ZADÁNÍ DIPLOMOVÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Sokolová** Jméno: **Nikola** Osobní číslo: **453422**

Fakulta/ústav: **Fakulta elektrotechnická**

Zadávající katedra/ústav: **Katedra počítačů**

Studijní program: **Otevřená informatika**

Specializace: **Datové vědy**

II. ÚDAJE K DIPLOMOVÉ PRÁCI

Název diplomové práce:

Aplikace strojového učení na energetických trzích

Název diplomové práce anglicky:

Machine learning in energy markets

Pokyny pro vypracování:

The efforts of the European Union (EU) in the energy supply domain sets targets for cutting greenhouse gas emissions and increasing the share of renewable energy. Hydropower is among the most efficient technologies to produce renewable electrical energy.
This project aims to propose predictive methods that take advantage of the available historical data to predict hydro power plant production. Accurate energy production forecasts may help to improve scheduling and operation of power systems.

1. Obtain relevant data.
2. Describe the underlying dynamics of the problem and support your statements with data.
3. Familiarize yourself with the domain and review the state of the art methods in energy production prediction.
4. Propose multiple predictive methods and correctly evaluate them. Your predictive methods must include confidence intervals.
5. Determine whether your proposed methods can be used to predict hydro power plant output.

Seznam doporučené literatury:

Jarábek, Tomáš, Peter Laurinec, and Mária Lucká. "Energy load forecast using S2S deep neural networks with k-Shape clustering." 2017 IEEE 14th International Scientific Conference on Informatics. IEEE, 2017.
Peter Laurinec, Mária Lucká, “Comparison of Representations of Time Series for Clustering Smart Meter Data”, roceedings of the World Congress on Engineering and Computer Science 2016 Vol I WCECS 2016, October 19-21, 2016, San Francisco, USA
Gabriela Grmanová, Peter Laurinec, Viera Rozinajová, Anna Bou Ezzeddine, Mária Lucká, Peter Lacko, Petra Vrablecová, Pavol Návrat, Incremental ensemble learning for electricity load forecasting, 2016 Acta Polytechnica Hungarica 13, 97 – 117

Jméno a pracoviště vedoucí(ho) diplomové práce:

Ing. Matej Uhrín katedra počítačů FEL

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) diplomové práce:

Datum zadání diplomové práce: **11.07.2022** Termín odevzdání diplomové práce: **13.02.2023**

Platnost zadání diplomové práce: **19.02.2024**

Ing. Matej Uhrín
podpis vedoucí(ho) práce

podpis vedoucí(ho) ústavu/katedry

prof. Mgr. Petr Páta, Ph.D.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Diplomantka bere na vědomí, že je povinna vypracovat diplomovou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací.
Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v diplomové práci.

.	
Datum převzetí zadání	Podpis studentky

Insert here your thesis' task.

Master’s thesis

Machine learning in energy markets

Bc. Nikola Sokolová

Department of Computer Science
Supervisor: Ing. Matej Uhrín

May 26, 2023

Acknowledgements

I would like to express my deepest gratitude to my family for providing me the opportunity to pursue my studies and supporting me throughout them. I am also very thankful to my supervisor for all the provided help and his patience with me. In addition, I would like to express my thanks to this school for giving me the opportunity to study here and write this thesis.

Declaration

I hereby declare that the presented thesis is my own work and that I have cited all sources of information in accordance with the Guideline for adhering to ethical principles when elaborating an academic final thesis.

I acknowledge that my thesis is subject to the rights and obligations stipulated by the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular that the Czech Technical University in Prague has the right to conclude a license agreement on the utilization of this thesis as school work under the provisions of Article 60(1) of the Act.

In Prague on May 26, 2023 . . . .

Czech Technical University in Prague
Faculty of Electrical Engineering

Citation of this thesis

Sokolová, Nikola. *Machine learning in energy markets* . Master's thesis. Czech Technical University in Prague, Faculty of Electrical Engineering, 2023.

Abstrakt

Cílem naší práce bylo předpovědět hodnotu produkce elektrické energie pro následující den vodní elektrárny na říčním toku. Analyzovali jsme možnosti pro jednokrokové předpovědi pro denní průměrná data a vícekrokové předpovědi pro hodinová průměrná data produkce elektrické energie, přičemž jsme jako prediktory použili hydrometeorologická data. Poté jsme prozkoumali vztah mezi daty o počasí a produkcí elektrické energie. Navrhli jsme několik přístupů, konkrétně Exponenciální vyhlazování, ARIMA a Temporální konvoluční síť. Zjistili jsme, že nejlépe fungující model bylo jednoduché Exponenciální vyhlazování bez jakýchkoliv exogenních proměnných, a že tyto exogenní proměnné neposkytovaly žádné další informace.

Klíčová slova Energy markets, prediction of hydropower plant's electricity production, temporal convolutional network, time series, renewable energy sources

Abstract

The goal of our work was to predict the day-ahead value of electricity production from a run-of-river hydropower plant. We analysed potential options for single-step predictions for daily average data and multi-step predictions for hourly average data of electricity production, using hydrometeorological data as predictors. Then, we investigated the relationship between the weather data and electricity production. We proposed several approaches, namely Exponential Smoothing, ARIMA, and Temporal Convolutional Network. We found that the best-performing model was simple Exponential Smoothing without any exogenous variables, and that these exogenous variables did not provide any additional information.

Keywords Energy markets, prediction of electric production of a hydropower plant, temporal convolutional network, time series, renewable energy sources

Contents

List of Figures

List of Tables

Introduction

The Earth's natural resources such as wind, sunlight, and water, play a very important role in achieving the European Union's energy goals and climate objectives. The amount of installed capacity from renewable energy sources, especially wind and solar power, has increased in the past few years, following the directive [46].

In 2020, the European Parliament adopted a resolution on the Green Deal, proposing actions against climate change. This implies the necessity of transitioning towards a decarbonized energy system in order to achieve an economy with net-zero greenhouse gas emissions. The Renewable Energy Directive requires the European Union to achieve a 32% share of renewable energy source by 2030.

Increasing share of such sources, has a bad influence on electricity markets and transmission grids. Power plants based on traditional sources can be planned and controlled. For instance, meeting the current energy demand when needed. This is not the case for power plants based on renewable resources. [10]

Energy production from renewable sources, especially from wind, solar, and run-of-river power plants [30, 26, 27], depends on external factors such as weather conditions, i.e wind speed, daylight, temperature. Hence, their output cannot be predicted and planned as accurately as that of more traditional energy sources that use fossil fuels. It means that volatile renewable energy sources can only generate power when their environmental conditions meet with their operational needs.

However, this condition does not apply to all run-of-river hydropower plants. Some create a water drop and ensure a relatively steady flow of water, with variations based on the water level 1. As less volatile renewable energy sources, they produce more power per installed capacity.

The power plant under analysis is aiming to offer its daily energy contracts for sale. In order to achieve this goal, it is necessary to determine the lowest (lower limit) and highest (upper limit) energy production for the next day.

In the thesis, we will analyze the predictability of the day-ahead output of a run-of-river hydropower plant based on meteorological conditions.

Thesis is divided as follows. In the first chapter 1, we discuss the mechanisms of energy production, concentrating specifically on the operations of hydropower plants. In the second chapter 2, we summarize existing approaches used in forecasting energy production from renewable sources and time series forecasting in general. The third chapter 3, is focused on analyzing the hydrological cycle in order to determine meteorological features used for the experimental part. Chapter 4 discusses the theoretical background. Finally, chapter 5 summarizes our findings, and presents methods for single-step (daily) and multi-step (hourly) day-ahead forecasting of energy production.

Chapter 1

Introduction to Energy production

This chapter aims to describe the basic principles of electricity production, with a focus on hydropower.

We will briefly mention energy sources as traditional fossil fuel-based thermal power and nuclear energy to renewable sources such as hydroelectricity, solar, wind, and geothermal power. Then we discuss the hydropower plants and analyse the one from which we collected our dataset.

1.1 Fundamentals of Electricity Generation

Electricity can be generated either chemically or mechanically [19]. Chemical methods involve the conversion of energy through reactions, such as in photovoltaic panels, while mechanical methods involve the rotary movement of a generator, driven by different forces like steam expansion or flowing water.

Power production can be accomplished through various methods, each more suited to specific conditions and circumstances and the output of a power plant may vary according to many factors such as local weather conditions, the availability of required resources, or electricity prices.

Existing solutions include thermal power based on fossil fuels such as coal, oil, and gas, as well as thermal power generated by nuclear fission and energy produced from natural resources, also known as renewable energy.

In the European Union, the term **'renewable energy'** refers to energy from renewable non-fossil sources. These include wind, solar, geothermal, and ambient energy, along with tidal, wave, and other ocean energies, hydropower, biomass, landfill gas, sewage treatment plant gas, and biogas. Solar energy can be divided into two additional categories: solar thermal and solar photovoltaic [46].

For instance, geothermal power plants harness steam emitted from the Earth to drive a steam turbine. Wind power plants use wind to spin copper wires within a generator, creating electricity. Hydroelectric dams use the force of falling or flowing water to rotate their generators. Finally, solar power utilize from converting sunlight directly into electricity. However, the application of such sources is limited to specific geographical regions.

In the following text, we describe the basic concepts in the domain of electricity production, such as capacity, energy, and demand load curve. For the following chapters, we need to understand the difference between the terms **capacity** and **energy**.

1.1.1 Capacity

Capacity represents the electricity produced or consumed instantaneously [19], measured in kilowatts (kW). In other words, it is maximum output a power plant can produce. Power plants often do not operate at their full capacity all the time. In other words, they have a capacity to produce a certain amount of power during a given time period but other factors such as maintenance or refueling can take power plants offline.

1.1.2 Energy

Energy is the amount of electricity that is actually generated or consumed over time, measured in watt-hours (kWh) [19]. To distinguish between the concepts of capacity and energy, we will demonstrate two examples.

- A typical small hydropower plant has a capacity of between $1,000$ kW and $10,000$ kW.
- The total amount of electricity generated in the Czech Republic was 84.9 TWh.

1.1.3 Demand Load Curve

The demand load curve represents the total electricity demand at a given time by various consumers, such as residential, commercial, and industrial [14]. This demand fluctuates throughout the day, an it is affected by the time of day, weather conditions, and seasonal variations.

The load curve defines the electricity demand's different aspects: **base load** (load needed all year), **peak load** (load needed only a few hours a day), and **intermediate load** for operating hours between base load and peak load, see Figure 1.1.3.

For instance, consider a power plant with a capacity of $50,000$ MW. Initially, the regional power system seemed well-prepared to handle the estimated summer peak demand of $35,000\ MW$. However, a change in weather conditions affected the plant's actual performance, reducing energy output. Then, there wasn't enough energy produced to meet the demand.

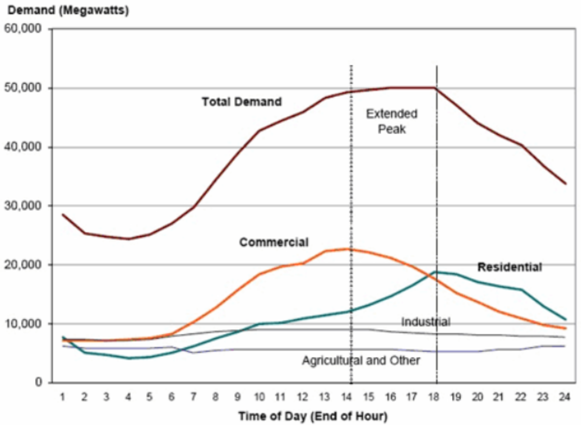

Figure 1.1: Example of electricity demand load across agricultural, commercial, and other sectors during the day. [33]

1.2 Hydropower plants

In this section, the principles of operation of hydropower plants are discussed. We then describe how distinct power plants work and finally, we will analyze the power plant for which our goal is to predict future output.

Hydropower is a renewable source of energy that harnesses the hydrological cycle 3. Hydropower plants range in size from small installations with a power output of a few kilowatts, to large dam-based power plants with a capacity of thousands of megawatts.

1.2.1 Basic principle

Hydropower plants convert the kinetic and potential energy of water into electricity. Water flowing towards the plant transfers its kinetic and potential energy to a turbine, which spins a generator. The rotational energy in the generator is converted into electrical energy through electromagnetic induction. The power output of the turbine depends on the size of the water fall, the water flow through the turbine, and the efficiency of the turbine.

The operational nature of hydro storage power plants differs significantly. Some produce baseload power and have comparably high capacity factors. Others are peak-load power stations with much lower capacity factors and operate only in times of high demand or high prices. The size of reservoir, the water flow into the reservoir and the turbine capacity are factors that determine how a hydro storage power plant operates.

1.2.2 Types of Hydropower Plants

The section focuses on the types of hydropower plants, specifically run-of-river, hydro storage, and pumped storage plants. Primary interest of this thesis is a small run-of-river plant, and the other two types are just briefly described.

1.2.2.1 Run-of-river

Run-of-river hydropower plants are a type of power station that uses the flow of a river to generate electricity. Their capacities range from just few kW to hundreds of MW. Unlike other types of hydropower plants, they have no water storage or very limited one, known as pondage. Water is accumulated in the pondage during periods of low demand, and can then be used for electricity production when demand is high. Their limitations lie in the absence of a large water reservoir, which affects their scalability, flexibility and make them dependent on river flows.

For instance, run-of-river power plants situated near mountain rivers have larger electricity production in the spring and summer months thanks to increased water flow by melting snow. [19] Additionally, in South Asia, the river flow increases during the rainy season. [2]

Hence, the amount of electricity a run-of-river hydropower plant can generate is determined by the following factors.

- The volume of the water flow in the river
- The change in elevation, often referred to as head

That means, the greater the water flow and the higher the head, the more electricity a hydropower plant can produce. This is due to the fact that a larger potential energy can be converted into electricity.

1.2.2.2 Hydro storage

Hydro storage power plants use a dam to store water in a reservoir. [19] They are typically situated in mountainous regions due to the specific geographical and geological conditions they require. However, this does not necessarily imply that these are the most optimal locations. Several factors must be taken into account including the potential output and the water rights. Such locations are chosen for the sharp drop in river elevation levels and the topographical suitability for a reservoir. The construction of these reservoirs often leads to large expanses of what used to be dry land becoming submerged under water.

An important distinction between run-of-river plants and hydro storage is the control of water flow. In hydro storage plants, operators have the ability to determine the quantity and timing of electricity production. This feature is significant for our analysis.

1.2.2.3 Pumped storage

Pumped storage power plants serve as energy storage for other sources and cover peak load demand. They utilize two differently elevated water reservoirs and store energy in the form of potential energy of water. Surplus electrical energy, primarily from high renewable energy production, is used to pump water into the higher reservoir. Conversely, when energy is needed, water flows through the turbine, and the generator supplies electricity to the grid [16].

1.2.3 Small hydro power plant in Želiezovce

Our dataset, referred to in 5.2, was collected from a small run-of-river hydropower plant in Želiezovce, geographically situated in the southern part of Slovakia, in the lower Hron region. The location provides suitable conditions for hydropower plant.

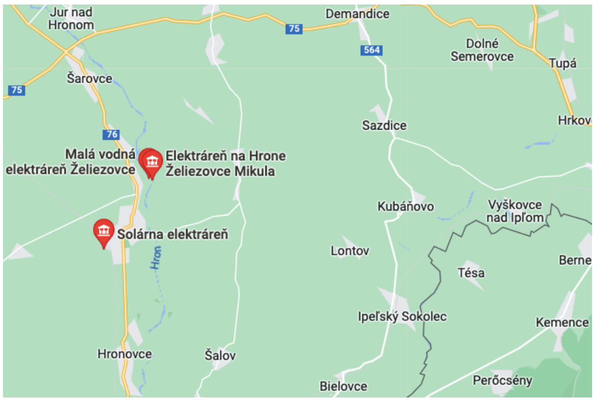

Figure 1.2: Location of the small power plant in Želiezovce [35]

The Power plant (see Figure 1.2.3) has maximum capacity is 2, 8 MW and expected annual production is 13, 5 GWh.

It processes the flows of the Hron river, which, after converting kinetic energy to electricity, flow back into the river. The necessary water drop did not exceed the levels of major water flow, thereby not creating a reservoir. Another factor affecting the power plant output is the restriction of the maximum amount of water it can process, 62.0 m^3s^-1, due to fish migration from April to June.

1.2.3.1 Summary

A potential problem for our prediction is the fact that the power plant can react quickly to changes in the power system. This means that the power

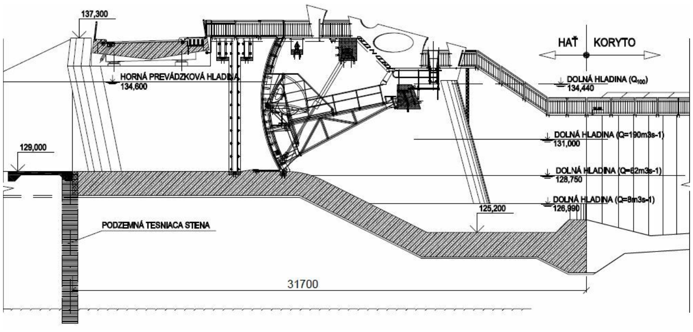

Figure 1.3: Key features include an embankment dam, a movable weir, a bio-corridor, the hydroelectric power plant itself, a bridge, a canoe slide, and flood embankments. An integral part of the construction was the implementation of a pumping station for an irrigation system that supplies water from the Hron River to the surrounding fields. [36].

plant's output can be controllable in some way, and fish migration can also introduce variability.

Aspects that we did not analyze include the pumping station, which can also decrease the amount of water going to the power plant, and the maximum volume of water that the power plant can process.

Chapter 2

Hydropower Energy Forecasting - Existing Approaches

In this chapter, we will discuss various existing approaches for time series analysis and electricity production forecasting.

The Autoregressive Integrated Moving Average Model (ARIMA), which has been in use for time series analysis since the 1970s, was introduced by Box and Jenkins [5]. ARIMA is employed for predicting stationary data. The performance of this model can be improved by applying various preprocessing techniques to time series data.

Zhang et al.'s study [45] revealed that feedforward neural networks struggle to effectively handle seasonality or trends in unprocessed raw data. According to their research, preprocessing techniques such as detrending or deseasonalization can substantially reduce forecasting errors. These neural networks yield more robust forecasting performances compared to ARIMA [43]. Several studies [39, 23] have found feedforward neural networks to be effective for inflow forecasting. For instance, Kicsi [29] performed short-term daily streamflow forecasting experiments using various ANN models.

Certain studies utilizing feedforward neural networks incorporated precipitation and temperature forecasts as feature variables [40]. These studies show that forecast performance can be enhanced using recurrent neural networks. For example, Yongsheng et al. [44] used RNN for short-term renewable energy prediction, and Busseti et al. [7] employed RNN for energy load forecasting. They tested several methods, including Kernelized Regression, Frequency NN, Deep Feedforward NN, and Deep Recurrent NN. The Deep Recurrent NN model demonstrated the best results.

A promising alternative is the Support Vector Machines (SVM). Case studies have indicated that SVM can outperform methods like feedforward neural networks [28, 41]. The SVM model has demonstrated proficient capability for short-term forecasting and predicting peaks in a time series process, outperforming neural network-based models [11].

Chapter **3**

Weather elements and Riverflow

The following chapter of the thesis is dedicated to discussing the weather elements that will be used later in our analysis. The basic characteristics of the collected hydrometeorological data can be found in Chapter 5.2.

We investigate the Earth's hydrological cycle to better understand how hydrometeorological data can affect the river flow and in the end the performance of a run-of-river hydropower plant. [32, 15]

Our assumption is that two weather elements should have the most influence on streamflow: precipitation and temperature [38, 13, 18].

For instance, when temperatures drop below 0 $°C$, water in the stream may freeze, and any snowfall does not melt, thus reducing the water supply to the stream. On the other hand, precipitation can increase the volume of water in the stream, which can lead to an increase in power production.

In addition to precipation and temperature, we have chosen other parameters that are closely related to them.

3.1 The Earth's hydrological cycle

The following text provides a brief overview of the hydrologic (water) cycle, as depicted in Figure 3.1. Based on the processes outlined in the hydrologic cycle, we selected specific hydrometeorological data for our analysis.

3.1.1 Precipitation

Hydropower generation is sensitive to changes in river flow, which are influenced by accumulated precipitation measured in mm. There is a strong relationship between hydropower station generation and precipitation data, according to many researches from our analys of existing approaches 2. The lowest flows typically occur during the winter months due to snow storage, while low flow events in the summer result from a deficit in precipitation

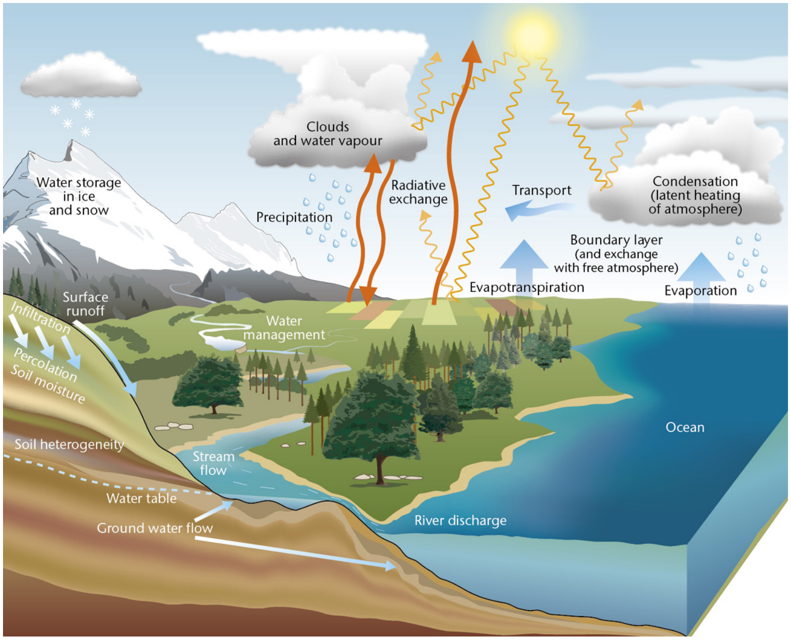

Figure 3.1: Diagram of the hydrological cycle [37].

and high evaporation. Liquid precipitation is less common during periods of snowfall.

3.1.2 Air Temperature

Air temperature affects the physical state of water in the environment. For instance, when temperatures drop below freezing, the water in the river may turn into ice. This results in a decrease in the flow of liquid water in the river. Snowfall that occurs during these low-temperature periods does not readily melt, instead accumulating and later providing an additional source of water to the river once temperatures rise.

High temperatures lead to liquid precipitation, which can flow directly into the river.

Furthermore, studies such as the one cited in [24] show that air temperature accelerates evaporation from the water surface. Consequently, we can expect a lower water level in the river during hotter periods due to increased rate of evaporation.

3.1.3 Atmospheric preassure

The rising of air in low-pressure leads to its cooling and condensation into clouds and precipitation. Therefore, it is reasonable to expect a dependency between atmospheric pressure and precipitation. A sudden drop in atmospheric pressure usually signifies a storm. If atmospheric pressure remains steady, it suggests that weather changes are not very probable. It is measured in hPa.

3.1.4 Evapotranspiration

Evapotranspiration, measured in mm, represents the combined process of evaporation and plant transpiration to the atmosphere, converting liquid water into water vapor.

During warmer periods, a higher rate of evapotranspiration could potentially lead to lower water levels in the river due to increased water loss to the atmosphere. This could impact the volume of water available for power generation.

3.1.5 Soil moisture index

The soil moisture index indicates the wetness of the soil, i.e., the amount of water present. Index is interpreted as the ratio between the volume of water the soil is currently holding and the maximum volume it can hold.

Typically, the index ranges between values of $[0, 1]$. However, following excessive precipitation, the index can exceed 1. This indicates that the ground is saturated and cannot hold more water, potentially leading to water runoff across the surface. For our purposes, this suggests a high water level in the river and we can, therefore, anticipate electricity production to be close to maximum output. In such cases, we expect a correlation between accumulated precipitation and the soil moisture index.

The soil moisture index is largely dependent on the type of soil, as different soils have distinct properties and can retain varying amounts of water. Top layers tend to be drier due to the evaporation process, while lower layers often contain more water.

We have chosen to gather data specific to the top layers of the soil. Although we already possess information about precipitation up to the current time, it could be beneficial to have an additional indicator to estimate the potential volume of water in the river.

3.1.6 Soil water content

Soil water content provides additional information to the soil moisture index. Soil water content refers to the volume of water contained within different soil layers, measured in m^3m^3, and is therefore dependent on the same factors as the soil moisture index, as well as on the level of the groundwater.

During periods of low precipitation, groundwater can be a consistent supply of water for rivers, as the soil releases water into them.

Soil type data was not collected because all the measurement stations are located in close proximity to the location of hydro-power plant. Additionally, in Slovakia, soil types tend to be quite consistent within specific regions, in this case it is mostly brown earth or chernozem, rendering the collection of this particular data unnecessary for our purposes. [25]

Soil water content data was collected with the expectation that it could provide insights about the water level in the river, especially during hotter months.

3.1.7 Snow depth

The snow depth may not have an immediate effect on river flow. However, it does indicate the potential for future water level increases in the river once the snow starts melting due to higher temperatures. It is important to note that the effects of increased snow depth on river flow can be delayed as well as for other mentioned.

Chapter **4**

Theoretical Background

4.1 Time Series Analysis

In this introductory chapter, we will introduce the fundamentals of time series analysis and outline the fundamental techniques for understanding, modeling, and forecasting data.

4.2 Time Series

We can think of a time series as a list of measured values, accompanied by information about when each value was recorded. They simply represent data points over time. Let us introduce the formal definition of a time series.

Let $(\Omega, \mathcal{F}, P)$ is probability space. The time series is set $\{Y_t, t \in T\}$, that $Y_t \in (\Omega, \mathcal{F}, P)$ and T is a set of time indices. If $t \in \mathbb{Z}$, we are talking about a discrete time time series. Otherwise, if $t \in \mathbb{R}$, then a time series is said to be continuous when observations are made continuously in time [8]. In this thesis, we consider the case of discrete time for our analysis.

Examples of time series can be found across numerous domains, from finances to engineering. We will introduce some of them.

4.2.1 Time series patterns

Traditional methods of time-series analysis are mainly concerned with decomposing a series into several patterns: a trend (T), a seasonal pattern (S), and 'irregular' fluctuations (E). The cyclical pattern (C) is usually combined with the trend, creating a trend-cycle component [8]. For ease of understanding in the upcoming chapters of this thesis, the trend-cycle component will be referred to simply as the 'trend'.

Time series typically include some of these patterns. To choose the most suitable forecasting model, it is important to identify these time series patterns in the dataset for futher analysis.

4.2.1.1 Trend

The trend is a long-term increase or decrease in the data [22]. Simply put, it is the long-term direction of the time series. The trend is not always linear, and a change in direction can occur when it shifts from an increasing trend to a decreasing one. For instance, in the context of company sales, a trend that was previously showing an increase might reverse its course, possibly due to factors such as increased competition or changes in market conditions.

4.2.1.2 Seasonal pattern

The seasonal pattern of a time series is a pattern that repeats with a known periodicity. This could manifest as a weekly pattern, repeating every 7 days, or a monthly pattern repeating once a month. It can also display annual patterns, repeating once a year. For instance, sales often increase before Christmas Eve, demonstrating an annual seasonal trend. Seasonality can be identified by regularly spaced peaks or periods of flatness in the data, with the same magnitude recurring over the specified period.

4.2.1.3 Cycle

In a timeseries, cycle is a pattern that repeats with a certain regularity, but with unknown and changing periodicity. In the field of economics, a prime example of this pattern is the business cycle. The main difference between seasonal and cyclical behavior lies in the frequency of fluctuations. If the fluctuations do not occur at a fixed frequency, then they are considered cyclical rather than seasonal. A seasonal pattern is characterized by a constant frequency that corresponds with specific calendar-based intervals or events [22].

4.2.1.4 Other irregular fluctuations

The irregular fluctuations represent the unpredictable pattern of the series.

4.2.2 Time series decomposition

The starting point for most time series decompositions is often a classical decomposition method. However, due to several problems associated with this method, it is not typically recommended. We will talk about this method first so that it is easier to understand a more robust method, known as Seasonal and Trend decomposition using Loess (STL), which will be used later in our analysis.

4.2.2.1 Additive and Multiplicative models

Let us assume y_t is the observed variable at time t, T_t is a trend value, S_t seasonal component and E_t is unpredictable component. Time series components can be then combined into two following models.

- **Additive model**

$$y_t = T_t + S_t + E_t \tag{4.1}$$

- **Multiplicative model**

$$y_t = T_t \times S_t \times E_t \tag{4.2}$$

In order to continue with the text, it is necessary to define several terms beforehand.

4.2.2.2 Seasonally adjusted data

The seasonal component is removed from the data. Then, seasonally adjusted data, contains the trend and the remainder component, reffered to as unpredictable component.

- **Additive model**

$$y_t - S_t \tag{4.3}$$

- **Multiplicative model**

$$\frac{y_t}{S_t} \tag{4.4}$$

4.2.2.3 Mooving average

Moving avereage is used as a first step in a classical decomposition method. [22]

Let us assume order of moving avere to be m, where $m = 2k + 1$. Then moving average of order m can be computed as follows.

$$\hat{T_t} = \frac{1}{m} \sum_{j=-k}^{k} y_{t+j} \tag{4.5}$$

It is an estimation of trend-cycle at a specific point in time t. Observations that occur close together in time tend to have similar values, this averaging process helps in smoothing out some of the random fluctuations in the data, leaving behind a clearer view of the trend-cycle component.

4.2.3 Additive decompostion

In additive decomposition, the moving average is calculated first, then the detrended series i.e $y_t - \hat{T}_t$. Next step is averaging the detrended values based on the specified period. That means for daily data with weekly pattern, all weeks are averaging. These seasonal component values are then adjusted to ensure that they add to zero. The seasonal component is obtained by stringing together these monthly values, and then replicating the sequence for each year of data. After such computation we obtain the seasonal component $\hat{S}_t$. And finally, we will obtain the remainer component by substracting seasonal and trend-cycle component from previous steps $\hat{R}_t = y_t - \hat{T}_t - \hat{S}_t$.

4.2.4 Multiplicative decomposition

A classical multiplicative decomposition operates in a similar way, but it uses division instead of subtraction.

4.2.4.1 STL decomposition

We will briefly mention the Seasonal and Trend decomposition using Loess, shortly STL, is a robost method which can handle any type of seasonality in the data. Loess is a method for estimating non-linear relationships. [9]

4.3 White noise

A time series that has a mean equal to 0, a variance that is constant over time, and we expect each autocorrelation to be close to 0 (as there is some random variation in time series), is referred to as white noise. [22]

4.4 Stacionarity

Time series with trends or seasonality are not stationary because they depend on the time at which they are observed. Trend or seasonality will affect the value at differents points in time. Time series $\{Y_t, t \in T\}$ is stationary, then for all s, the distribution of $(y_t, ..., y_{t+s})$ does not depend on t [22].

White noise is a stationary process. Time series without trend and seasonality patterns, but with cycle is stationary. Cycles do not have fixed lenght and we are not sure where the peaks of the cycles are.

Stationary time series do not have predictable patterns in the long term.

4.5 Pearson correlation coefficient

Let us assume two random variables X, Y positive variances σ_x^2 and σ_y^2 and covariance $cov(X, Y)$. Mathematically, the Pearson correlation coefficient is defined as follows [4].

$$\rho_{XY} = \frac{\mathbb{E}[(X - \mu x)(Y - \mu y)]}{\sigma_x \sigma_y} = \frac{cov(X,Y)}{\sigma_x \sigma y}, \quad \rho \in [-1,1] \tag{4.6}$$

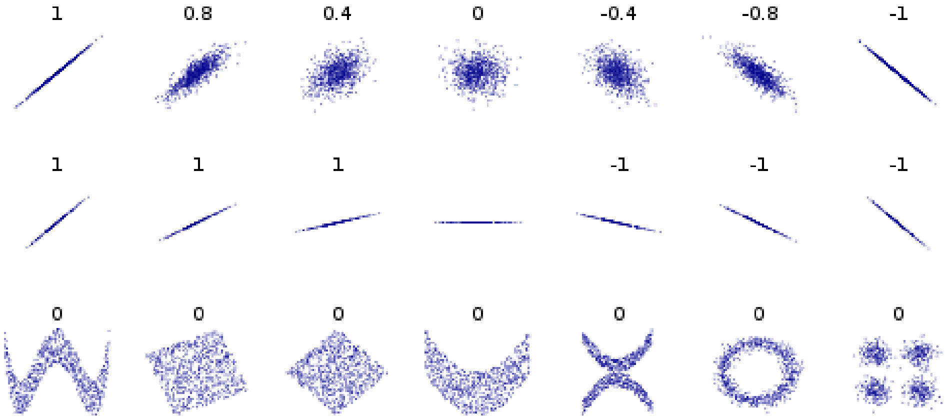

Figure 4.1: Visualised plots illustrate the correlation between different pairs of variables X and Y [12].

The Pearson correlation coefficient is a measure of the linear dependence between two random variables X and Y. Specifically:

- Value $\rho_{XY} = 0$ means that there is no linear relationship between variables X and Y.
- Value $\rho_{XY} = \pm 1$ means that variables are lineary dependent.
- Value $\rho_{XY} \in (-1,1)$ indicates the degree of linear dependence between the variables.

Given n samples $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, the sample correlation coefficient can be computed as follows.

$$r = r_{xy} = \frac{1}{n-1} \sum_{i=1}^{n} (\frac{x_i - \overline{x}}{s_x})(\frac{y_i - \overline{y}}{s_y}), \tag{4.7}$$

where $\overline{x}$, $\overline{y}$ are sample means, and s_x, s_y are sample deviations.

4.6 Autocorrelation

Autocorrelation is one of the key features of the time series and it measures the linear relationship between lagged values.

Let $\{Y_t, t \in T\}$ be a time series, μ_t is a mean value and $\sigma_t^2 > 0$ is a variance for every t. Autocorrelation coeficient for times s and t is defined as follows.

$$R(s,t) = \frac{\mathbb{E}[(X_t - \mu_t)(X_s - \mu_s)]}{\sigma_t \sigma_s}, \quad R(s,t) \in [-1,1] \tag{4.8}$$

Given n samples $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, the sample autocorrelation coefficient $\widehat{R}(k)$ can be computed as follows.

$$\widehat{R}(k) = \frac{1}{(n-k)} \sum_{t=1}^{n-k} \frac{(x_t - \mu)(x_{t+k} - \mu)}{\sigma^2}. \tag{4.9}$$

The autocorrelation can be used to identify non-stationary time series. For a stationary time series, the ACF will drop to zero relatively quickly, while the ACF of non-stationary data decreases slowly.

4.7 Differencing

The approach for making non-stationary data stationary is called differencing. Differenced time series represents the change between consecutive observations. We can obtain differenced series y'_t as follows.

$$y'_t = y_t - y_{t-1} \tag{4.10}$$

Resulted timeseries will have $T - 1$ values. Using differencing we stabile the mean value by reducing trend and seasonality components.

4.7.1 ARIMA models

ARIMA, which stands for AutoRegressive Integrated Moving Average, is a combination of two main components: the autoregressive (AR) aspect and the moving average (MA) parts. It's a statistical model commonly used for analyzing and forecasting time series data. Together with Exponential Smoothing they are the most widely known approaches for forecasting data.

4.7.2 AR(p) - Autoregressive model

Autoregresive models forecast the the prediction based on linear combination of past values. That means, lagged values of the y_t are predictors. The parameter q is the order of moving average model.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \tag{4.11}$$

Order of the AR model is p, ε_t is a white noise and $\phi = (\phi_0, ..., \phi_p)$ are regression coeficients.

4.7.3 MA(q) - Moving average model

Mooving average model uses past forecast errors in a regression model.

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \tag{4.12}$$

where ε_t is a white noise.

4.7.4 ARIMA(p,d,q)

Based on models defined from previous sections we can formulate ARIMA model.

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t, \tag{4.13}$$

where y'_t is differenced series. The parameter d is a degree of differencing.

4.7.5 SARIMA

The SARIMA is a seasonal ARIMA model, including seasonal term. It can be written as follows.

$$ARIMA(p, d, q)(P, D, Q)_m, \tag{4.14}$$

where m is the seasonal period, first part is defined ARIMA model and second part represents the seasonal part of the model.

SARIMA model is defined as follows.

$$(1 - \phi_1 B)\left(1 - \Phi_1 B^4\right)(1 - B)\left(1 - B^4\right) y_t = (1 + \theta_1 B)\left(1 + \Theta_1 B^4\right) \varepsilon_t \tag{4.15}$$

4.7.6 SARIMAX

The SARIMAX is an extension of defined SARIMA model and allows to include the exogenous regressors into the model.

4.8 Exponential smoothing

In the following text we will define the Exponential smoothing methods.

4.8.1 Simple exponential smoothing

Simple exponential smoothing - SES, often reffered to as a Holt Linear method is a method, which can be used when there is no clear trend or seasonal pattern. All forecasts for the future are equal to the last observed value of the series.

Let us asume time series $\{Y_t, t \in T\}$. Forecasts are then computed using weighted averages. The weights exponentially decrease as they come from further in the past.

Component form is defined as follows.

$$Forecast : \widehat{y}_{t+h|t} = l_t, \tag{4.16}$$

$$Smoothing : l_t = \alpha y_t + (1 - \alpha) l_{t-1}, \tag{4.17}$$

where l_t is the level and α is the smoothing parameter and $0 \leq \alpha \leq .1$

4.8.2 Holt's linear trend method

Simple exponential smoothing is suitable when there is no trend or seasonality. To accommodate these elements, an extension to simple exponential smoothing was introduced, known as Holt's linear trend method [20].

$$Forecast : \widehat{y}_{t|h} = l_t + hb_t, \tag{4.18}$$

$$Level : l_t = \alpha y_t + (1 - \alpha)(l_{t-1} + b_{t-1}), \tag{4.19}$$

$$Trend : b_t = \beta(y_t - l_{t-1}) + (1 - \beta) b_{t-1}. \tag{4.20}$$

4.8.3 Holt-Winters's method

Method has two variants, additive and multiplicative. Let us first introduce component form for the additive method.

$$Forecast : \widehat{y}_{t|h} = l_t + hb_t + s_{t+h-m}(k+1), \tag{4.21}$$

$$Level : l_t = \alpha(y_t - s_{t-m}) + (1 - \alpha)(l_{t-1} + b_{t-1}), \tag{4.22}$$

$$Trend : b_t = \beta(y_t - l_{t-1}) + (1 - \beta) b_{t-1}, \tag{4.23}$$

$$Seasonal : s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma) s_{t-m}. \tag{4.24}$$

Multiplicative model is computed as follows.

$$Forecast : \widehat{y}_{t|h} = (l_t + hb_t) s_{t+h-m}(k+1), \tag{4.25}$$

$$Level : l_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1}), \tag{4.26}$$

$$Trend : b_t = \beta(l_t - l_{t-1}) + (1 - \beta) b_{t-1}, \tag{4.27}$$

$$Seasonal : s_t = \gamma \frac{y_t}{l_{t-1} - b_{t-1}} + (1 - \gamma) s_{t-m}, \tag{4.28}$$

4.9 TCN - Temporal Convolutional Networks

Convolutional Neural Networks (CNNs), are typically used in classification tasks. However, demonstrated that they can be used to sequence modeling and forecasting when suitably modified. In this section, we will describe how Temporal Convolutional network works and the network architecture which will be used for further analysis. In the study [3] was demonstrated that convolutional networks offer several advantages over recurrent models in various tasks. Additionally, convolutional networks enable parallel computation of outputs, which can result in performance enhancements.

4.9.1 Network Architecture

The TCN network 4.9.1 consists of dilated and 1D convolutional layers and can take a sequence of any length and map it to an output sequence of the same length, similar to Recurrent Neural Networks.

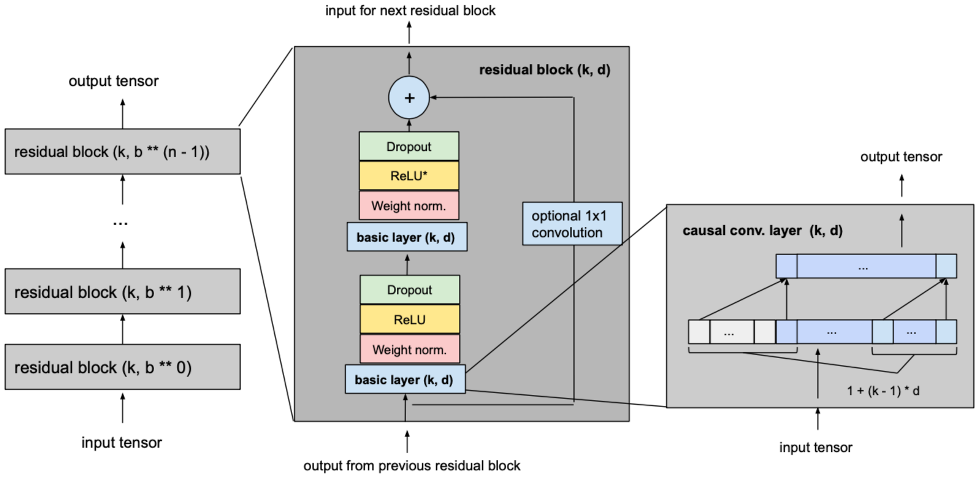

Figure 4.2: The following architecture visualise architecture of the TCN model. [34].

4.9.1.1 Sequence modeling

Let us define the sequence modeling task. Asumme an input sequence $x_0, \ldots, x_T$, and corresponding outputs $y_0, \ldots, y_T$ at each point of time. Formally, a sequence modeling network is any function $f : \mathcal{X}^{T+1} \rightarrow \mathcal{Y}^{T+1}$ that produces the following mapping:

$$\hat{y}_0, \ldots, \hat{y}_T = f\left(x_0, \ldots, x_T\right), \tag{4.29}$$

if the causal constraint that y_t depends only on $x_0, \ldots, x_t$ and not on inputs $x_{t+1}, \ldots, x_T$.

The goal is to find a network f that minimizes expected loss between the actual outputs and the predictions.

$$L\left(y_0, \ldots, y_T, f\left(x_0, \ldots, x_T\right)\right), \tag{4.30}$$

where the sequences and outputs are drawn according to some distribution.

4.9.1.2 Causal Convolutions

The TCN model is designed based on following principles. First, it ensures that the network generates an output sequence of the same length as the input sequence. Second, it prevents any information leakage from the future to the past.

The TCN architecture employs a 1D fully-convolutional network, where each hidden layer has the same length as the input layer. Zero padding of length (kernel size - 1) is applied to maintain consistent layer lengths throughout subsequent layers. To satisfy the second principle, the TCN utilizes causal convolutions, where each output at time ”t” is convolved exclusively with elements from time ”t” or earlier in the preceding layer.

4.9.1.3 Dilated layer

A simple causal convolution can only consider a history of limited size, which grows linearly with the depth of the network. To overcome this limitation, dilated convolutions are incorporated to enable an exponentially larger receptive field.

Formally, for a 1D sequence input $\mathbf{x}$ of length n and a filter $f : 0, \ldots, k-1 \rightarrow \mathbb{R}$, the dilated convolution operation F at position s in the sequence is defined as follows:

$$F(s) = \left(\mathbf{x} * df\right)(s) = \sum i = 0^{k-1} f(i) \cdot \mathbf{x}_{s-d \cdot i} \tag{4.31}$$

Here, d represents the dilation factor, k denotes the filter size, and $s - d \cdot i$ accounts for the past direction. Dilation introduces a fixed step between adjacent filter taps. When $d = 1$, a dilated convolution reduces to a regular convolution. However, using larger dilation factors allows the output at each level to capture a wider range of inputs, effectively expanding the receptive field of the convolutional network [31].

4.9.1.4 Residual connection

A residual block contains a branch leading out to a series of transformations $\mathcal{F}$, whose outputs are added to the input $\mathbf{x}$ of the block:

$$o = \text{Activation}(\mathbf{x} + \mathcal{F}(\mathbf{x})) \tag{4.32}$$

Since a TCN's receptive field depends on the network depth n as well as filter size k and dilation factor d, stabilization of deeper and larger TCNs becomes important.

Chapter 5

Experiments

The chapter begins with task formulation, followed by a discussion on the basic characteristics of the collected dataset, i.e electricity production and weather data. Afterward, we will describe how the dataset was preprocessed. Finally, we will present our experiments and next day forecasts using different approaches.

5.1 Task formulation

The task is to forecast the load production for a next day. Mathematically, the problem can be defined as follows.

5.1.1 Single-step prediction

The task for single step prediction is to forecast the load production at time $t+1$ as $\hat{y}(t+1)$ given the past observations.

5.1.2 Multi-step prediction

The definition is as follows. Assume the load production given the previous values and the forecasted load production at time t as $\hat{y}(t)$. Here, t is measured in hours, so $t \in \{0, 1, 2, ..., 23\}$ for a next day forecast using hourly data.

We will use the load production from previous times and possibly other variables, meteorological data, to forecast the load production 24 hours into the future.

The goal is to minimize the difference $D(y_t, \hat{y_t})$ between the actual load production and our forecasted load production over the course of a day.

Figure 5.1.2 illustrates the production load in kilowatts (kW) for a specific day. It can be observed that the majority of the hourly values throughout the day cluster around the mean value, indicating relatively low daily variance.

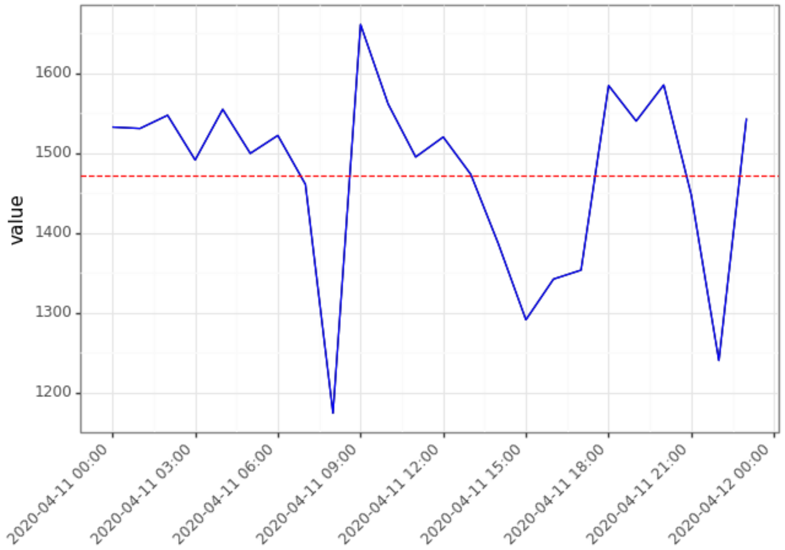

Figure 5.1: The hourly values of electricity production (kW).

We can expect the predictions to have less variation or deviation from the mean value due to the observed low daily variance.

5.2 Dataset description

The production load dataset includes information collected over the period **'2020/01/01 00:00:00' - '2022/05/01 23:45:00'** i.e format $(YY/MM/DD)$. Measurements were collected every 15 minutes. More details about the power plant are discussed in section 1.2.3.

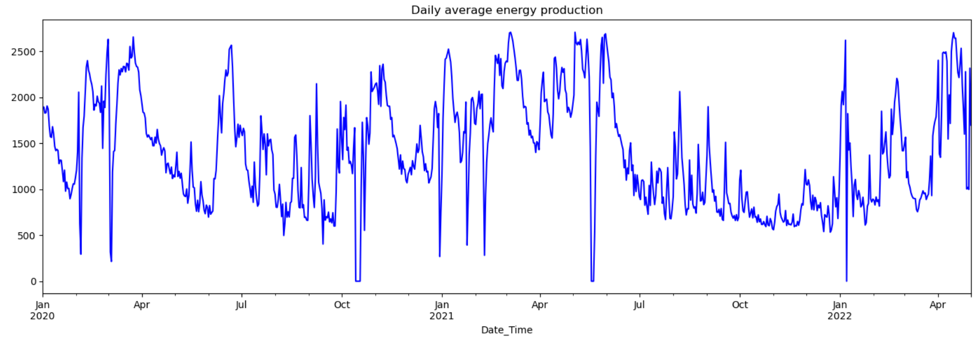

Figure 5.2: Average daily power production (kW)

The average daily power production of the power plant (Figure 5.2) over the specified time period, measured in kW, represents power plant's operational performance and output on a day-to-day basis. It should be noted that any zero values in the graph signify periods of maintenance or times when the power plant was shut down. The values of the load are lower from spring to the end of summer compared to the rest of the year. This pattern can suggests the presence of yearly seasonality and even with 2.32 years of data, it might still be possible to identify some patterns in the data.

In addition to the production data of the power plant, meteorological data used for prediction was sourced from six measuring stations around the power plant and they are available at [1]. The meteorological data includes following parameters, that were previously discussed in Chapter 3.

- **Precipation** (mm)
  The mean amount of precipitation over the last hour.
- **Snow depth** (m)
  The measurement of snow depth at a selected location.
- **Temperature** ($°C$)
  The mean temperature over the last hour.
- **Atmospheric preassure** (hPa)
  Value of atmospheric pressure, available up to 10 km.
- **Soil water content** (m^3m^3)
  We gathered data on the volume of water in each available soil layer i.e from 0 to 289 cm..
- **Soil moisture index**
  The index was collected for soil layer depths ranging from 0 to 100 cm.
- **Evapotranspiration** (mm)
  This parameter, accumulates evapotranspiration over the previous hour.

5.3 Data preprocessing

The following section provides a description of the data preprocessing procedures used for the collected dataset.

5.3.1 Electricity production

The collected dataset does not include any missing values. It was first resampled into daily average values of production load. As mentioned in Section 5.2, zero values, which indicate periods of maintenance or shutdown of the power plant, present a potential issue. Including these values in our analysis could negatively impact our results and reduce the predictive accuracy. The reason is that these periods are subject to a range of other factors and, as such, are not predictable.

First, we calculated the daily mean power production, which was found to be 1404.144 kW and represents the typical daily performance of the plant. Then, we set a threshold value for replacing the zero values; this was defined as the minimum of the 7-day moving average. Any value below this threshold is treated as noise.

This choice was made because the analyzed power plant typically has similar output across consecutive days, and we do not expect any substantial changes that could affect production within such a short time period. The minimum value we observed from the 7-day moving average was 385.308 kW, so we set our threshold to be 400 kW for both hourly and daily data. Such values are then replaced with the corresponding value after the moving average computation.

Figure 5.3.1 presents a comparison of the daily mean power production, the 7-day moving average, and the actual data. As can be seen, the 7-day moving average effectively smooths out the data while maintaining a similar production load and preserving the data patterns.

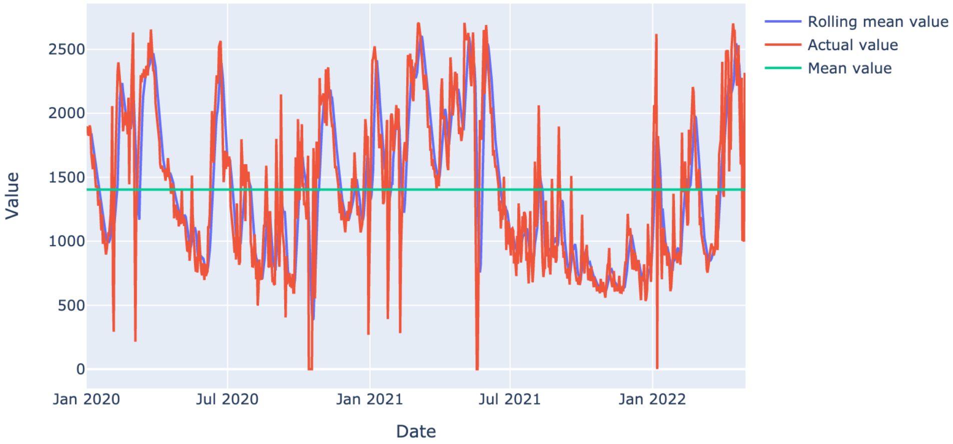

Figure 5.3: Mean and 7-day Moving Average for Daily Load Production (kW)

Given that the task is to predict the production for the next day, we also resampled the dataset into hourly intervals by taking the mean of the 15-minute intervals. Then, we replaced all zero values according to a threshold value, as explained above for daily data.

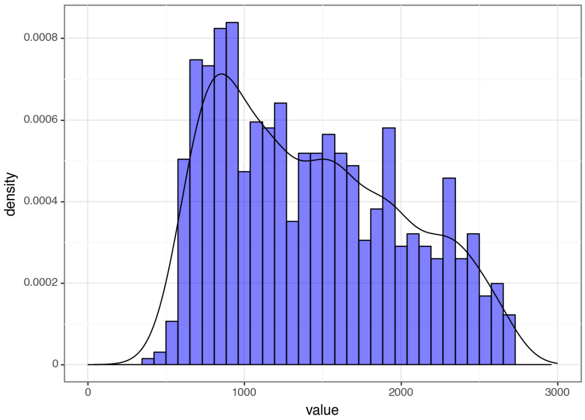

Figure 5.4: Distribution plot of daily average load (kW) after replacing zero values

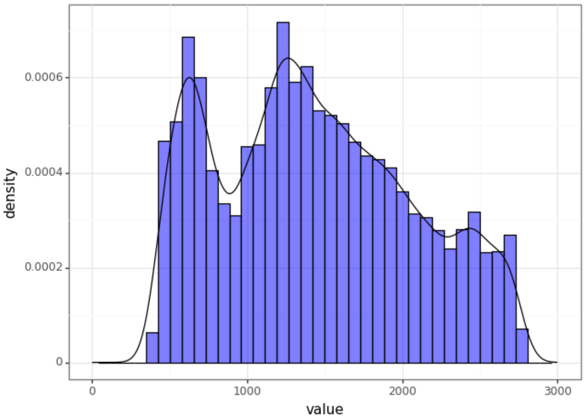

Figure 5.5: Distribution plot of hourly average load (kW) after replacing zero values

In the following text, we will evaluate some basic statistical properties. The density plots are displayed in Figures 5.3.1 and 5.3.1. The mean value of 1437.212 kW shows the average hourly energy production. This is the central value around which the individual hourly production values are distributed. The standard deviation of 626.149 kW signifies higher variability in the data over time period. The minimum value of 388.607 kW is the lowest hourly energy production recorded and the maximum value of 2772.000 kW is the highest hourly energy production recorded.

	count	mean	std	min	25%	50%	75%	max
hourly mean value	20448.000	1437.212	626.149	388.607	934.438	1381.625	1896.062	2772.000
daily mean value	852.000	1425.290	580.847	403.562	912.573	1359.125	1886.586	2707.198

Table 5.1: Hourly average energy production values statistics

The first quartile indicates that a quarter of the values are below 934.438 kW. The median for our dataset is 1381.625 kW. This value can be a helpful measure of the ”typical” production, especially when the data is skewed. And it seems that the majority of values are below 1896.062 kW. For daily data, the values are similair.

5.3.2 Hydrometeorological data

The following text will discuss the analysis and selection of individual features that will be used as exogenous variables for further analysis. Features were analyzed on the hourly data intervals, which were then resampled to daily intervals for the daily data. Based on the findings in Section 3, it is evident that the majority of the selected meteorological variables exhibit a delayed effect on electricity production.

The relationship between the selected hydrometeorological features and the production load is discussed in the next section 5.4 and only the most significant findings will be presented.

We began with **precipitation**. The delayed effect can be explained by various factors such as the time it takes for the water to infiltrate the soil, contribute to the groundwater, and eventually enter the river.We calculated a two-week cumulative sum for precipitation data from all locations. What was found to be the most relevant for our target value. This time period was used for other data as well.

A similar approach was applied to **evapotranspiration**, although we expected an inverse relationship in this case, that could indicate lower water level in the river.

The mean **temperature** across all locations at a specific point in time was not found to have impact on our present target value. However, we speculate that there could be a relationship between the target value and the average temperature over a certain time period, which might provide more useful information. The same approach was applied to **atmospheric pressure**. Atmospheric preassure values were found very similair for all locations.

Snow depth was calculated as a mean value over hourly time period and cummulatively summed for all locations.

The **soil moisture index** was calculated as the mean value over a specified time period, aggregating from data across all locations. This gave us additional insights about the soil wetness around the power plant over a larger geographical range and extended time period.

We aggregated the data from all soil layers to calculate the **volumetric soil water content** as a mean value over a specific time period. This measurement informs us about the amount of water potentially available to contribute to the river.

All missing values were removed in preparation for analyzing the relationships between the features and our target value, the production load.

Figure 5.3.2 visualizes these features after the data preprocessing. The Soil Moisture Index was first removed from the features because it was redundant, given the availability of the volumetric soil water content.

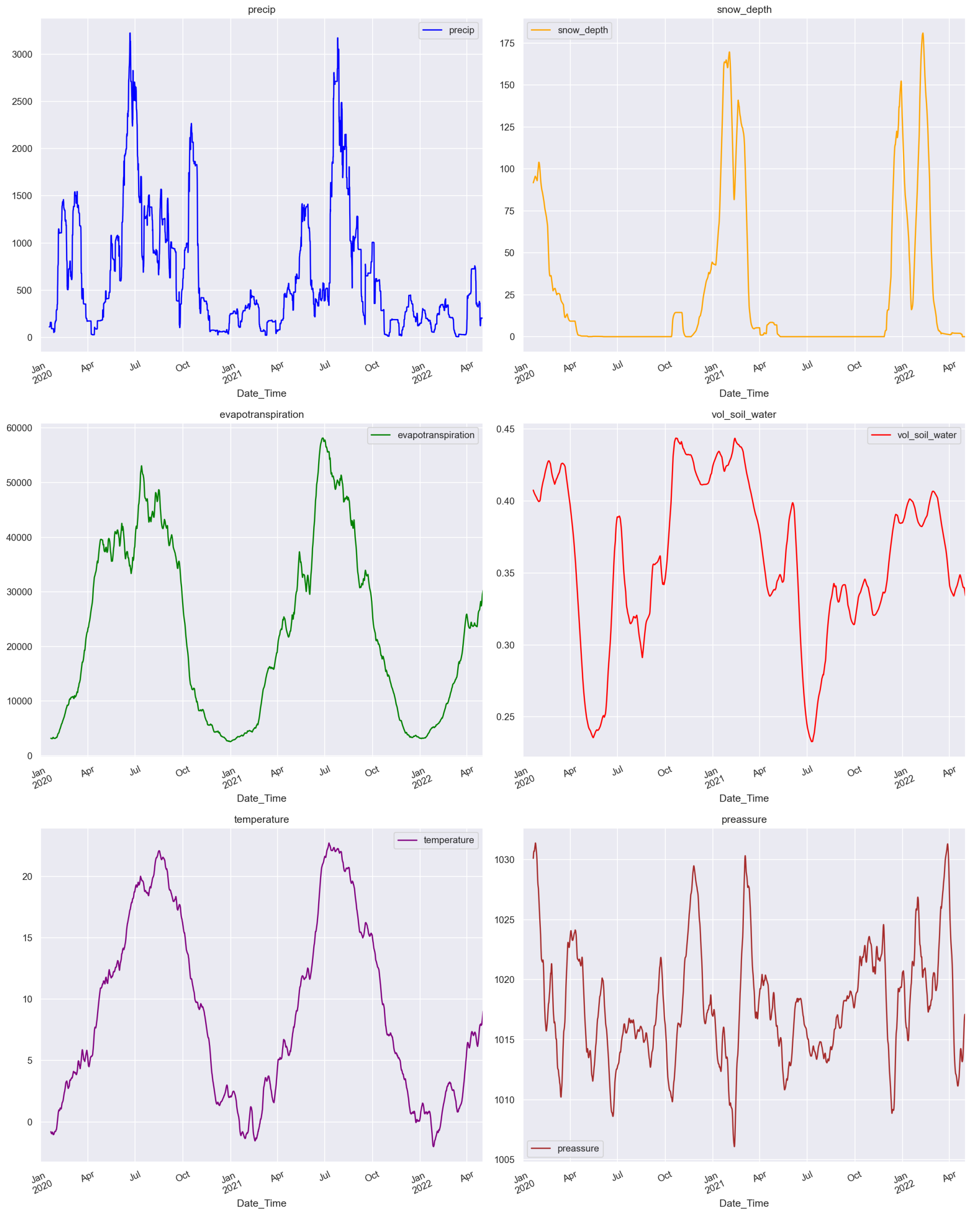

Figure 5.6: We can observe an obvious yearly pattern in all hydrometeorological features. (kW)

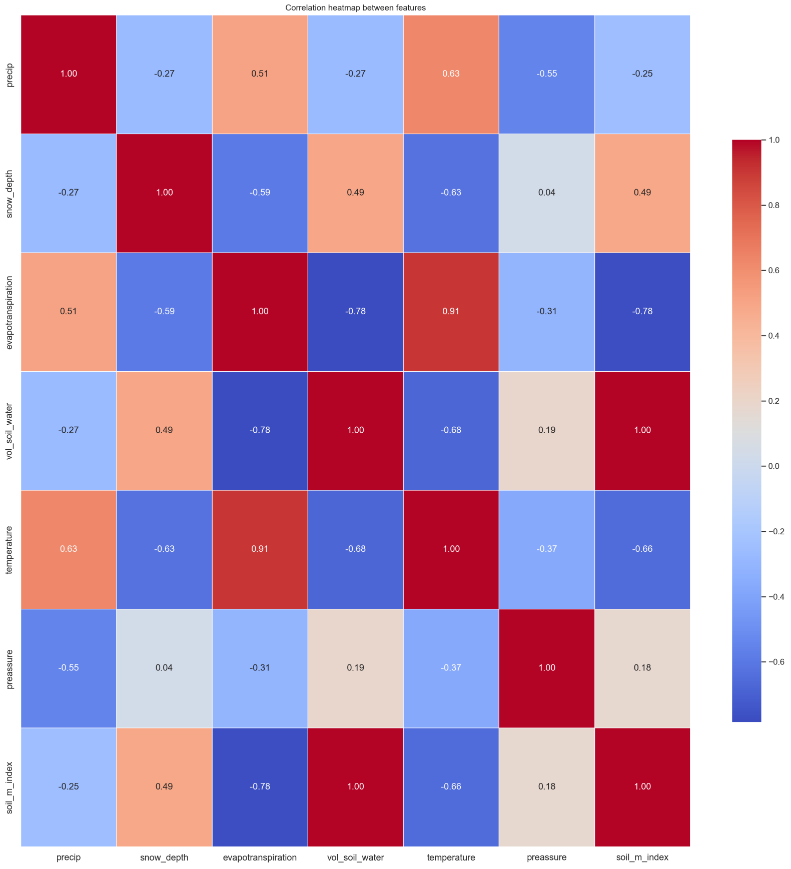

Figure 5.7: Visualization of relationships based on correlation coefficients for hourly data.

5.3.3 Relationships between hydrometorological features

Based on the possible relationships 5.3.2 between the features we have identified the following variables to be used as exogenous variables.

The variable **precipitation** demonstrates strong positive correlations with both temperature and evapotranspiration. Moreover, previous studies 2 have indicated that models incorporating precipitation data over time show strong performance.

Given the very strong positive correlation between the **soil moisture index** and **volumetric soil water**,indicating that they essentially measure the same attribute, we have opted to use volumetric soil water content in our model.

Temperature has strong relationship with evapotranspiration and moderate positive correlation with precipitation.

While **pressure** appears to have relatively weak correlations with most of the variables.

We have decided to exclude **snow depth** from our dataset, despite its significant correlations with evapotranspiration and temperature. The strong negative correlations may introduce unnecessary complexity into our model.

Similarly, we've chosen to omit **evapotranspiration** from our feature set, despite its very strong positive correlation with temperature and strong negative correlation with volumetric soil water content. This decision aligns with our goal to avoid multicollinearity, as evapotranspiration appears to be highly predictable from other selected variables, particularly temperature.

5.3.3.1 Summary

The features selected for the analysis of the relationship between power production and meteorological data include precipitation, volumetric soil water content, temperature, and atmospheric pressure.

5.4 Relationship between power production and meteo data

Our analysis indicates a very small dependency, specifically 0.043, between the production load and precipitation. This suggests that there is little to no linear relationship between these two variables. However, the relationship might be of a different form than linear. It is likely that a more comprehensive dataset of precipitation measurements would be beneficial, considering more locations and rivers flowing into the Hron River. Based on Figure 5.4, we have come to the conclusion that we probably do not have a enough data. these sources highlight approaches where models that use precipitation values achieve high accuracy.

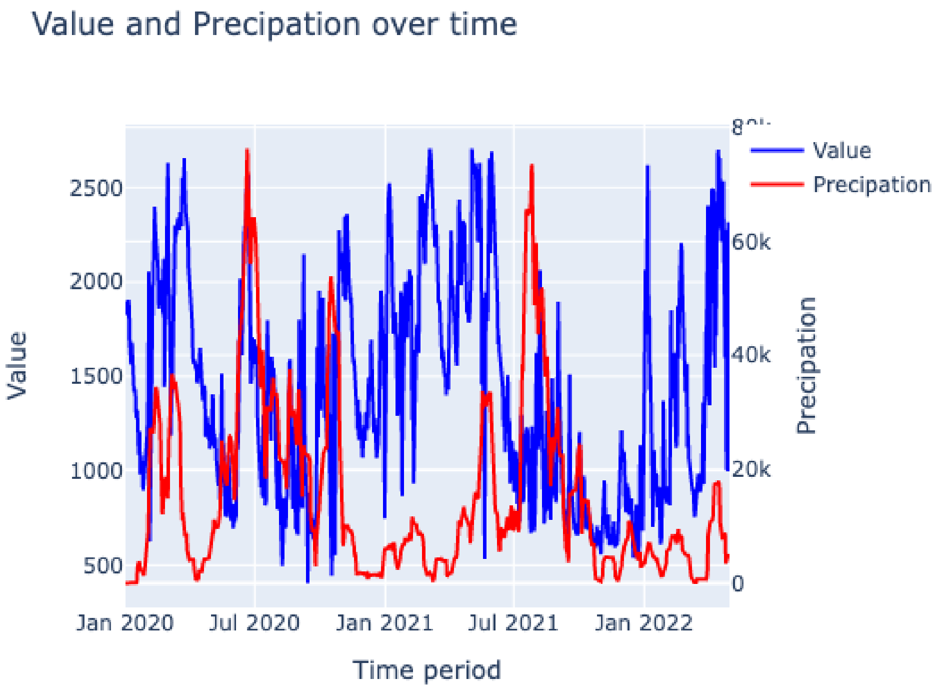

Figure 5.8: Value and Precipation over time

Surprisingly, the correlation coefficient between value and volumetric soil water content is higher than we expected, at 0.361. This suggests a degree of linear relationship. As the volumetric soil water content increases, the value tends to increase as well. Given this, we will consider including volumetric soil water content as a feature in our model. Figure 5.4 validates our prior assumptions regarding 5.3.3.

As expected, the production load tends to decrease as the temperature increases. This might be due to the evaporation process, confirming our assumptions from previous chapters. As can be seen in Figure 5.4, the production tends to increase with lower temperature. However, numerous other factors also influence production. The reason might also be the fact that some of the water from the river is used for irrigating fields close the power plant.

Based on the very low correlation coefficient of −0.11, we've decided to exclude pressure as a feature in our model.

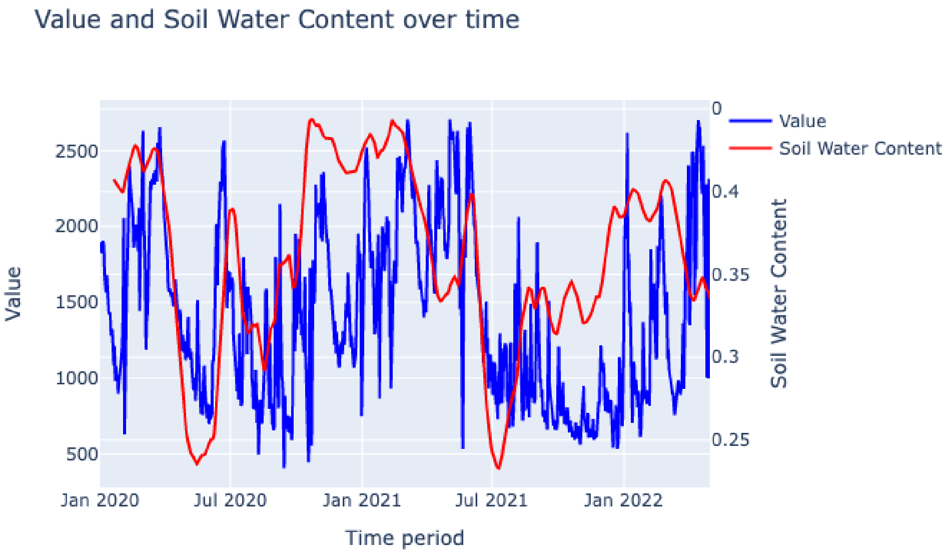

Figure 5.9: Value and Soil water content over time for daily data

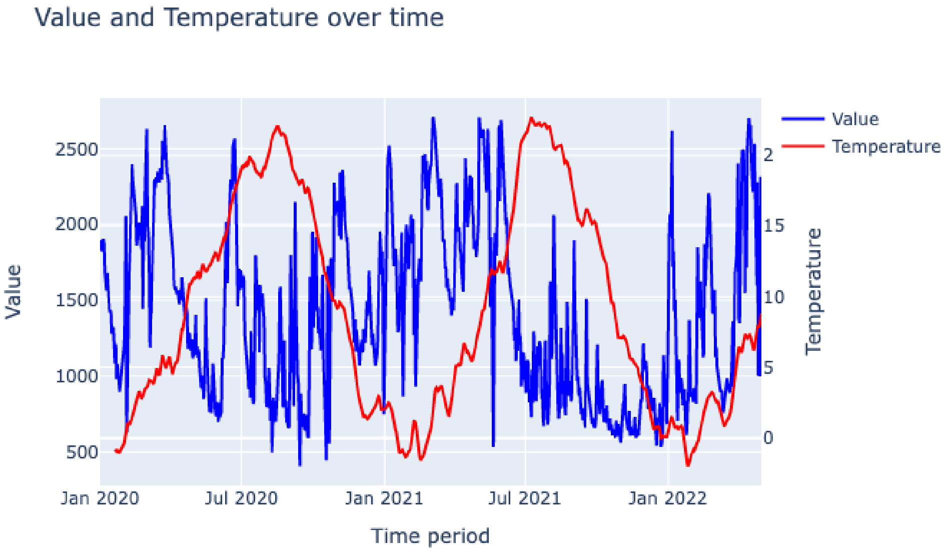

Figure 5.10: Value and Temperature water content over time for daily data

5.5 Traditional statistical methods

We used two models Exponential smoothing (ETS), and $SARIMAX$ for forecasting using Traditional statistical methods. In the following text we discuss the models selection, training, computation of prediction intervals and selected metric. Both models, Exponential Smoothing (ETS) and Seasonal Autoregressive Integrated Moving Average with Exogenous Regressors (SARIMAX), were implemented using the Statsmodels library in Python.

5.5.1 Selection of the model

In study [6] was concluded that if we want to prevent underfitting a model, we should use AIC as an criterion for model selection.One of the commonly used metrics for model selection in forecasting is the Akaike Information Criterion. It is computed as follows.

$$AIC = 2k - 2ln(\theta) \tag{5.1}$$

where k represents the number of parameters in the model and θ denotes the likelihood of the model.

The proposed models were chosen based on the lowest AIC value.

5.5.2 Training and Test

The were trained using the walk-forward method [22]. Basic process is as follows.

- The dataset is splitted into training test set.
- The model is fitted on the training data, then used to forecast the next time step.
- The prediction is evaluated against the actual value.
- The next observation from the test set is included in the training set, and the model is refitted.
- These steps are repeated until all observations in the test set have been included in the training set and forecasted.

Following the approach, the dataset was divided into a training set and a test set. To maintain consistency and ensure that the same test set was used for other solutions.

5.5.3 Validation metric

Root Mean Square Error (RMSE) is a commonly used metric in regression analysis and forecasting to measure the amount of variance in the prediction error.

RMSE is calculated by taking the square root of the average squared difference between the predicted and actual values. The formula for RMSE is as follows.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{5.2}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and n is the number of observations. The lower the RMSE, the better the model's performance.

5.5.4 Prediction Intervals

Prediction intervals are intervals within which we expect our predicted value $\hat{y}_t$ will lie, with a specified probability, or the uncertainty of our predictions. Exponential smoothing can generate prediction intervals, and the calculation differs depending on whether the model is additive or multiplicative [21]. Depending on the prediction intervals, we can evaluate the accuracy of the prediction model.

Prediction interval is defined as follows

$$\hat{y}_{T+h|T} \pm c\hat{\sigma}_h, \tag{5.3}$$

where $\hat{\sigma}_h$ is an estimation of the standard deviation of the h-step forecast distribution, and c is a confidence interval depending on the coverage probability.

Then we can evaluate accuracy of the prediction model based on such intervals.

We used bootstrapping, assuming that residuals are uncorrelated, have constant variance, and the true values are randomly distributed around the predicted value.

The calculation for an additive model:

$$y_t = \hat{y}_{t|t-1} + e_t, \quad e_t \sim \mathcal{N}\left(0, \sigma^2\right) \tag{5.4}$$

The calculation for a multiplicative model:

$$y_t = \hat{y}_{t|t-1} \cdot (1 + e_t), \quad e_t \sim \mathcal{N}\left(0, \sigma^2\right) \tag{5.5}$$

Where e_t is sampled from the residuals.

For the case of multiple steps ahead, that means a larger error and wider prediction intervals, as $\hat{\sigma}_h$ will increase with the number of prediction steps.

5.6 TCN solution

TCN model was implemented using Darts library [42], as proposed in the section 5.6 for 50 epochs.

5.6.1 Data preprocessing

In the following section the preprocessing of dataset before feeding into the network will be described.

5.6.1.1 Dataset splitting

A standard ratio of $(70\%, 20\%, 10\%)$ was followed for splitting the dataset into the train, validation, and test sets, respectively. In the case of recurrent neural networks, the order of the data points can matter due to their memory.

5.6.1.2 Data normalization

Based on prior research [31], we decided to prioritize normalization over standardization.

Normalization rescales the values to a range, typically $(0, 1)$. Technique is useful when the ranges of the features are significant for the model's performance and some algorithm do not perform good on different scales [17].

Standardization centers the values around zero by subtracting the mean and then scales them by dividing with the standard deviation. This approach ensures that the resulting distribution has a unit variance. Standardization is less influenced by outliers, making it suitable for cases where robustness to extreme values is important.

For Standardization first, the mean is subtracted each feature, resulting in standardized values with a zero mean. Then, the values are divided by the standard deviation to achieve a unit variance distribution. This procedure allows the features to have comparable scales.

$$\text{Standardization}(x) = \frac{x - \text{mean}(x)}{\text{std}(x)} \tag{5.6}$$

For normalization, we employed the Min-max scaler technique, which rescales the values to a predefined range, typically between 0 and 1. To accomplish this, we subtracted the minimum value from each feature and divided the result by the difference between the maximum and minimum values. The Min-max scaler was fitted exclusively on the training set, ensuring that the validation and test sets remain independent of the specific values used for normalization. By applying the same transformation to the validation and test sets, we maintain consistency while preserving the integrity of these datasets.

$$\text{Min-max scaling}(x) = \frac{x - \min(x)}{\max(x) - \min(x)}. \tag{5.7}$$

5.6.1.3 Handling Time Series patterns

Time series models such as ARIMA and other traditional models often require the removal of trend and seasonality components before modeling. In the case of neural networks, this step is not necessary as the model has the capability to learn and capture trend and seasonality patterns inherently [31].

However, incorporating trend and seasonality information as additional inputs can potentially enhance the model's performance.

We discussed in previous sections that our data might have yearly seasonal pattern and we tranformed 'Date time' index using sine and cosine transformations to get signal about time of the year and then feed them to the network as exogenous variables.

Let t represent the timestamp in days, and *year* denote the length of one year in days (approximately 365.2425). The transformation is then performed as follows.

$$\mathrm{sin(t)} = \sin\left(\frac{2\pi t}{\mathrm{year}}\right) \tag{5.8}$$

$$\mathrm{cos(t)} = \cos\left(\frac{2\pi t}{\mathrm{year}}\right) \tag{5.9}$$

5.6.1.4 Data windowing

We created a window of consecutive samples from the data.

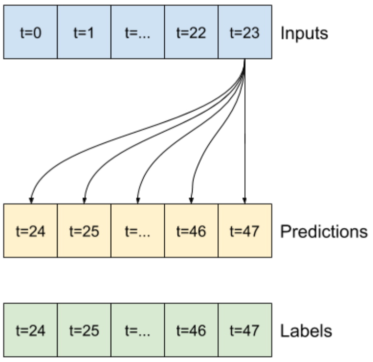

Figure 5.11: Demonstration example where given the past 24 values, the model will predict next 24 values. [17]

During the experiments, we generated multiple windows with different lengths for both the input and label sequences in order to predict the output sequence. It was observed that longer sequences were computationally too expensive to compute within a reasonable time frame. Therefore, the best performance was achieved using a daily input width. The label width, representing the length of the output sequence, was set to 60 previous days.

5.6.2 Training and Validation

5.6.2.1 Loss function

We choosed the Mean Square Error (MSE) loss function because our goal is to predict a continuous variable, i.e value of production load.

MSE is defined as a sum of squared differences between predicted and actual values divided by number of samples.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{5.10}$$

where n is number of our samples, y_i is an actual value and $\hat{y}_i$ is our prediction.

5.6.2.2 Validation metric

We used $RMSE$ as validation metric, which we discribed in previous section.

5.6.2.3 Overfitting prevention

The early stopping strategy, as described in book [31], is a commonly used technique to prevent neural networks from overfitting and ensure that the model generalizes well to unseen data. The principle is as follows.

At the end of each epoch during training, the loss on the validation set is evaluated. If the validation loss does not decrease compared to the previous epoch, the current model's performance is compared to the best model achieved so far.

The best model is determined based on having the lowest validation loss. If the current model performs better, it replaces the previous best model. The early stopping algorithm continues until a stopping criterion is met.

In our implementation, we have chosen to stop training if there is no improvement in the validation loss for 10 epochs. The number of epochs to wait before stopping training was chosen based monitoring the behavior of the validation loss over multiple training iterations.

It was observed that significant improvements in the validation loss typically occurred within the first few epochs. Then, if the loss does not decrease after 10 epochs, we consider it as a sign of potential overfitting. At this point, we store a copy of the model's parameters, representing the best performance achieved so far.

5.7 Results

In this section, we evaluate our results from experiments for single-step prediction on daily average data and multi-step prediction on hourly average data. More details and a discussion of the results can be found in the Conclusion of this thesis (see Chapter 6).

The daily average dataset was too short for using neural networks, so we included the TCN model into single-step forecasts only for comparison with our best-performing model.

Tables 5.2 and 5.3 provide information about average $RMSE$ errors and present models which we used in our analysis. The ETS model is specified by three components: error, trend, and seasonality, each of which can be either additive (A) or multiplicative (M), or none (N). In this case, all are set to None, indicating a simple exponential smoothing model without trend or seasonality adjustments. The SARIMAX model is specified by two sets of three parameters: one for the non-seasonal components of the model (p, d, q), and one for the seasonal components (P, D, Q). In the case of our SARIMAX(1,1,1)(0,0,0) model, this suggest an ARIMA(1,1,1) model with no seasonal components.

5.7.1 Single-Step Forecast

The presented results suggest that the Simple Exponential Smoothing model provided the most accurate single-step forecasts for daily average data. It had the lowest Root Mean Square Error (RMSE) of 225.19 kW, making it the best-performing model. It seems that the daily dataset was too short for accurate forecasting using the Temporal Convolutional Network (TCN). This is indicated by the higher RMSE of the TCN model at 318.34, which is significantly higher than the RMSE of the Simple Exponential Smoothing model.

Model Number	**Model**	**Exogenous variables**	**Test RMSE (kW)**
1	ETS(N,N,N)	None	225.19
2	SARIMAX(1,1,1)(0,0,0)	None	227.2
3	SARIMAX(1,1,1)(0,0,0)	Precipitation	229.3
4	SARIMAX(1,1,1)(0,0,0)	Soil Water Content	233.8
5	SARIMAX(2,0,1)(0,0,0)	Temperature, Precipitation	245.9
6	SARIMAX(1,1,1)(0,0,0)	Temperature, Soil Water Content	246.4
7	TCN	None	318.34

Table 5.2: Single-step forecast experiments results

Interestingly, the inclusion of exogenous variables, such as precipitation, soil water content, and temperature, in the SARIMAX models did not improve the forecasting accuracy. In fact, models that used these variables had higher RMSE values than the SARIMAX model without any exogenous variables, indicating that these variables might have acted more as noise, reducing the model's performance.

Figure 5.7.1 provides a visualization of the predictions made by the Exponential Smoothing method on the training dataset. The fact that the majority of actual values lie within the determined confidence interval suggests that the model's predictions are reasonably accurate, and the estimated uncertainty is appropriate.

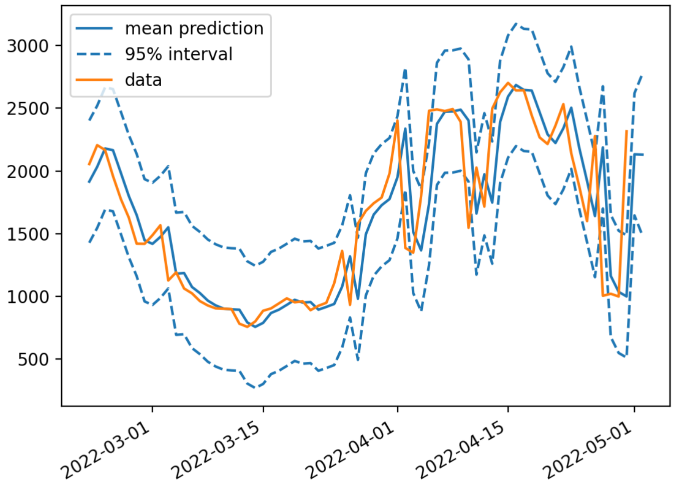

Figure 5.12: Daily forecast for Exponential smoothing method with 95% confidence interval.

The TCN model's predictions are shown in the following figures 5.7.1 and 5.7.1. Despite the higher RMSE, the TCN's forecasts remaind those from the Exponential Smoothing model, which might suggest that even if our daily average dataset was short, the TCN model still managed to capture some meaningful patterns for prediction. However, it's clear that the Simple Exponential Smoothing and SARIMAX models without exogenous variables outperformed the TCN model.

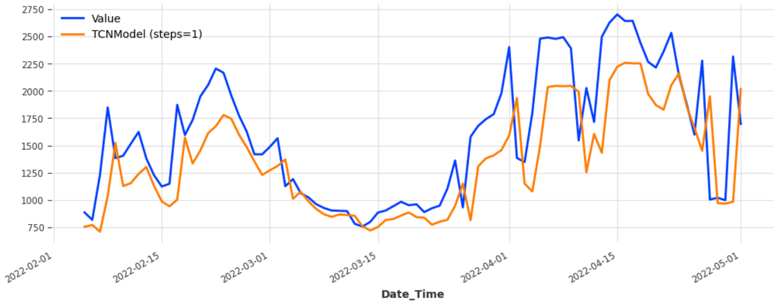

Figure 5.13: Daily forecast for Temporal Convolutional Network.

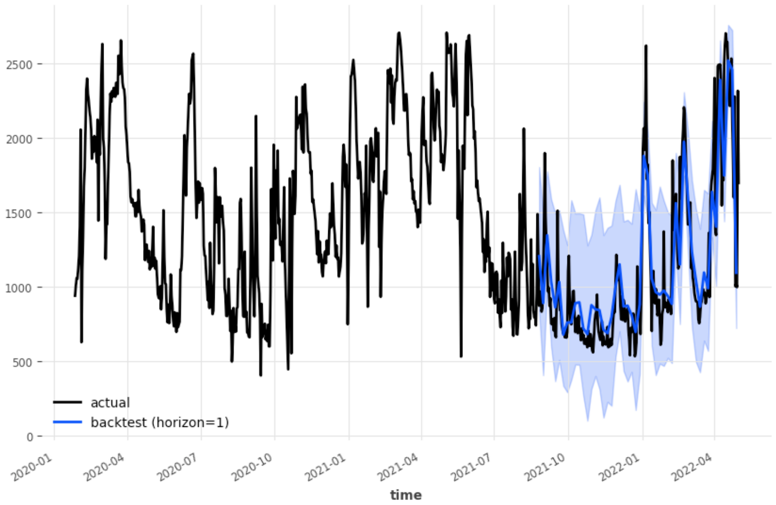

Figure 5.14: Predicted values for TCN with 95% interval.

5.7.2 Multi-Step forecast

In the case of multi-step predictions, or forecasts made 24 hours in advance, a similar pattern can be observed. The Simple Exponential Smoothing model again demonstrated the highest performance, providing the most accurate forecasts with an RMSE of 355.37 kW. As for the TCN model, despite utilizing different combinations of exogenous variables, it performed significantly worse than the Simple Exponential Smoothing model. The reason behind such poor results could be that we did not give the model a sufficient amount of history for predictions, or we did not set the parameters accurately, which might have resulted in inadequate receptive fields.

Model Number	**Model**	**Exogenous variables**	**Test RMSE** (kW)
1	ETS(N,N,N)	None	355.37
2	SARIMAX(1,1,1)(0,0,0)	None	359.2
3	SARIMAX(1,1,1)(0,0,0)	Precipitation	362.1
4	SARIMAX(1,1,1)(0,0,0)	Soil Water Content	367.2
5	SARIMAX(2,0,1)(0,0,0)	Temperature, Precipitation	381.9
6	SARIMAX(1,1,1)(0,0,0)	Temperature, Soil Water Content	383.4
7	TCN	None	560.3
8	TCN	Sin Yearly + Cos Yearly	551.2
9	TCN	Precipation + Sin Yearly + Cos Yearly	611.4
10	TCN	Precipation	618.6

Table 5.3: Multi-step forecast experiments results

The inclusion of exogenous variables did not improve the models' performance very much of the TCN models performance. This suggests that these variables may not provide additional useful informatio. It may also suggest that the models are struggling to appropriately incorporate this additional information into their forecasts and we need to propose different solutions.

The figures 5.7.2 and 5.7.2 provide visual representations of the multi-step predictions made by the Exponential Smoothing and TCN models, respectively. Once again, most of the actual values lie within the 95% confidence intervals, implying that the uncertainty estimation of these models is reasonable. However, the width of the confidence intervals for the TCN model appears to be larger than those for the Exponential Smoothing model, suggesting that the TCN model has a higher level of uncertainty about its predictions.

These results suggest that for both single-step and multi-step forecasting tasks, simpler models like Simple Exponential Smoothing and SARIMAX without exogenous variables outperform more complex models like TCN. Moreover, the inclusion of exogenous variables appears not to improve, but rather to impair, the models' performance.

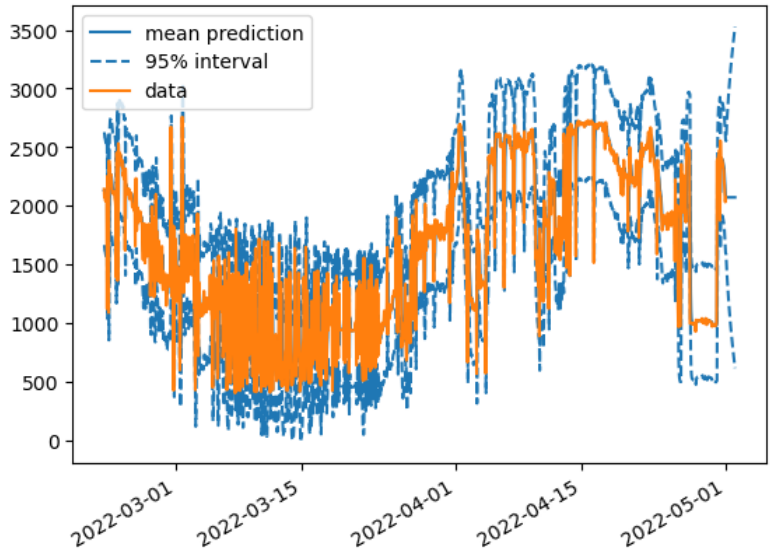

Figure 5.15: Predicted values for Exponential smoothing with 95% interval.

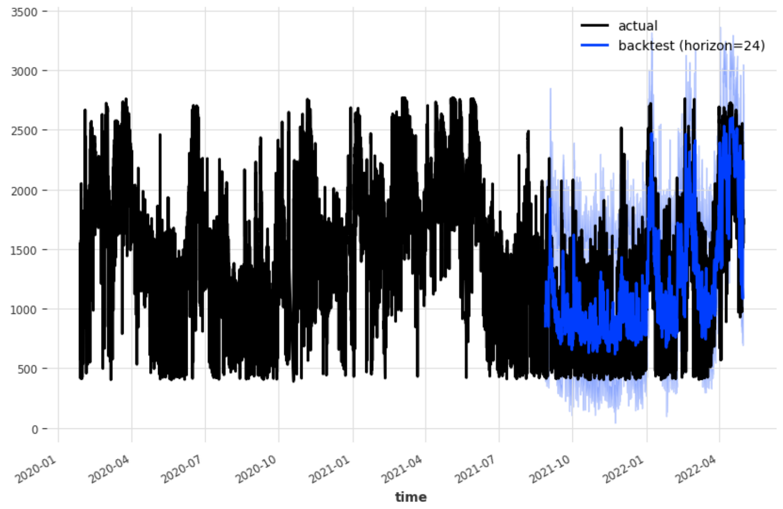

Figure 5.16: Multi-step forecasts for TCN with sin yearly and cos yearly exogenous variables with 95% interval.

Chapter 6

Conclusion

The objective of this thesis was to develop predictive models for next-day forecasts of hydroelectric power production. We initiated our work with an analysis of the domain, including an understanding of the operational mechanisms of hydropower plants and the hydrological cycle.

Collecting the necessary data was a challenging task. Electricity production data were gathered from a small hydro power plant in Želiezovce, and hydrometeorological data were sourced from nearby locations. Our hydrometeorological dataset consisted of features such as precipitation, atmospheric pressure, temperature, snow depth, soil water content, and the soil moisture index.

Next, we analysed the collected dataset, identifying potential relationships between production and the features. We chose temperature, precipitation, and soil water content as exogenous variables.

In the theoretical section of this thesis, we discussed the theoretical background and evaluated existing techniques for time series prediction and hydro power prediction.

We proposed various predictive methods for single-step, for daily average production, and multi-step forecasting, for hourly average production. We selected two traditional forecasting methods, SARIMAX (as exogenous variables can be included) and Exponential Smoothing. After analyzing other potential methods, we chose the Temporal Convolutional Network due to its potential to outperform other neural networks such as LSTM.

Our findings revealed that the statistical exponential smoothing model, without any exogenous variables, delivered the best performance for both multi-step and single-step forecasts. This outcome is likely due to the fact that the hydrometeorological data were gathered from locations near the power plant, not directly from it, which may have impacted our results. For single-step prediction, our most accurate model achieved an average RMSE of 225.19 kW. For multi-step prediction, the Exponential Smoothing model performed best with an RMSE of 355.37 kW. The SARIMAX model performed similarly to the ETS model, outperforming the Temporal Convolutional Network. The TCN model had an RMSE of 318.84 kW for single-step prediction and 551.2 kW for multi-step prediction, with the model using exogenous variables to determine the time of the year.

Interestingly, the inclusion of exogenous variables such as precipitation and temperature, which we initially thought would improve model performance, seemed to act as noise and impair accuracy. Even the inclusion of soil water content, which correlated more closely with electricity production, did not enhance our results.

One of our goals was to determine whether our forecasts could assist the hydropower plant in selling electricity contracts. These contracts are legal agreements between an electricity generator (in our case, a hydro power plant) and a buyer (such as an energy trader).They specify details like the amount of electricity to be supplied. Accurate forecasts of hydroelectric power production can be extremely valuable in this context. For the seller, accurate forecasts can assist in planning and negotiating the amount of power they can commit to delivering under the contract. It can also help manage risk and price contracts more accurately. For the buyer, access to accurate forecasts from the seller can provide confidence in the reliability of the power supply, assist in planning their own power grid management and distribution, and facilitate more informed decisions about which contracts to sign.

The ETS solution for forecasting hydroelectric power production can be used in the process of selling contracts for a hydro power plant. Most of the actual values fell within the 95% confidence interval, suggesting our model's uncertainty estimation accurately reflects potential outcomes. However, a good fit within the confidence interval does not necessarily imply accurate predictions. We considered other metrics, such as RMSE, to evaluate prediction accuracy.

All of the goals of this thesis were fulfilled.

6.1 Future Work

The impact of exogenous variables worsens the model's accuracy. Our initial analysis suggested that including exogenous variables such as precipitation and temperature could enhance the model's performance. However, even the inclusion of soil water content, which correlates more closely with electricity production than precipitation and temperature, did not improve our results. This finding shows the importance of collecting data from a larger number of stations in the vicinity of the run-of-river power plant.

In terms of different forecasting model solutions, incorporating probabilistic models like Gaussian processes or Bayesian networks could potentially yield more accurate results. Additionally, it is important to investigate the reasons behind the poor accuracy of our proposed TCN network.

We intend to refine our data collection process and expand the scope of our analysis. Greater emphasis will be placed on the selection of relevant exogenous variables and fine-tuning our models to better incorporate the hydrometeorological features. Integrating advanced probabilistic and hybrid models into our methodology will be a significant focus of our future work. Given the potential of these models to handle complex patterns and uncertainties in data, we think that they will bring us closer to achieving more accurate and reliable predictions for hydroelectric power production.

Bibliography

[1] Meteomatics AG. *Weather API.* [accessed September 01, 2022]. 2022. URL: `https://www.meteomatics.com/en/api/getting-started/`.

[2] Mauricio E Arias et al. “Impacts of hydropower and climate change on drivers of ecological productivity of Southeast Asia’s most important wetland”. In: *Ecological modelling* 272 (2014), pp. 252–263.

[3] Shaojie Bai, J Zico Kolter, and Vladlen Koltun. “An empirical evaluation of generic convolutional and recurrent networks for sequence modeling”. In: *arXiv preprint arXiv:1803.01271* (2018).

[4] L. Baker. *Beginner’s Guide to Correlation Analysis: Learn The One Reason Your Correlation Results Are Probably Wrong.* Bite-Size Stats. Lee Baker. URL: `https://books.google.cz/books?id=PPlfDwAAQBAJ`.

[5] George EP Box et al. *Time series analysis: forecasting and control.* John Wiley & Sons, 2015.

[6] Hamparsum Bozdogan. “Model selection and Akaike’s information criterion (AIC): The general theory and its analytical extensions”. In: *Psychometrika* 52.3 (1987), pp. 345–370.

[7] Enzo Busseti, Ian Osband, and Scott Wong. “Deep learning for time series modeling”. In: *Technical report, Stanford University* (2012), pp. 1–5.

[8] Christopher Chatfield. *The analysis of time series: theory and practice.* Springer, 2013.

[9] Robert B Cleveland et al. “STL: A seasonal-trend decomposition”. In: *J. Off. Stat* 6.1 (1990), pp. 3–73.

[10] Lars Dannecker. *Energy time series forecasting: efficient and accurate forecasting of evolving time series from the energy domain.* Springer, 2015.

[11] Maria Grazia De Giorgi et al. “Comparison between wind power prediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN)”. In: *Energies* 7.8 (2014), pp. 5251–5272.

[12] DenisBoigelot. *The correlation between different pairs of X and Y.* [accessed April 22, 2023]. 2011. URL: `https://en.wikipedia.org/wiki/File:Atomic_force_microscope_block_diagram.svg`.

[13] Yagob Dinpashoh et al. “Impact of climate change on streamflow timing (case study: Guilan Province)”. In: *Theoretical and Applied Climatology* 138 (2019), pp. 65–76.

[14] Davis W Edwards. *Energy trading and investing: Trading, risk management and structuring deals in the energy market.* McGraw-Hill Education, 2009.

[15] Saeid Eslamian. *Handbook of engineering hydrology: modeling, climate change, and variability.* CRC Press, 2014.

[16] K.H. Fasol. “A short history of hydropower control”. In: *IEEE Control Systems Magazine* 22.4 (2002), pp. 68–76. DOI: `10.1109/MCS.2002.1021646`.

[17] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow.* ” O’Reilly Media, Inc.”, 2022.

[18] Geoffrey D Gooch, Alistair Rieu-Clarke, and Per Stalnacke. *Integrating water resources management.* Iwa Publishing, 2010.

[19] Manfred Hafner and Giacomo Luciani. *The Palgrave Handbook of International Energy Economics.* Springer Nature, 2022.

[20] Charles C Holt. “Forecasting seasonals and trends by exponentially weighted moving averages”. In: *International journal of forecasting* 20.1 (2004), pp. 5–10.

[21] Rob Hyndman et al. *Forecasting with exponential smoothing: the state space approach.* Springer Science & Business Media, 2008.

[22] Rob J Hyndman and George Athanasopoulos. *Forecasting: principles and practice.* OTexts, 2018.

[23] Asif Iqbal et al. “A fuzzy expert system for optimizing parameters and predicting performance measures in hard-milling process”. In: *Expert Systems with Applications* 32.4 (2007), pp. 1020–1027.

[24] Yuji Ito and Kazuro Momii. “Potential effects of climate changes on evaporation from a temperate deep lake”. In: *Journal of Hydrology: Regional Studies* 35 (2021), p. 100816.

[25] B. Surina J. Hrasko V. Linkes. *Soil Map Slovakia - podne typy.* [accessed January 12, 2023]. 2011. URL: `https://esdac.jrc.ec.europa.eu/content/soil-map-slovakia-podne-typy`.

[26] Henriette I Jager and Mark S Bevelhimer. “How run-of-river operation affects hydropower generation and value”. In: *Environmental Management* 40 (2007), pp. 1004–1015.

[27] Georges Kariniotakis. *Renewable energy forecasting: from models to applications*. Woodhead Publishing, 2017.

[28] Kyoung-jae Kim. “Financial time series forecasting using support vector machines”. In: *Neurocomputing* 55.1-2 (2003), pp. 307–319.

[29] Özgür Kişi. “Streamflow forecasting using different artificial neural network algorithms”. In: *Journal of Hydrologic Engineering* 12.5 (2007), pp. 532–539.

[30] Alban Kuriqi et al. “Influence of hydrologically based environmental flow methods on flow alteration and energy production in a run-of-river hydropower plant”. In: *Journal of Cleaner Production* 232 (2019), pp. 1028–1042.

[31] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *nature* 521.7553 (2015), pp. 436–444.

[32] Ray K Linsley Jr, Max Adam Kohler, and Joseph LH Paulhus. “Hydrology for engineers”. In: (1975).

[33] Woodbank Communications Ltd. *Demand load Curve*. [accessed May 03, 2023]. 2005. URL: `https://www.mpoweruk.com/electricity_demand.htm`.

[34] Francesco Lässig. *Temporal Convolutional Network*. [accessed May 05, 2023]. URL: `https://unit8.com/resources/temporal-convolutional-networks-and-forecasting/`.

[35] Google Maps. *MVE Želiezovce*. [accessed April 18, 2023]. 2023. URL: `https://www.google.com/maps/search/elektr%C3%A1re%C5%88+bl%C3%ADzko+%C5%BDeliezovce,+Slovensko/@48.0677038,18.6669806,17z/data=!3m1!4b`.

[36] Informační centrum ČKAIT s. r. o. *Vodní dílo Želiezovce*. [accessed May 04, 2023]. 2017. URL: `https://www.casopisstavebnictvi.cz/clanky-vodni-dilo-zeliezovce.html`.

[37] Met Office. *The water cycle*. [accessed May 20, 2023]. 2023. URL: `https://www.metoffice.gov.uk/weather/learn-about/weather/how-weather-works/water-cycle`.

[38] Folakemi Ope Olabiwonnu, Tor Haakon Bakken, and Bokolo Anthony Jr. “Achieving sustainable low flow using hydropower reservoir for ecological water management in Glomma River Norway”. In: *Sustainable Water Resources Management* 8.2 (2022), p. 53.

[39] Min Qi and Guoqiang Peter Zhang. "An investigation of model selection criteria for neural network time series forecasting". In: *European journal of operational research* 132.3 (2001), pp. 666–680.

[40] Hira Singh Sachdev, Ashok Kumar Akella, and Niranjan Kumar. "Analysis and evaluation of small hydropower plants: A bibliographical survey". In: *Renewable and Sustainable Energy Reviews* 51 (2015), pp. 1013–1022.

[41] Francis EH Tay and LJ Cao. "Modified support vector machines in financial time series forecasting". In: *Neurocomputing* 48.1-4 (2002), pp. 847–861.

[42] Unit8. *Darts*. 2015. URL: `https://unit8co.github.io/darts/`.

[43] Berlin Wu. "Model-free forecasting for nonlinear time series (with application to exchange rates)". In: *Computational Statistics & Data Analysis* 19.4 (1995), pp. 433–459.

[44] Deng Yongsheng et al. "A short-term power output forecasting model based on correlation analysis and ELM-LSTM for distributed PV system". In: *Journal of Electrical and Computer Engineering* 2020 (2020), pp. 1–10.

[45] G Peter Zhang and Min Qi. "Neural network forecasting for seasonal and trend time series". In: *European journal of operational research* 160.2 (2005), pp. 501–514.

[46] Anna Zygierewicz and Lucia Sans. "Renewable Energy Directive: Revision of Directive (EU) 2018/2001". In: *European Parliamentary Research Service, nd* 10 (2021).

Appendix A

Implementation

In this section we present the implementation details and notes of concepts that were discussed in the previous chapters.

The implementation was divided into two parts: tasks and reports. All performed experiments can be found in the '/reports' file.