converged. Convergence is the state when all nodes have met the mismatch tolerance. The main power flow solution methods are:

- Gauss-Seidel method updates the voltage one node at a time until all nodes are within the mismatch tolerance.
- Newton-Raphson method uses a first order expansion of the power flow equations to approach convergence. Generally faster than the Gauss-Seidel method and able to converge to small tolerances. However, the method is prone to the phenomenon of **divergence**, when mismatches increase instead of decrease from iteration to iteration. This occurs when the solution vector exits outside the feasible solution space at any point during the algorithm. Once outside

feasible space, the solution gradient tends to further increase mismatches, leading to solutions that “blow-up” in the numerical sense. This method requires calculating the first order approximation matrix (known as the Jacobian). Several variations on the Newton-Raphson are in use, including:

Fast Decoupled: separates the loosely linked real and reactive components of the power flow equations in order to speed up solution.

Fixed Newton: does not update the first order approximation matrix every iteration to reduce computational burden.

Non-divergent power flow: applies a reduction to the Jacobian multiplier whenever the solution appears to exit feasible space. In certain situations, this may prevent divergence, or at least stop it before blow-up.

- Interior-Point Newton method forces the solution inside feasible space to avoid divergence. The interior point method uses a second order expansion of the power flow equations as a basis for its algorithm. The method is more computationally intensive than either the Gauss-Seidel or Newton-Raphson, but is less susceptible to numerical divergence.

Divergence

Divergence is the condition of the power network when the numerical solution can not be found any more due to some possible issues:

- the power system is going to “blow-up.”
- the power system is in voltage collapse.
- the power system is unstable.
- the initial conditions defined were bad or poor.
- some issues related to software or input data.

Divergence of the power flow solution has traditionally been associated with the singularity of the Jacobian matrix. Since some methods require an inverse of the Jacobian as part of its solution algorithm, singularity of the Jacobian means division by zero [9].

3.1.3 Power system reliability

Reliability is an important factor concerning the quality of energy supply.
Power reliability can be defined as the degree to which the performance of the

elements in a system results in electricity being delivered to customers within accepted standards and in the desired amount [10].
Reliability indices typically consider such aspects as:

- the number of customers;
- the connected loads;
- the duration of the interruption measured in seconds, minutes, hours, or days;
- the amount of power interrupted;
- and the frequency of interruptions.

These factors depend on variable such as reliability of individual items of equipment, circuit length and loading, network configuration, distribution automation, and available transfer capacity [8].

For reliability purposes, it is important to know the maximum voltage that can be transferred with transmission lines to meet the anticipated load demand. It is also important to know the levels of power through various transmission lines under certain contingency outage conditions to maintain the continuity of service. Knowledge of power flows and voltage levels under normal operating conditions are necessary in order to determine fault currents and the ensuing consequences on the stability of the system [9].

[American National Standards Institute (ANSI) C84.1-2016 [4], voltage standards for service voltage limits are classified as Range A and Range B limits. The voltage between 0.950 p.u. and 1.050 p.u. of nominal voltage lies under Range A, and the voltage between 0.917 p.u. and 1.058 p.u. of nominal voltage for 240 V service voltage lies under Range B. Note that the voltage can be within Range B for only a short duration and frequency, and thus corrective measures are necessary to constrict]

Reliability criteria

The goal of a distribution system operator (DSO) is to ensure a reliable system. Unfortunately, a completely reliable electricity supply is not feasible to obtain since it comes at an infinite cost. So, network operators need to determine an acceptable reliability level, by balancing the costs and benefits, where acceptable reliability level means that all the elements in a network have an acceptable voltage range.

The European GARPUR project (**G**enerally **A**ccepted **R**eliability **P**rinciple with **U**ncertainty modelling and through probabilistic **R**isk assessment) developed reliability management approaches and criteria. One of these criteria used by system operators is the N-1 criterion.

The basic principle of N-1 security in network planning states that if a component, for example a transformer or circuit, should fail or be shut down in a network operating at the maximum levels of transmission and supply, the network security must still be guaranteed. This means that the safety of the system is guaranteed and the spreading of the failure is avoided.
It is possible that there may be another contingency before restoring the network after the fail of one element, this criterion is known as N-1-1 criterion.

With the increasing of network complexity more than one element may fault, for this reason there exists other levels of reliability, like the N-2 criteria. In this case, even if in the network two components fail, the network security is guaranteed.
This N-2 criteria requires much more computational power since, the system operator must calculate what happens to the network for any combination of two fault elements. So, the problem becomes a combination problem, where the possible combination are given by: $\binom{N}{2}$, with N the number of elements in the network.

In general, the calculation can be extended to any generic k elements, but the complexity of the problem increases with the value of k. Indeed, the possible combination in a N-k contingency are: $\binom{N}{k}$, with $2 < k < N$

3.2 Pandapower

This thesis project will be developed with the help of Pandapower.
Pandapower is a Python based power system analysis tool aimed at automation of static and quasi-static analysis and optimization of power systems [11].
Pandapower is a powerful tool that allows to easily create a model for any power network using customizable predefined data structures, it can solve the OPF problems, perform the state estimates, topological graph searches and diagnose the system for possible errors.

3.2.1 Data structure

Pandapower is based on a tabular data structure, where every element type is represented by a table that holds all parameters for a specific component. After the calculation of the power flow, a result table, which contains the element specific results of the different analysis methods, is added to the structure.
The tabular data structure is based on the Python library pandas. It allows storing variables of any data type, so that electrical parameters can be stored together

with status variables and meta-data, such as names or descriptions. The tables can be easily expanded and customized by adding new columns without influencing the Pandapower functionality. All inherent pandas methods can be used to efficiently read, write and analyse the network and results data.

A Pandapower network is a Python dictionary that holds all information about the network. Most importantly, it includes element and a result tables for each element type, such as line, transformer, switch, loads. The element table holds all input parameters that are specified by the user, while the result table is used to store the results of the power flow calculation. Input and output parameters are identified by the same index in both tables [11].

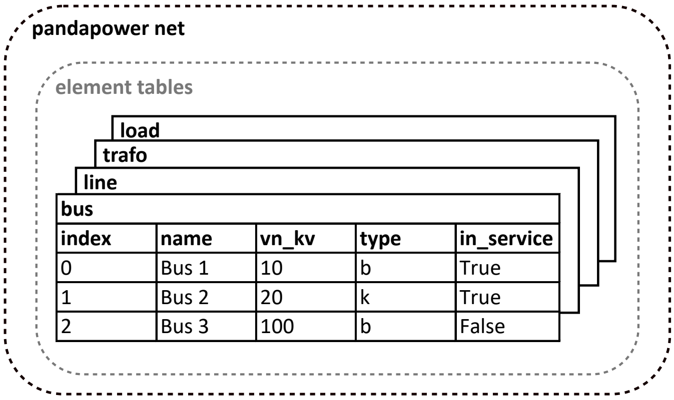

Figure 3.2: Pandas data frame representation of the Pandapower network

3.2.2 Network models

There are two main ways of how a power system can be defined. A commonly used approach is the bus-branch model (BBM), which defines the network as a collection of buses which are connected by generic branches. Branches are modelled with a predefined equivalent circuit and are used to model multiple elements connected to that branch (multi-pole), like lines or transformers. Buses are attributed with power injections to model single-pole elements like loads, generators or capacitor banks. Since the BBM is an accurate mathematical representation of the network, electric equations for power systems analysis can be directly derived from it, but the need to calculate the impedances for each branch and summed power injections at each bus manually can be cumbersome and error-prone, especially for complex elements.
Instead of a BBM, Pandapower uses an element-based model (EBM) to model electric grids. An element is either connected to one or multiple buses and is defined with characteristic parameters. This allows defining the network parameters, such as

length and relative impedance for lines, or short circuit voltage and rated apparent power for transformers. While BBM allows only the definition of a summed power injection at each bus, single-pole elements (such as load or generation elements) can be connected to buses independently. This also allows connecting multiple elements at one bus. The element models are then processed internally with the appropriate equivalent circuits to derive a mathematical description of the grid [11].

3.2.3 Optimal power flow

The power flow is the most important electric analysis function for power system planning. It allows calculating the current flows and voltages in the network.
The Pandapower power flow solver is based on the Newton-Raphson method. The implementation is based on PYPOWER Python library. To solve the OPF, the bus constraints include maximum and minimum voltage magnitude, active and reactive power limits can be defined for PV and slack-elements like external grids and generators, but also for PQ-elements, such as loads and static generators.
After defining all the network elements, to run the power flow solver it is just need to execute the command:

```
pandapower.runpp(net, ...)
```

this function takes as input the Pandapower network data structure and some other optional values (for example the algorithm solver, max number of iterations, tolerance and so on).

Internally, Pandapower solves the following optimization problem:

$$
\begin{aligned}
\min \quad & \sum_{i \,\in\, gen, sgen, load, external grid} P_i \cdot f(P_i) \\
\text{s.t.} \quad & \\
& \text{load flow equations} \\
& \text{branch constraints} \\
& \text{bus constraints} \\
& \text{operation power flow equations}
\end{aligned}
$$

where P_i is the active power of any element and $f()$ is the cost function. Few example of the possible constrains are:

$$
\begin{aligned}
P_{min,i} \leq P_g \leq P_{max,i},\ g \in gen \\
Q_{min,i} \leq Q_g \leq Q_{max,i},\ g \in gen \\
V_{min,i} \leq V_g \leq V_{max,i},\ i \in bus
\end{aligned}
$$

It is possible to customize the cost function and choosing between a piece-wise or polynomial cost function. Detailed information about the optimization problem, cost function and network constrains are available in the Pandapower documentation [12].

After running the power flow calculation, new tables are added to the network data frame.

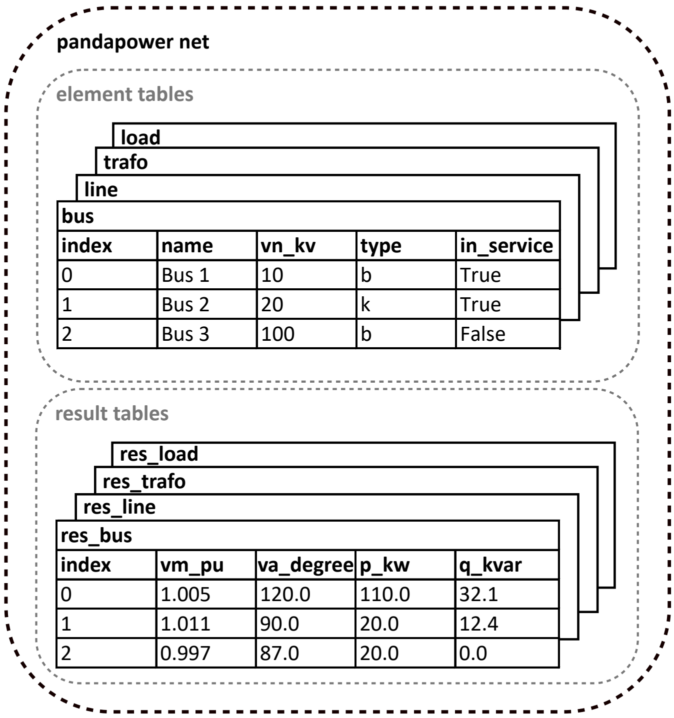

Figure 3.3: Pandas data frame representation of the Pandapower network after the power flow calculation

3.2.4 Time series

Pandapower allow running time series analysis for a given network. There are two main requirements for time series calculations:

- a Pandapower network

- some time series (in a pandas data frame for example)

To execute the time series calculation, the loads, generators and other elements' active and reactive power time series have to be passed to a controller that will be in charge to change the elements' values according to the time series.

The time series calculation can be run with the command:

```
pandapower.timeseries.run_time_series.run_timeseries(net, ...)
```

this command will start a loop that iterates over every **time_step**. For each step, a control loop is started for each controller by **run_control**. The controller updates the elements' values at each step with the values given in the time series.

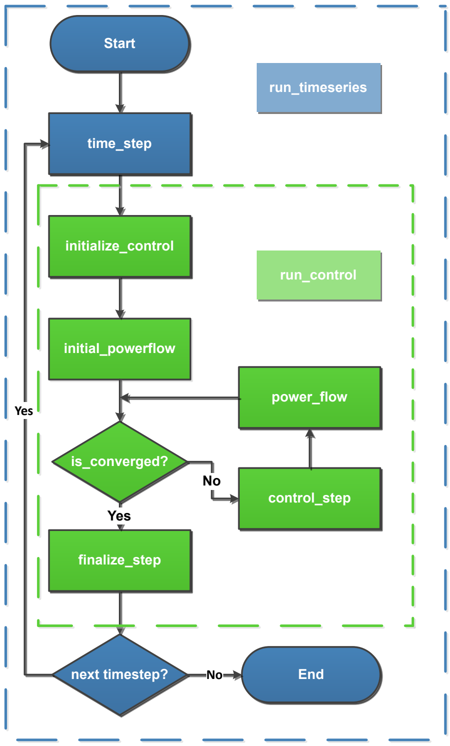

Figure 3.4: Pandapower time series calculation loop [13]

After each step, the elements' values are stored in an output writer object and this allows, after the full calculation is finished, to easily save the values on disk.

3.2.5 Other functionality

Pandapower has some other features:

- Predefined Networks. In addition to creating custom networks through the application programming interface (API), 66 predefined, published test and benchmark networks can be directly accessed through Pandapower. One of these networks, MV Oberrhein, is the one used in this thesis.

- Plotting features. Pandapower comes with extensive plotting features using the Matplotlib library. All Pandapower elements can be translated into different Matplotlib collections that can be customized with respect to shape, size and colour to allow highlighting and create individual network plots. It is also possible to use colour maps to codify information, like the loading of lines or the voltage at buses.

- Converter. Pandapower includes converters in order to export a Pandapower grid as a MATPOWER or PYPOWER casefile or the other way.

Chapter 4

Problem analysis

The principle of active network management ANM is to address congestion and voltage issues via short-term decision making policies [14].
ANM creates a smarter network infrastructure providing automated control of various components in the network and provides the information needed to ensure that every device performs in an optimal manner. This automated control allows grid companies to avoid reinforcing the network with expensive upgrades, so reducing the costs. For example, in case of energy generation from the renewable devices higher than what a single line can handle, a grid company, to avoid congestions and possible overvoltages, has three main options:

- Replace the existing line with a line that can handle a higher voltage.
- Add another parallel line.
- Use ANM.

The first two solution require some infrastructure investment that can be expensive and problematic, especially in the case of overhead or underground lines.
The solution with ANM does not require construction cost for the grid company; to keep the network working, in this case, the output of the renewable devices can be curtailed to reduce lines' overloading.

In these references, ANM schemes maintain the system within operational limits by relying on the curtailment of the generator devices, PVs, WPs and other DER devices.
Curtailment of renewable energy may be considered as counter-intuitive on the environmental point of view, and it may be considered as last option. Indeed, this process can slow down the switch to clean energy, because of the lost of the curtailed energy.

In this mindset, ANM could also be used to control flexible loads and reduce the curtailment effects. These flexible loads, also known as virtual batteries, such as water heaters, air conditioning systems, electric vehicles, can be controlled to be turned on if the energy production is higher than the energy consumption so to avoid curtailment on the generators [15].
Another way to reduce the energy curtailment is to use Flexible Alternating Current Transmission System (FACTS) devices. They offer some level of power flow control and enhance the transfer capability over the existing network. This flexibility can be utilized for congestion mitigation and renewable energy integration. Particularly, FACTS devices allow controlling all parameters that determine active and reactive power transmission: nodal voltages magnitudes and angles and line reactance. These devices replace the mechanical switches with semiconductor switches allowed much faster response times. One problem with these devices is the cost a system operator should sustain to implement them in the network [16].

4.1 Network topology

A distribution network can be represented as a direct graph $\mathcal{G}(\mathcal{N}, \mathcal{E})$, where $\mathcal{N}$ is a set of positive integers representing the buses (or nodes) in the network, $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of directed edges linking two buses together. The notation $e_{ij} \in \mathcal{E}$ refers to the directed edge with sending bus i and receiving bus j. Each bus might be connected to several devices, which may inject or withdraw power from the grid. The set of all devices is denoted by $\mathcal{D}$ that can be either loads $\mathcal{L}$ or generators $\mathcal{G}$.
Several variables are associated with each bus $i \in \mathcal{N}$: a bus voltage $\mathcal{V}_i$, a bus current injection I_i, an active power injection P_i and reactive power injection Q_i. The complex powers S_i, $S_d \in \mathbb{C}$ injected into the network at bus i, or device d, can be obtained by the relation $S_i = P_i + \mathbf{i}Q_i$ or $S_d = P_d + \mathbf{i}Q_d$.
Similarly, variables I_{ij}, P_{ij}, Q_{ij} and S_{ij} refer to the direct flow of the quantities in branch e_{ij} as measured at bus i [4].
I think this part should be somewhere but I am not sure where to put it. Here it looks like it is not properly linked to the previous or following part.

4.2 Problem statement

This thesis will focus on ANM and the problem faced by a system operator to maintain the network within its operational limits. In particular, the DSO will evaluate whether in a given moment there will be a voltage problem and to apply (*maybe*) curtailment to generator devices to maintain the voltage inside a safe range.

The DSO knows some information about the network and several sets of measurements:

- The network topology: the number of lines, buses, loads and the generators, the lines' length, the distance between buses, and the distance between each load and generator from the external grid. Moreover, lines' impedance are known.
- The active and reactive power of the loads at each time step (*due to uncertainty, some measure may be missing*).
- The type of DER device, their active power for each time step and the maximum power they can generate.

The DSO will consider the behaviour of the network over a set of discrete time steps $t \in 1,2,...,T-1,T$ with $T \in \mathbb{N}$ and he will predict if the system in a given time step $t+1$ will be in a critical condition, knowing the system information of the previous time steps. In particular, the DSO will consider only the r preceding steps. The system critical state C_{t+1} can be formulated as follows:

$$C_{t+1} = f(S_{t+r-1}, S_{t+r-2}, \dots, S_{t-1}, S_t) \tag{4.1}$$

where:

- S_i is the state of the system at one generic time step i, $S_i \in \mathbb{R}^n$ with n the number of elements.
- $f(\dots)$ is a forecasting function that maps the input state system to the critical system evaluation $f : \mathbb{R}^{r,n} \mapsto \{0,1\}$.
- C_{t+1} is the forecasted value of the system ($C_{t+1} \in \{0,1\}$) stating if the system is critical ($C_{t+1} = 1$) or not ($C_{t+1} = 0$).

The state of the system is represented by the loads and generators' active and reactive power and/or the buses' voltage magnitude:

$$S_t = [V_t^1, V_t^2, \dots, V_t^{n-1}, V_t^n] \tag{4.2}$$

where V_t^i represents the voltage magnitude at time t for bus i.

4.3 Solving methodology

It is common in time series forecasting problems to use artificial neural networks (ANNs) to find a solution, thanks to their capacity to learn an approximate mapping function from the input space to the output space. In this case, the ANN will take as input the information from the network, and it will output a binary value, stating if there will be or not a critical situation in the network.

Model inputs consist of state variables from time instance $t-r+1$ to t. According to 4.1 and 4.2, the input can be expressed as a matrix of state variable:

$$\begin{aligned} \text{input} &= [S_{t+r-1}, S_{t+r-2}, \ldots, S_{t-1}, S_t] \\ &= \begin{bmatrix} V^1_{t-r+1} & V^1_{t-r+2} & \cdots & V^1_{t-1} & V^1_t \\ V^2_{t-r+1} & V^2_{t-r+2} & \cdots & V^2_{t-1} & V^2_t \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ V^{n-1}_{t-r+1} & V^{n-1}_{t-r+2} & \cdots & V^{n-1}_{t-1} & V^{n-1}_t \\ V^n_{t-r+1} & V^n_{t-r+2} & \cdots & V^n_{t-1} & V^n_t \end{bmatrix} \end{aligned} \tag{4.3}$$

Given this input, the ANN must output binary value stating whether the system is safe or in critical situation at the time step $t + 1$.

The strategy behind this thesis is to use supervised learning techniques that may extract optimal solutions, decision-making patterns to be applied to the network.

Chapter 5

Project implementation

The project is developed in Pandapower a, Python based, power system analysis tool aimed at automation of static and quasi-static analysis and optimization of balanced power systems [11].

5.1 GYM-ANM

[*Old part, you can skip this section and jump to 5.2*] The network employed for these experiments is the network used in the paper [4]; it consists of: one external grid, five buses, one transformer, four lines, three loads, two static generators and one DES device.

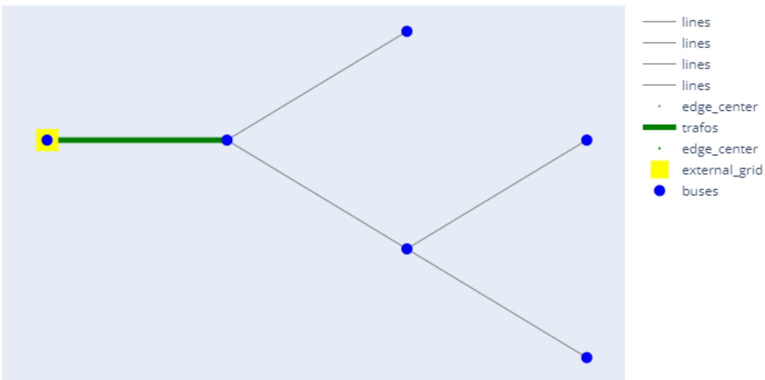

Figure 5.1: GYM-ANM network

Following the paper, 3 situations are tested obtaining similar results:
Situation 1
This situation (figure: 5.2) characterises a windy night, when the consumption is low, the PV production null, and the wind production at it is near maximum. Due

to the very low demand from the industrial load, the wind production must be curtailed to avoid an overheating of the transmission lines connecting buses 0 and 4.

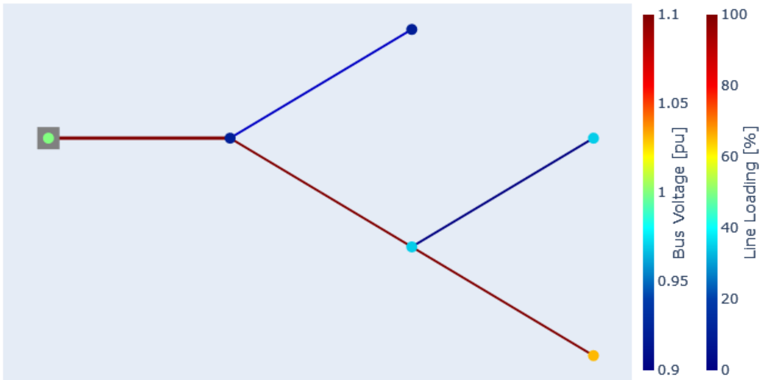

Figure 5.2: GYM-ANM network situation 1. Interactive image at the following link

Situation 2

In this situation (figure: 5.3), bus 5 is experiencing a substantial demand due to a large number of EVs being plugged-in at around the same time. This could happen in a large public EV charging garage. In the morning, workers of close-by companies would plug in their car after arriving at work and, in the evening, residents of the area would plug in their cars after getting home. In order to emphasise the problems arising from this large, localised demand, we assume that the other buses (3 and 4) inject or withdraw very little power into/from the network. During those periods of the day, the DES unit must provide enough power to ensure that the transmission path from bus 0 to bus 5 is not overrated, which would lead to an overheating of the line. For this to be possible, the agent must strategically plan ahead to ensure a sufficient charge level at the DES unit.

Situation 3

Situation (figure: 5.7), represents a scenario that might occur in the middle of a sunny windy weekday. No one is home to consume the solar energy produced by residential PVs at bus 1 and the wind energy production exceeds the industrial demand at bus 2. In this case, both renewable generators should be adequately curtailed while again storing some extra energy to anticipate the EV late afternoon charging period, as depicted in Situation 2.

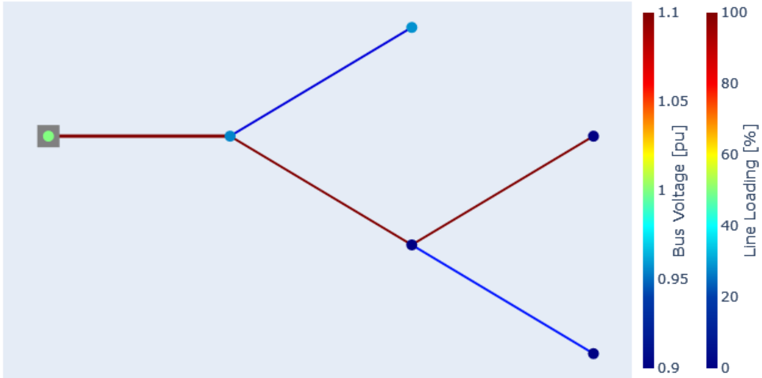

Figure 5.3: GYM-ANM network situation 2. Interactive image at the following link

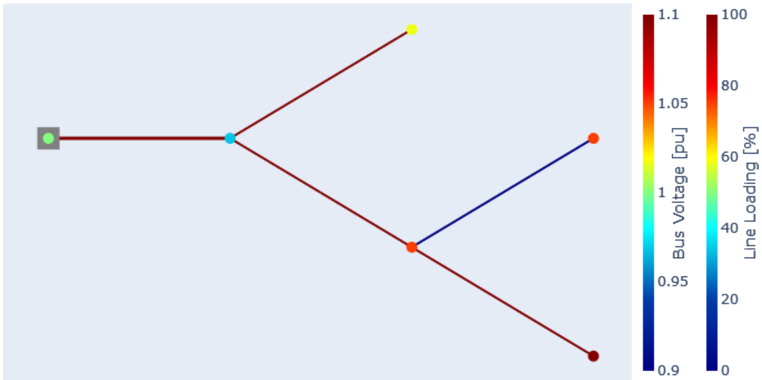

Figure 5.4: GYM-ANM network situation 3. Interactive image at the following link

5.1.1 Generate dataset

Pandapower allows running time series simulations of a network. When calling the time series function, a loop starts to iterate for each time step and the power flow of the network is calculated.
For this simulation, a time Δt of 15 minutes if used for a temporal window of one year, in total of 35,040 time steps. To be compliant to Pandapower requirements, the time series must be passed to a controller that will change the network's devices values for each loop.
The results in this section are obtained with the following time series:

Algorithm 1 Time series of active power ?? for loads $\mathcal{L}$ and generators $\mathcal{G}$

$n\_timesteps = 365 * 24 * 4$

$\triangleright$ Loads active power

$Load0\_p = cos(range(n\_timesteps))^2 * rand(n\_timesteps) * 4$

$Load1\_p = normal(7{,}0.5, size = n\_timesteps)$

$Load2\_p = normal(14{,}0.5, size = n\_timesteps)$

$\triangleright$ Generators active power

$PVgen\_p = normal(4{,}0.5, size = n\_timesteps)$

$Windgen\_p = cos(range(n\_timesteps))^2 * rand(n\_timesteps) * 12$

After the simulation is run, some output values are exported as *.xlsx* files and these constitute the dataset, the story of the network. These output values are:

- loads active power in MW.
- buses voltage magnitude in pu.
- the lines loading in percentages and the current in kA.

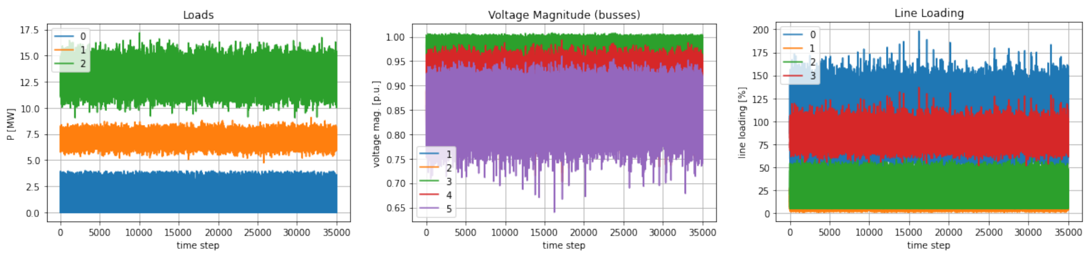

Figure 5.5: Loads active power, buses voltage magnitude and line percentage loading

5.1.2 Voltage lines forecasting

Splitting the dataset

The aforementioned dataset is divided in training, validation and test set in 70, 20 and 10 percent respectively, without a random shuffle before the splitting. This has two main advantages:

1. It ensures that chopping the data into windows of consecutive samples is still possible.
2. It ensures that the validation/test results are more realistic, being evaluated on the data collected after the model was trained.

Normalization

The data is normalised subtracting the mean and dividing by the standard deviation of each feature. The mean and standard deviation are computed using the training data so that the models have no access to the values in the validation and test sets.

Windowing

The training, validation and test sets are divided in windows to then fed to the neural network.
The main features of the windows are:

- The width (number of time steps) of the input and label windows.
- The time offset between them.
- Which features are used as inputs, labels, or both.

The result are obtained using an input window of 12 time steps (3 hours in the past) and an output window of 6 time steps (1.5 hours in the future). The offset is 1, this means that the predictions are from time step 13 to 18. The training features are:

$[load0\_p, load1\_p, load2\_p, PVgen\_p, Windgen\_p, L0, L1, L2, L3]$

and the label features are:

$[L0, L1, L2, L3]$

where Li is the loading in percentage of line i.

Training

The model used for training is an artificial neural network with two hidden layers with 128 and 64 neurons respectively.
The neural network is trained for 20 epochs with Adam optimizer and early stopping callback.

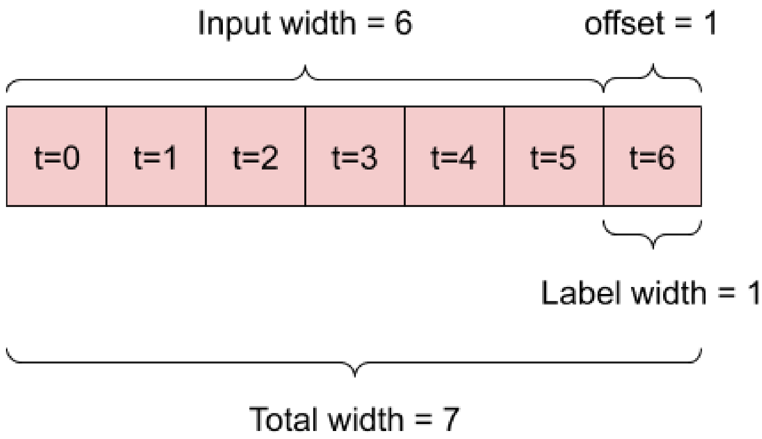

Figure 5.6: Window visual representation for the input and output time steps TODO: add an image that shows the real situation (This is taken from Tensorflow website)

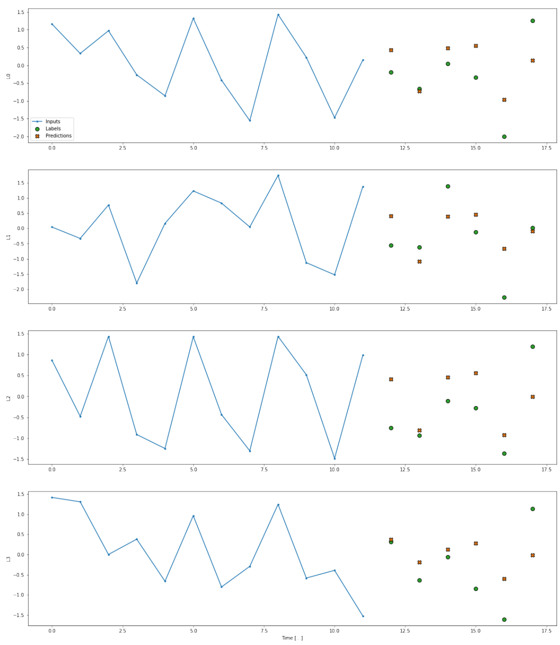

Figure 5.7: Forecasting of the values
TODO: denormalise the data for visualisation purposes. Decrease height or split it: too large for this template (it messes up with the layout)

The metric used for the evaluation is the mean absolute error (MAE):

$mean\_absolute\_error : 0.6201$
$\triangleright$ MAE for each feature
$\{L0 : 0.5625071, L1 : 0.613873, L2 : 0.5569194, L3 : 0.7471818]$

Given this voltage forecast, it would be possible to understand if the network risks overloading for the next time steps and proceed accordingly to avoid damaging the network.

5.2 MV Oberrhein

The network used for these experiments is the MV Oberrhein network from Pandapower; a generic 20 kV network serviced by two 25 MVA HV/MV transformer stations. The network supplies 141 HV/MV substations and 6 MV loads through four MV feeders. The network layout is meshed, but the network is operated as a radial network with 6 open sectioning points.
To simplify the situation, the network can be divided in 2 independent parts [ref].

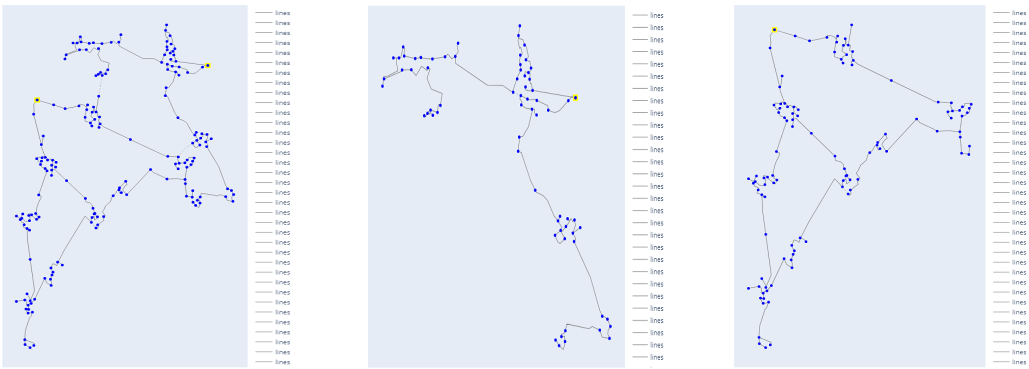

Figure 5.8: MV Oberrhein network. Used network: middle one
TODO add letters

The model consist of: one external grid, one transformer, 70 buses, 61 loads and 60 renewable generators.

5.2.1 Simbench database

Q: Better to move it to the 'Background' chapter?
The time series dataset used is taken from the Simbench database. This database refers to some real distribution networks in Germany in the year 2016. SimBench includes multiple time series for one year with 15 min resolution for load, generation and storage units. All time series came as active and reactive power. The time series were grouped by element type, reducing the total number of required time

series to a reasonable number, while retaining the possibility to model individual nominal power [17]. All active power values are normalized to the maximum active power value.

Power utilities commonly use generic load profiles to group commercial customers with similar load shapes into categories or standard load profiles (SLPs). The most commonly used profiles set is developed by the German Association of Energy and Water Industries (BDEW). It comprises eleven aggregated profiles, one for residential consumers (H0), three for agricultural (L0-L3), and seven for commercial consumers with different opening hours (G0-G6). They are differentiated into workdays, Saturdays and Sundays as well as three seasonal categories winter, summer, and transitional. The set also includes two profiles for street lightning (B0) and band load (G7).
The generation time series for photovoltaics (PVs), wind energy and biomass generated for the SimBench dataset are created using the agent-based simulation tool for optimized grid expansion planning SIMONA. SIMONA's power plant models receive real weather data of Germany from the German Weather Service in 2011 for Wind and 2012 for PV time series as input data.
For 2011 and 2012 generation data, the time axis is adjusted to 2016 by shifting days so that they correspond to the nearest weekday[18].

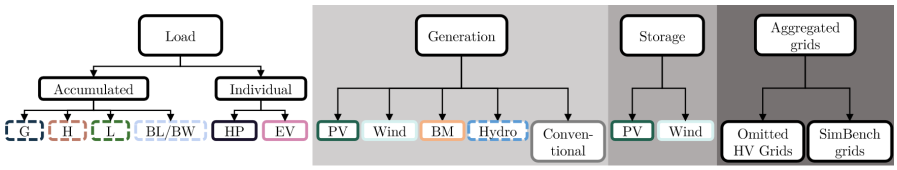

Figure 5.9: Overview of the SimBench time series type

The load time series were distinguished between real measured accumulated, highlighted with a dash, and simulated individual consumers, marked with a solid frame in Figure 5.9.

5.2.2 Time series

Some time series from the Simbench database are taken to adapt to the number of loads and DERs of the MV Oberrhein network in consideration.

As said, each element (load or generator) falls under a specific profile type that represent the consumption or generation over time.

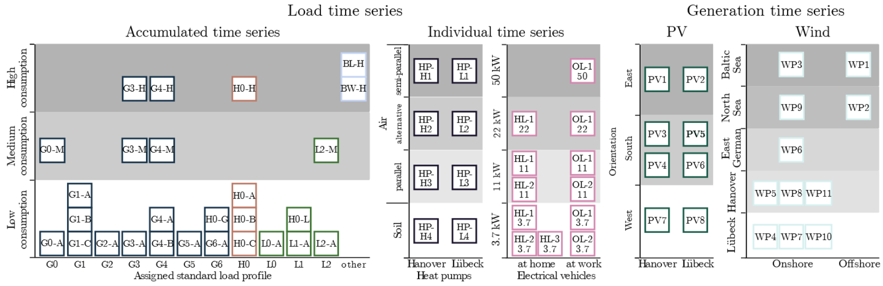

Figure 5.10: Load and generations profiles

In this case, the profiles' distribution in the Pandapower network is as follows:

Load elements by type: {'H0-A': 17, 'H0-B': 12, 'H0-C': 10, 'L2-A': 1, 'H0-G': 6, 'H0-L': 10, 'L2-M': 1, 'L1-A': 1, 'G2-A': 1, 'G6-A': 1, 'G1-B': 1}

RES elements by type: {'PV7': 8, 'PV8': 3, 'PV6': 7, 'PV5': 9, 'PV3': 16, 'WP4': 5, 'PV4': 8, 'PV1': 1, 'WP7': 3}

where the letters stand for: commercial enterprises (G), households (H), agricultural holdings (L) and industrial companies (BL/BW)'; with last letters –A to C indicating low consumption, -M medium consumption, and -H high consumption customers.
For the DES device there are photovoltaics PVs and wind parks WPs. It is possible to notice a bigger presence of PVs over WPs.

The loads and DERs are chosen so that different profile types are present.

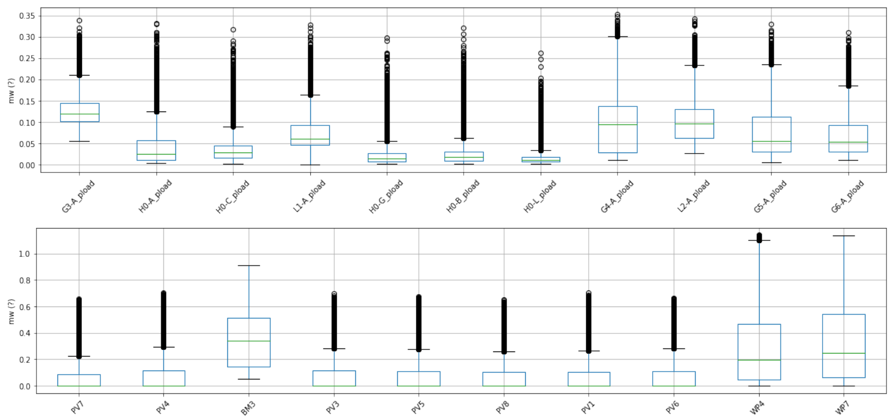

Figure 5.11: Box plots of load and generation for each profile type

Q: Is it ok that elements of the same type have the same profile, so same consumption/generation over time? R: add some noise
Q: Also from the documentation the values of the profiles are in MW, but it looks strange that the values never go above 1MW, is it common that facilities (small or big) consumes less than a 1MW? R: it depends on the device. Actually, the values are normalized, so a scaling factor can be used to change the values.

Since the profiles are equal for every element of that type, some noise is added to increase randomness. In particular, the noise added is a scaling factor in the range [0.85,1.15]. The scaling factor allows avoiding negative values in case of a value lower than 1; subtraction may result in a negative value of reactive power for a particular load.

With a similar approach, it is possible to easily create different cases, changing the scaling factors.

Table 3: Active power scaling factors of relevant study cases in SimBench

Acronym	Scenario description	Load	Generation		
			Wind	PV	Others
		p	p	p	p
hL	high load, low DER generation	1.00	0	0	0
n1	high load, low DER generation & contingency case	1.00	0	0	0
hW	high load, very high wind, high PV, high other DER generation	1.00	1.00	0.80	1.00
hPV	high load, high wind, very high PV, high other DER generation	1.00	0.85	0.95	1.00
ℓW	low load, very high wind, high PV, high other DER generation	0.25 (HV), 0.10 (MV/LV)	1.00	0.80	1.00
ℓPV	low load, high wind, very high PV, high other DER generation	0.25 (HV), 0.10 (MV/LV)	0.85	0.95	1.00

Figure 5.12: Possible cases. [ref]

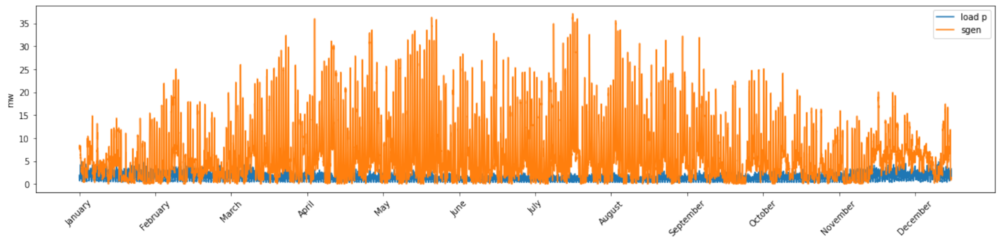

Figure 5.13: Sum of load energy consumption and energy generation over the considered year

[old] It is possible to notice a higher consumption of energy over the production, especially at the beginning and end of the year, when the sun intensity is not too high.

Case: High generation

Some cases are tested using scaling factors for load and generation. The high resolution case refers to scaling factors, as follows:

scale_factor_load = 0.6
scale_factor_sgen = 1

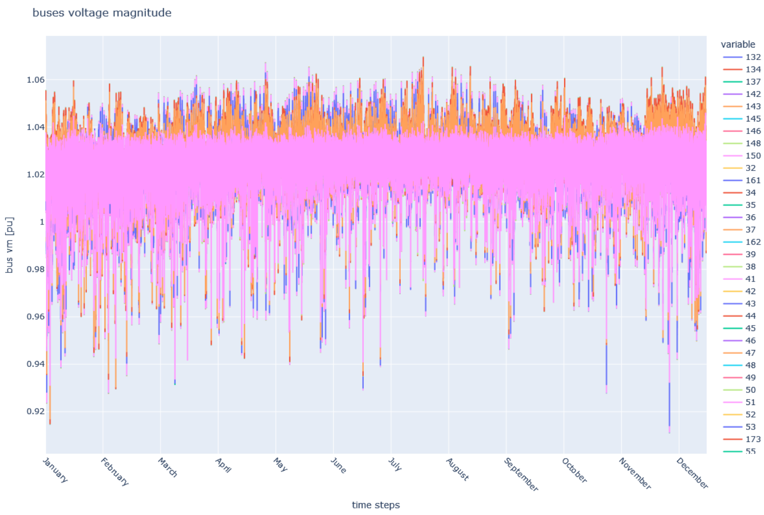

Figure 5.14: Case high generation. Voltage buses results obtained running the time series

From the plot above, it is possible to see that there are some overvoltage problems and few under voltage problems. A time step is critical when the voltage of any bus is out of the boundaries $V_i < v^{\min}$ (under voltage) or $V_i > v^{\max}$ (over voltage), where $v^{\min}$ is the minimum acceptable voltage, 0.95, and $v^{\max}$ is the maximum one, 1.05.
The problems are highlighted in the following plot.

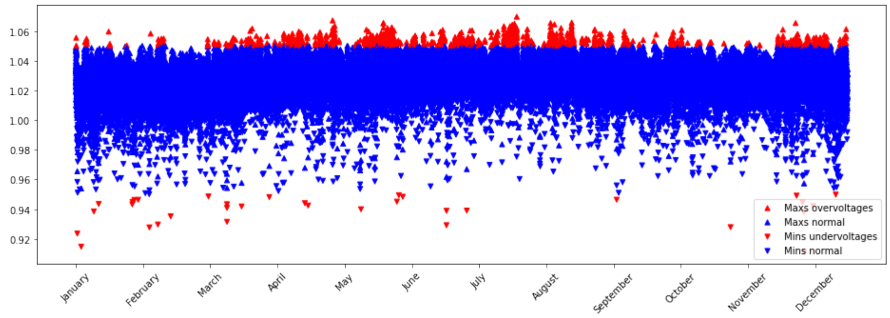

Figure 5.15: In this plot, the maximum and minimum values found for each time step and plotted, so for each time step there are two values representing the highest and lowest value among each bus. The colours, blue and red, represent the normal or critical condition

The full dataset is divided in train, validation and test set with the following proportions: 0.7, 0.2, 0.1 respectively.
The dataset is highly imbalanced, since usually networks perform well most of the time and the critical conditions are few:

Number of critical situations: 1161, over 35040 time steps, ratio: 3.3%

Number of critical instants in Train set: 907, ratio: 3.7%

Number of critical instants in Val set: 95, ratio: 1.4%

Number of critical instants in Test set: 159, ratio: 4.5%

The model tested is a multi layer perceptron with two hidden layers of size 192 and 64 and the output size is just one neuron with a Sigmoid activation function. The data input shape is as follows:

Inputs shape (batch size, time steps, features): (512, 16, 69)

with a temporal window r of 16 time steps (4 hours in the past). This input is flatten before being passed to the MLP.
The output shape is just:

Labels shape (batch size, time steps): (512, 1)

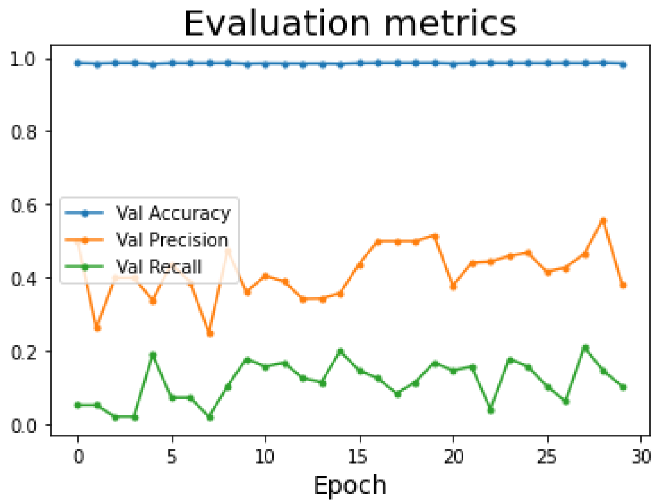

Figure 5.16: Training history on the validation set

As expected, the results are poor.
I also tried to add some class weights for the less present class, set bias weights in the last fully connected layer, applied differencing method (subtract time step value at time t with the value at time $t+1$; this removes the time dependency and stabilize the mean. Trend and seasonality are reduced in this way) but the results were still not acceptable.

Any suggestion? :'(
Moreover, what would be an acceptable level of recall/precision/accuracy? I was thinking that since the network can handle a critical situation for some time, if the DSO would curtail some generators, these values should be >0.8 so that it makes sense to trust the model and apply some action.

Case: High load

The high load case refers to scaling factors, as follows:

scale_factor_load = 1
scale_factor_sgen = 0.1

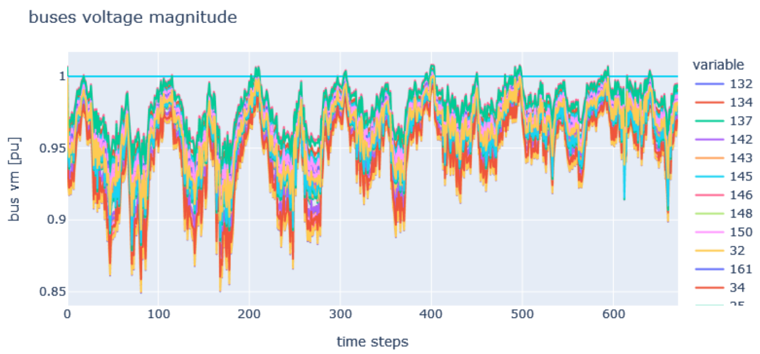

Figure 5.17: Case high load. Voltage buses results obtained running the time series

In this case, there are only under voltage problems.
The problems are highlighted in the following plot.

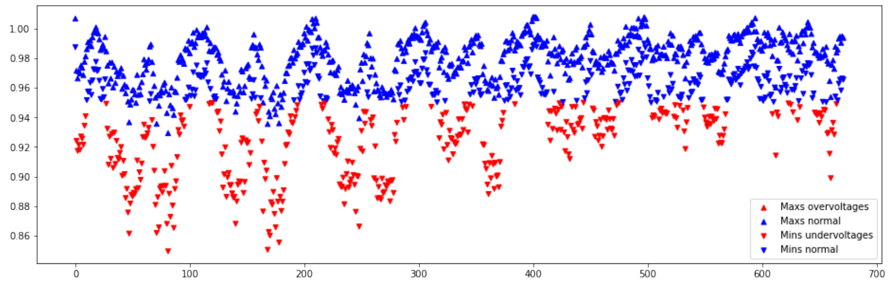

Figure 5.18: Same as above, the maximum and minimum values found for each time step and plotted, so for each time step there are two values representing the highest and lowest value among each bus. The colours, blue and red, represent the normal or problem condition.

Bibliography

[1] Hannah Ritchie and Max Roser. «CO and Greenhouse Gas Emissions». In: *Our World in Data* (2020). https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions (cit. on p. 3).

[2] United Nations Framework Convention on Climate Change (UNFCCC). *THE PARIS AGREEMENT*. 2015 (cit. on p. 3).

[3] Hannah Ritchie and Max Roser. «Energy». In: *Our World in Data* (2020). https://ourworldindata.org/energy (cit. on p. 3).

[4] Robin Henry and Damien Ernst. «Gym-ANM: Reinforcement learning environments for active network management tasks in electricity distribution systems». In: *Energy and AI* 5 (Sept. 2021), p. 100092. ISSN: 2666-5468. DOI: `10.1016/j.egyai.2021.100092`. URL: `http://dx.doi.org/10.1016/j.egyai.2021.100092` (cit. on pp. 5, 20, 23).

[5] Jianhong Wang, Wangkun Xu, Yunjie Gu, Wenbin Song, and Tim C. Green. *Multi-Agent Reinforcement Learning for Active Voltage Control on Power Distribution Networks*. 2022. arXiv: `2110.14300 [cs.LG]` (cit. on p. 5).

[6] Elia. «Our infrastructure». In: (2022). URL: `https://www.elia.be/en/infrastructure-and-projects/our-infrastructure` (cit. on p. 6).

[7] T.A. Short. *Electric Power Distribution Handbook, 2nd Edition*. 2004 (cit. on p. 6).

[8] E. Lakervi and E. J. Holmes. *Electricity distribution network design*. English. IEE power series. Contribution: organisation=svt,FACT1=1. Peter Peregrinus Ltd., 1995 (cit. on pp. 8, 12).

[9] MOHAN Ned. *Power electronics: a first course*. 2012 (cit. on pp. 11, 12).

[10] JD Kueck. «Measurement Practices for Reliability and Power Quality». In: (May 2005). DOI: `10.2172/885961`. URL: `https://www.osti.gov/biblio/885961` (cit. on p. 12).

[11] Leon Thurner, Alexander Scheidler, Florian Schafer, Jan-Hendrik Menke, Julian Dollichon, Friederike Meier, Steffen Meinecke, and Martin Braun. «Pandapower—An Open-Source Python Tool for Convenient Modeling, Analysis, and Optimization of Electric Power Systems». In: *IEEE Transactions on Power Systems* 33.6 (Nov. 2018), pp. 6510–6521. ISSN: 1558-0679. DOI: `10.1109/tpwrs.2018.2829021`. URL: `http://dx.doi.org/10.1109/TPWRS.2018.2829021` (cit. on pp. 13–15, 23).

[12] Fraunhofer IWES. «Convenient Power System Modelling and Analysis based on PYPOWER and pandas». In: (May 2019). URL: `https://pandapower.readthedocs.io/en/v1.3.0/_downloads/pandapower.pdf` (cit. on p. 16).

[13] «Timeseries Module Overview». In: (). URL: `https://pandapower.readthedocs.io/en/v2.2.2/timeseries/timeseries_loop.html` (cit. on p. 17).

[14] Quentin Gemine, Damien Ernst, and Bertrand Cornélusse. «Active Network Management for Electrical Distribution Systems: Problem Formulation and Benchmark». In: (May 2014) (cit. on p. 19).

[15] Quentin Gemine, Damien Ernst, and Bertrand Cornélusse. «Active network management for electrical distribution systems: problem formulation, benchmark, and approximate solution». English. In: *Optimization and Engineering* 18.3 (2016-10). ISSN: 1389-4420. DOI: `10.1007/s11081-016-9339-9`. URL: `http://dx.doi.org/10.1007/s11081-016-9339-9` (cit. on p. 20).

[16] A. Adamczyk, R. Teodorescu, R.N. Mukerjee, and P. Rodriguez. «Overview of FACTS devices for wind power plants directly connected to the transmission network». In: (2010), pp. 3742–3748. DOI: `10.1109/ISIE.2010.5637987` (cit. on p. 20).

[17] Steffen Meinecke, Džanan Sarajlić, Simon Ruben Drauz, Annika Klettke, Lars-Peter Lauven, Christian Rehtanz, Albert Moser, and Martin Braun. «SimBench—A Benchmark Dataset of Electric Power Systems to Compare Innovative Solutions Based on Power Flow Analysis». In: *Energies* 13.12 (2020). ISSN: 1996-1073. DOI: `10.3390/en13123290`. URL: `https://www.mdpi.com/1996-1073/13/12/3290` (cit. on p. 30).

[18] Christian Spalthoff, Džanan Sarajlić, Chris Kittl, Simon Drauz, Tanja Kneiske, Christian Rehtanz, and Martin Braun. «SimBench: Open source time series of power load, storage and generation for the simulation of electrical distribution grids». In: (May 2019) (cit. on p. 30).