Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

Sakke Rantala

Assessing the Quality of Hydro-power Cascade Operation

Master's thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology in the Degree Programme in Engineering Physics and Mathematics.

Rovaniemi, May 30, 2016

Supervisor: Assistant Professor, PhD, Pauliina Ilmonen

Advisors: M.Sc. Juho Päiväniemi and M.Sc. (Eng) Hannu Korva

Aalto University
School of Science

Abstract of master's thesis

Author	**Sakke Rantala**		
Title of thesis	**Assessing the Quality of Hydro-power Cascade Operation**		
Master's programme	**Degree Programme of Engineering Physics and Mathematics**		
Thesis supervisor	**Assistant Professor Pauliina Ilmonen**		
Major or Minor/Code	**Systems and Operations Research / F3008**		
Department	**Mathematics and Systems Analysis**		
Thesis advisor(s)	**Juho Päiväniemi, M.Sc. and Hannu Korva, M.Sc. (Eng)**		
Date	**30.5.2016**	**Number of pages VI + 61**	**Language English**

Abstract

This thesis studies applicability of certain mathematical methods in the assessment of the quality of operation of a hydro-power cascade. The quality is approached from the aspects of maximization of the benefits and the assessment of the behaviour of the river.

The thesis proposes two new approaches that can be utilized in the consideration of the ex post optimality of the production. The first one is an ex post optimization methodology that is based on a mathematical river model. Using the methodology, the efficiency of the actual outcome can be assessed with the real constraints but without the uncertainties of the real decision-making situation. The methodology does not pay attention to the economic value of the different components of the production or to the production as a part of a larger portfolio. The methodology is demonstrated with a case study. The other approach is the operative buffer that depicts the water level of a reservoir as a time-based buffer with respect to the water flowing into the reservoir. In addition, with the aid of the buffer, one can determine and unify the risk levels regarding to the water management.

The study of the behaviour of the river is based on the assumption that the behaviour should be predictable. Therefore, the thesis focuses on the centrality problem of the curves, or, functional data on a general level. The selected tool is a field of mathematical statistics, Functional Data Analysis (FDA), which is a subject of growing interest. In the thesis, certain functional depth functions are applied in the assessment of typicality of multivariate functional data. We propose a new approach, Pareto-efficient depth, to define the most typical observations. It combines the Pareto-efficiency, robust statistical measures and FDA in a practical way. For the data of this thesis, the new approach seemed more suitable than the other depth functions, and it could be an interesting subject of further studies.

In all, the applicability of the methods discussed can be considered reasonable as long as their weaknesses are taken into account. The functional approach to the process quality assessment is justified, even though unambiguous definitions of the best or of the typical could not be agreed upon. In addition, monitoring the efficiency forms an incentive to continuously improve both the performance of the operation and the optimization model.

Keywords ex-post optimum, multicriteria decision-making, hydro-power, functional data-analysis, depth function, Pareto-depth

Aalto-yliopisto
Perustieteiden
korkeakoulu

Diplomityön tiivistelmä

Tekijä	**Sakke Rantala**		
Työn nimi	**Vesivoimaketjun operoinnin laadun arviointi**		
Koulutusohjelma	**Teknillinen fysiikka ja matematiikka**		
Valvoja	**Professori Pauliina Ilmonen**		
Pää- tai sivuaine/koodi	**Systeemi- ja operaatiotutkimus / F3008**		
Työn ohjaajat	**FM Juho Päiväniemi ja DI Hannu Korva**		
Päivämäärä	**30.5.2016**	**Sivumäärä VI + 61**	**Kieli Englanti**

Tiivistelmä

Tässä diplomityössä tutkitaan eräiden matemaattisten menetelmien soveltuvuutta vesivoimaketjun operoinnin laadunarviointiin. Laatua tarkastellaan sekä tuotannon hyötyjen että joen käyttäytymisen näkökulmasta.

Hyötyjen arviointiin esitetään kaksi menetelmää, joiden avulla tuotannon tehokkuutta voidaan arvioida jälkikäteen ilman todellisen päätöksentekotilanteen epävarmuuksia. Ensimmäinen on matemaattiseen jokimalliin perustuva optimointimenetelmä, jolla arvioidaan toteutuneen tuotannon tehokkuutta todellisen päätöksentekotilanteen reunaehdoilla. Menetelmän toimivuutta tarkastellaan käytännön esimerkin avulla. Menetelmä ei huomioi tuotannon eri komponenttien taloudellista arvoa eikä sen kokonaisarvoa osana laajempaa portfoliota. Toinen menetelmä on operatiivinen varaumapuskuri, joka kuvaa laitosaltaan pinnankorkeuden aikamääräisenä puskurina suhteessa altaaseen tulevaan vesimäärään. Puskurin avulla voi lisäksi määrittää ja yhtenäistää vedenhallintaan liittyviä riskitasoja.

Joen käyttäytymisen osalta työssä oletetaan, että sen ennakoitavuus on tärkeää muille joen käyttäjille. Siksi työssä paneudutaan menetelmiin, joiden avulla voidaan määrittää käyrämuotoisen, tai funktionaalisen, datajoukon tyypillisimmät havainnot. Arvioinnin työkaluksi työssä valittiin matemaattisen tilastotieteen osa-alue funktionaalinen data-analyysi. Työssä sovelletaan sen tunnettuja menetelmiä tyypillisen havainnon määrittämiseen moniulotteisesta funktionaalisesta otoksesta. Työssä esitellään lisäksi uusi lähestymistapa löytää tyypillisimmät havainnot. Menetelmä yhdistää Pareto-tehokkuuden, robustin tilastotieteen ja funktionaalisen data-analyysin käytännönläheisellä tavalla. Uusi lähestymistapa osoittautui tämän työn aineistoissa toimivammaksi kuin muut esitetyt menetelmät, ja menetelmä voisi olla mielenkiintoinen jatkotutkimuksen aihe.

Kaiken kaikkiaan esitettyjen menetelmien toimivuus arvioidaan kohtuulliseksi, kunhan niiden heikkoudet tunnistetaan. Prosessin laadunarvioinnissa funktionaalinen lähestymistapa on perusteltu, vaikka yksiselitteisiä määritelmiä parhaalle tai tyypilliselle havainnolle ei löytyisikään. Lisäksi tehokkuuden tarkkailu kannustaa sekä operoinnin että joki- ja optimointimallin jatkuvaan kehittämiseen.

Avainsanat ex-post optimi, monitavoiteoptimointi, vesivoima, funktionaalinen data-analyysi

To Lilian and Lukas.

Acknowledgements

I want to express my gratitude to Kemijoki Oy for the opportunity to work and study in the interesting environment. Especially I want to thank Juho, Hannu, Erkki, Ismo and all the operators for the comprehensive discussions during the process.

The thesis adviser Pauliina introduced me into the worlds of robust statistics and Functional Data Analysis. Your personality and energy will go on inspiring me. I want to thank you for your professional guidance through the last courses of the degree and, most importantly, with the thesis.

The credit for the idea of the ex post methodology belongs to the team which was formed by many experts representing the shareholders of Kemijoki Oy. In addition, I want to thank Anssi for the guidance into the scientific discourse related to the concept of *ex post*.

I want to thank all the great persons I had an opportunity to get to know during my years in the university. Especially, I would like to thank you Santtu, Oskari and Aleksi for these years. Learning with you and from you has been momentous for me.

I want to thank both my and Elina's families for their support in the different phases of our life.

My life has changed a lot during the five years of studies. I have learned a lot in the university but, I guess, even more at home. Lilian and Lukas, I want to go on growing up with you.

Above all, I want to thank you, Elina, for your endless love and support.

Rovaniemi, May 30, 2016

Sakke Rantala

Special Terms

bagplot A bivariate extension of *boxplot*. Visualizes distribution of the data data. 37

behaviour of the river All the dynamic things that can be observed from the river bank, including the water levels, the flows and their changes. 5, 30

convex Any two points of a *convex set* can be joined with a line whose all points belong to the set. A function is called convex if its epigraph forms a convex set. 8

depth Measure of centrality in multi-dimensional setting. 35, 40

ex post Based on forecasts rather than actual results, opposed to *ex ante*. [Oxford Dictionaries, 2016]. 1, 13, 56

Functional Data Analysis FDA, A field of mathematical statistics where functions, not only points, are seen as observations. 5, 31

Kemijoki Oy (KEJO) Hydro-power company in Finland. This thesis was done for this company. 2

Key Performance Indicator KPI, an indicator for measuring the performance of an organization in its pursue to reach its key target(s). 1

Linear Programming Algorithm that solves linearly formed optimization problems. Linear mapping $L(\cdot)$ satisfies the following: $L(ax + by) = aL(x) + bL(y)$. 12, 14

Pareto-optimality A state of multi-objective allocation of resources where one criteria cannot be improved without making another worse. 9, 33, 46, 56

Contents

1 Introduction

1.1 Problem statement

This thesis studies different mathematical methods to assess the quality of operation of multiple connected hydro-power plants, or, a cascade. The quality of operation refers to the level of excellence of the operational actions conducted under a wide range of uncertainties regarding to the future. Excellence is defined here rather vaguely as *responsibly optimal* and its evaluation in this thesis is restricted to the two following aspects. The first objective is to maximize the benefits gained from the production, and the second one is to maintain the river behaviour predictable. The former will focus on the short-term optimization problem and especially on the ex post optimal efficiency of the dispatch problem. The methodology proposed extends the current Key Performance Indicators. The latter discusses, on a general level, the centrality problem of multivariate curves, or functional observations in the context of the hydro-power cascade.

The definition of the quality of operation may be subjective, depending on the interests of the person. The change of quality, as a concept, instinctively incorporates juxtaposition of entities that should be comparable. In this thesis, the reference entity is the relevant history. Such an entity could be, for example, another cascade, which would result in a comprehensive benchmarking study. The thesis focuses on mathematically interesting and conceptually relevant measures. Intentionally, some aspects of the quality of operation are addressed only shortly or omitted if they relate more to other disciplines. For example, the quality of communication between different parties during an electricity network issue is essential when it comes to the safety issues but still is excluded from the scope of the thesis. Even though the thesis discusses certain quantitative measures and their properties, a qualitative motivation plus critique is provided for each one.

The thesis consists of two thematically independent main sections. Section 2 will study the assessment of the optimal operation both theoretically and empirically, whereas Section 3 focuses on the general question: how to define "normal" among multivariate curves?

1.2 Background and motivation

The hydro-power company Kemijoki Oy (later "KEJO") owns 20 hydro-power plants of which 16 are located in the Kemijoki river. The nominal power capacity is more than 1100 MW, and the energy production of the year 2015 ecxeeded 5 TWh. The stock of KEJO is divided into the capital and the energy shares. The owners of the energy shares obtain all the energy produced. KEJO manages the production abilities and produces the energy efficiently and responsibly for the shareholders. [Kemijoki Oy, 2016]

The shareholders are independent market players and mutual competitors, whereas KEJO is an impartial company between them. One of the shareholders is the *balance responsible party*[1] of the asset. The other shareholders order their share of energy from the balance responsible party according to certain governing rules.

The Kemijoki river system is illustrated in Figure 1, where circles correspond to the regulated reservoirs, boxes to the plants and the connecting lines to the main course of the river. The regulated lake in the middle divides the system into two controllable sections. It takes about 12 hours for water to pass through the first section and about 30 – 40 hours through the other. There are small reservoirs before each plant, which enable flexible production. The total production capacity exceeds 1000 MW and the daily changes pose a valuable balancing potential, as demonstrated in Figure 2. In all, the river is a dynamic system and its real time operation depends on – and is restricted by – the past decisions and is highly driven by the future expectations.

Earlier, the river was physically operated by KEJO. In 2014, the owners decided to transfer also the operation to the balance responsible party. The aim of the transfer was to improve the efficiency of the physical operation and the fluency of trading. During the transfer process, the questions related to the quality became urgent for two main reasons. Firstly, the other shareholders wanted to ensure that the quality would improve as promised. Secondly, as the legal responsibilities and the license to operate the river will remain at KEJO, it is essential for it to manage the quality improvement process and

[1] "An electricity market party which has a valid balance service agreement with Fingrid. [...]" Fingrid is the *transmission system operator* (TSO) in Finland. It owns the high-voltage grid and takes care of cross-border connections. For more information, see `http://www.fingrid.fi/en/`.

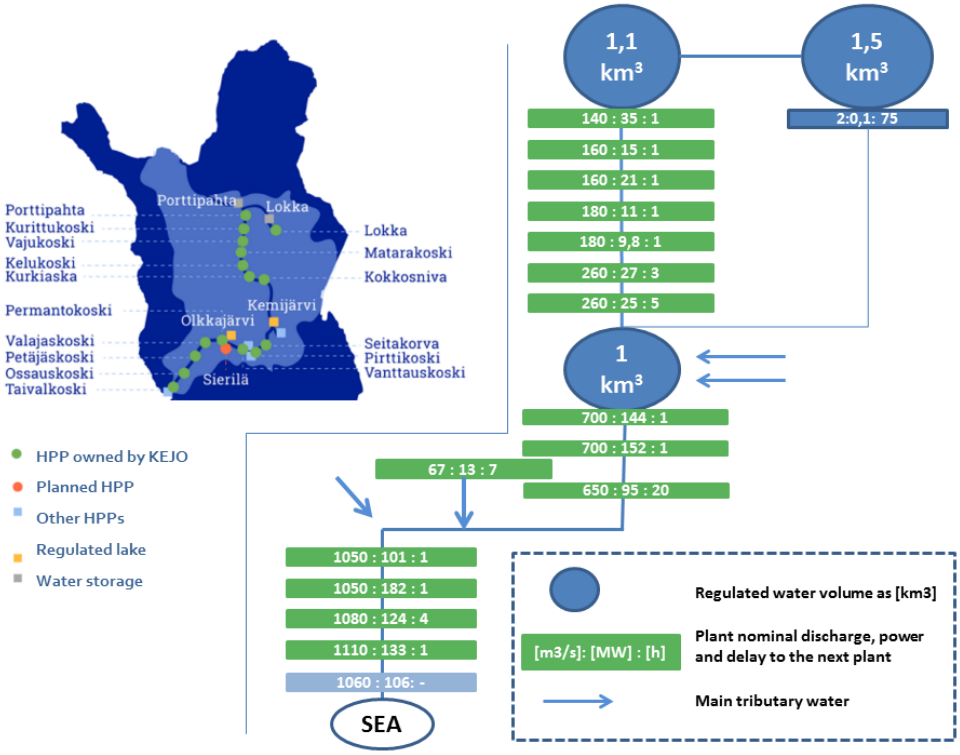

Figure 1: The map and the schematic picture of the Kemijoki river system. The source of the map and the information: Kemijoki Oy

take into account all the relevant stakeholders and users of the river in the long term. The new operator, as a shareholder, has a strong incentive to achieve the same objectives in both the short and the long term.

The operation of hydro-power is one part of the water resource management which includes assessment of contradicting interests, effects on the environment and legal issues, to name a few. There are advantages and disadvantages to water resource management in general. Äijö et al. [1992] provides a profound view on those questions. For example, Siivola [1992] discusses questions related to recreational use. She summarizes that advantages of recreational usability are not easy to valuate, but it often prevents different industries from gaining more additional value. Additionally, Kovanen [1992] summarizes the aspect of power engineering and concludes that hydro-power production affects the environment only locally, whereas, for example, coal

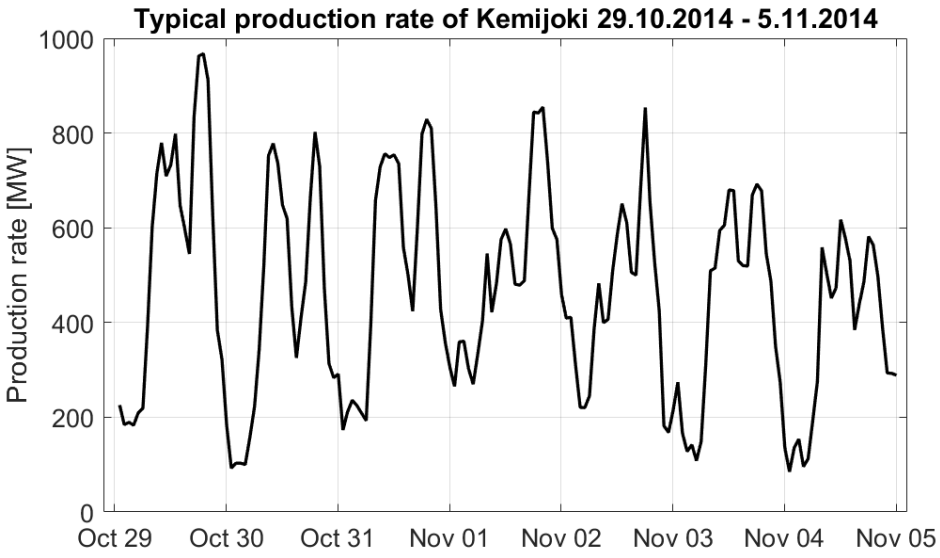

Figure 2: Typical production rate of the river Kemijoki during a week. Source: Kemijoki Oy

and peat affects the environment when mined, transported, and burned. In all, utilizing one resource for many purposes is a multicriteria problem.

The multicriteria decision analysis aims to holistically evaluate contradicting decision alternatives. It structures the problem and elicits the values of different objectives within an interest group and between different groups. Related to the water resource management, Marttunen [2011] presents a recent view on the subject. He proposes a methodology called Decision Analysis Interview that is implemented in five real cases in Finland. The cases consist mainly of the lake regulation problems that have a significant impact on a large group of different stakeholders and the environment. The decision analysis interview methodology takes into consideration the contradicting parties. Along with the methodology and the case studies, one of the key results is that the level of interaction during the process correlates highly with the integration of the analysis. Namely, if only a few people collaborate, the analysis probably does not cause the desired impact in the real world. To assess such an impact, Palonen [2006] presents a real-world case where the regulation rules were adjusted in order to better comply with the recreational use. Using surveys and pricing trends of the real properties on the water front, the

local impact of the changes were assessed as positive. However, not much attention was paid to the resulting additional or lost value of the hydro-power players. From the same point of view, Nurmi [2010] considers the effects of fluctuation of water levels on the environment and also general questions related to indicators. However, the daily fluctuation or changes in discharges are not covered.

Section 2 discusses the aspect of optimality. The holistic evaluation of optimality proves difficult even afterwards because it lacks the counter-factual outcome. In a real decision-making situation, there are many uncertainties regarding the future. An *ex post optimization* methodology is proposed in this thesis. It can be utilized in the assessment of the efficiency of the historical outcome, without the uncertainties. The optimality of the possible portfolios, or, of the bids of the owners are excluded from the scope of this thesis.

Section 3 focuses on the statistical study of curves, Functional Data Analysis. The high-level aim is to define the typical behaviour of the river. The behaviour includes matters that can be observed from the river bank, for example water levels, flows, and their changes. For an operator, it is important to comprehend the behaviour as the operation affects the environment and the utility of the river to its other users. The utility of this approach is to develop the operation so that undesirable effects could be avoided while maintaining an optimal production. However, this thesis mainly discusses the problem of normality among curves on a general level. The examples base on the contextual data.

2 Assessing optimality of the production

2.1 Optimization of hydro-power production

2.1.1 Physics of hydro-power production

The underlying source of energy for hydro-power is the sun. It enables the ongoing process of water gaining potential energy while rising up from lower levels to the upper soil. The hydro-power production transforms the potential energy of water into a relevant form of energy, in practice, electricity. The potential difference is captured at one location by constructing dams on rivers, and the hydro-power plant is often integrated into a dam. The main components of a hydro-power plant are depicted in Figure 3. The turbine transforms the kinetic energy of the flowing water, Q, into rotational energy which the generator eventually transforms into the electricity. The potential energy difference is determined by the geodetic head, H_g, that is the difference between the water levels upstream and downstream from the plant. The geodetic head is subtracted by the losses in the intake screen and in the waterways before and after the turbine. The resulting potential difference is called the net head, h. For a more comprehensive explanation of the physics of the hydro-power production, see for example INSKO [1978].

The turbine itself has an efficiency that depends on the net head and the flow, $\eta(h, Q)$. Common turbine types include Kaplan, Francis, and Pelton, each possessing different properties. In this river, the geodetic heads are quite low, and therefore, the Kaplan-type turbine is the most suitable ([INSKO, 1978, pp. 93 – 94]). Typically, the efficiency of a Kaplan-turbine is approximately 90 – 95 %. An example is illustrated in Figure 4, and it is possible to deduce that the efficiency, η, of the Kaplan turbine is quite robust against changes in the discharge, Q, or in the net head, h. The axial power of the turbine is then the net potential energy difference subtracted by the turbine losses. It is shown in the equation (1), where ρ denotes the density of the flowing water. The generator has losses of approximately $1 \ldots 3\%$ which are reduced from the axial power before the electricity grid.

$$P(Q, h) = g \cdot h \cdot Q \cdot \eta(h, Q) \cdot \rho \tag{1}$$

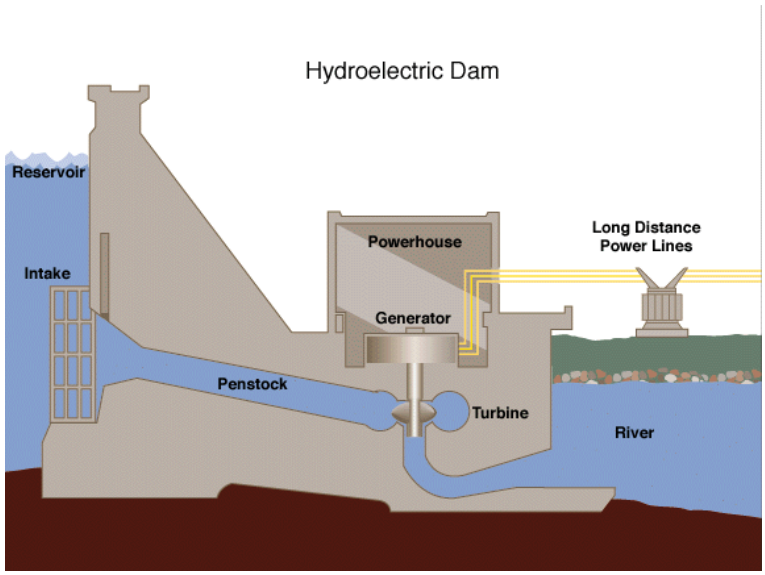

Figure 3: The main components of a hydro-power plant. *The figure is published under the Creative Commons Attribution 2.5 Generic license by Tomia (Own work)* `https://commons.wikimedia.org/w/index.php?curid=3302749`

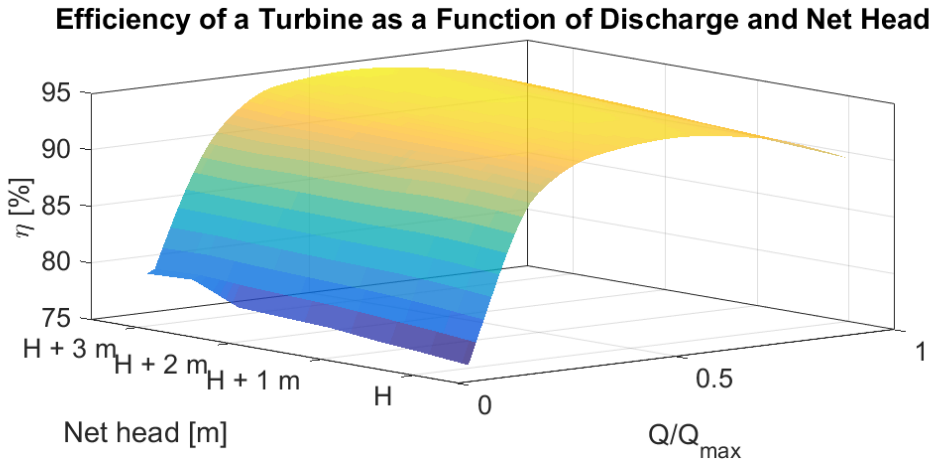

Figure 4: An example of a Kaplan turbine efficiency as a function of net head and discharge ratio. Source: Kemijoki Oy

In the river cascade illustrated in Figure 1, the plants are physically connected by the river course[2] as the water flows from the upper plant(s) to the lower ones. The connection system plays a crucial role in the operation, because the regulating volumes in almost all the reservoirs are so small that they could properly operate only a few hours without feeding water from the upper plants. The synchronous operation is important also because the empty reservoirs reduces the potential energy difference. To give a concrete example, an average decline of 5 cm in all the reservoir levels implies an annual loss of about 10 GWh.

2.1.2 Optimization in general

Optimization means actions of "making the best or most effective use of a situation or of a resource" [Oxford Dictionaries, 2016]. The phases of optimization could be structured as setting the objective, becoming aware of relevant boundaries and then determining the best alternative among the feasible solutions. If the two first phases can be stated mathematically, the last phase is a mathematical optimization problem. The mathematical optimization finds the solution (if any) to the *modelled* problem, and the solution is optimal with respect to that very model. The solution is not necessarily real-world optimum due to the erroneous model. Depending on the mathematical model, the optimization problem can be grouped for instance into linear or non-linear, convex or non-convex and integer problems.

The general optimization problem is stated in the equation 2, where $\mathbf{x}$ is a decision variable and $\mathcal{S}$ defines its feasible region. The minimization of an objective function can be turned into maximizing by changing the sign. If there is only one objective function ($m = 1$), then the result of the optimization is called an optimal solution. If there are many conflicting objectives ($m > 1$), the process is called multi-objective optimization, and its result is a set of Pareto-optimal solutions. [Miettinen, 2008, p. 2]

$$\min_{\mathbf{x} \in \mathcal{S}} \{f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_m(\mathbf{x})\} \tag{2}$$

[2]Another form of physical connection is the electricity grid. It is essential when operating, but it is excluded from the scope of the thesis.

Pareto-optimality is defined in Definition 2.1. All the members of the Pareto-set are mathematically equally appropriate, also known as *non-dominated* or *Pareto-efficient* solutions. If the set consists of more than one member, the last phase of the optimization procedure is to find the most preferred solution among the non-dominated ones. A prerequisite for decision-making is that the decision-makers are aware of the feasible alternatives. The problem of presenting the solutions to them in understandable way is discussed more for instance in Korhonen and Wallenius [2008]. A visualization of the Pareto-optimal solutions in the bivariate case is given in Figure 5. The sub-figure in the bottom-right corner demonstrates that sometimes the results may be intuitively surprising as quite distant solutions prove to be non-dominated. If the edge of the heart-like set were only a bit further away from the centre, the edge would be the only non-dominated solution.

Definition 2.1. *Pareto-optimality:*
Let $\mathcal{X} \in \mathbb{R}^{k \times n}$ denote domain, k-vector $\mathbf{x} \in \mathcal{X}$ observation and $\mathbf{X}_{PO} \subset \mathcal{X}$ the set of Pareto-optimal observations. Then, $\mathbf{x}^* \in \mathbf{X}_{PO}$, if $\nexists\, x \in X$ such that

$$f_i(x) \geq f_i(x^*)\, \forall\, i \,=\, [1, \ldots, m] \tag{3}$$

$$\exists\, j = [1, \ldots, m] \;\; s.t. \;\; f_j(x) > f_j(x^*), \tag{4}$$

where m is the number of objective functions $f_i(\cdot)$. If only the first condition (3) holds, $\mathbf{x}^*$ is called weakly Pareto-optimal.

The multi-objective problem can be transformed into single objective problem by mapping all the objectives into one. However, the objectives are not necessarily commensurable. For example, one might prefer two star hotel for €150 to one star hotel for €80 but prefer four star hotel for €250 to five star hotel for €290. In other words, the stars cannot be weighted consistently with money. So, it is difficult to determine the trade-offs between different alternatives before the alternatives are known. Moreover, the *a priori* preferences might not correspond to the solution that the decision-maker would have selected from the non-dominated solutions. Furthermore, the feasible region may diminish if the constraints on the objectives are set before the alternatives are known. [Branken et al., 2008, Preface]

In practice, however, the preferences are often determined a priori. Then the multi-objective optimization problem is reduced to single-objective by minimizing the weighted sum of normalized objective functions, f_i^* (eq. 5).

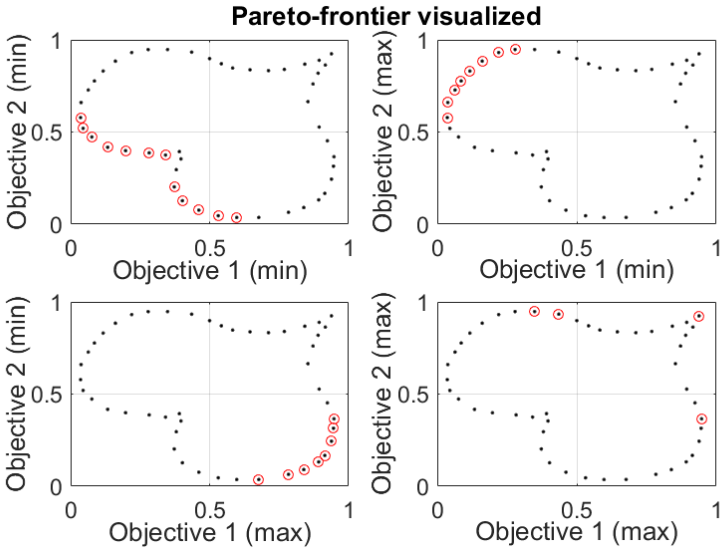

Figure 5: Pareto-optimal solutions with different objective combinations shown on a bivariate non-convex set. Pareto-optimal solutions are mathematically equally appropriate.

Its result is at least weakly Pareto-optimal[3]. There are serious drawbacks in this method. Firstly, all the Pareto-efficient solutions are not necessarily found in non-convex problems. Secondly, it is difficult to normalize the weight functions and to determine globally relevant weights. Moreover, in multi-objective linear problems, the weights that result in a non-dominated solution are not unique. As a consequence of that, very different weights can produce similar Pareto-optimal solutions, or slightly different weights may produce remarkably different solutions. [Miettinen, 2008, pp. 8–13]

$$\min_{\mathbf{x}\in\mathcal{S}} \sum_{i=1}^{m} w_i f_i^*(\mathbf{x}),\ w_i \geq 0 \,\forall\, i \tag{5}$$

[3]The result is Pareto-optimal if $w_i > 0 \,\forall\, i$. For proofs, see e.g. [Miettinen, 2012, pp. 78 – 79]

2.1.3 Hydro-power optimization

The objective of hydro-power optimization is to maximize the sum of the net present values of different future value streams. Each value stream depends on time and consists of the production-related quantity and the corresponding unit price or cost. The most important value stream of the hydro-power is the flexible energy production whose value is determined in the Nordic energy market at different time-scales. The other value streams of the hydro-power are the balancing power, the imbalance power, the frequency controlled reserves and its value as a part of a production portfolio. For more information about the free electricity markets in the Nordic countries, see for example [Kerttula, 2012, Chapter 2] or the web pages of the market place Nord Pool and the transmission system operators[4].

Let k-vector $\boldsymbol{\mathcal{X}}(t)$ denote the magnitude of the production related quantity, k-vector $\boldsymbol{\beta}(t)$ the corresponding unit value or cost, and r the discount rate. Then, the objective can be stated very generally as in the equation 6. The elements of $\boldsymbol{\mathcal{X}}(t)$ are not independent as, for example, selling the frequency controlled reserves reduces the capacity of regulating power. In other words, one has to dynamically allocate the resource into many markets.

$$\max_{\mathcal{X}\in\mathcal{S}} \int_{t=0}^{\infty} e^{-rt} \boldsymbol{\mathcal{X}}(t)^T \cdot \boldsymbol{\beta}(t) dt \tag{6}$$

One can control only restricted amount of the factors related to the different elements of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\beta}$. For example, forecasting of the prices includes forecasting of both the total consumption and the aggregated production that again depend, inter alia, on the weather and on both the economic and the political circumstances, at least on the long term. Moreover, in the efficient market, the magnitude and the price are theoretically interdependent; the more is supplied the less is the price. In all, this formulation is not very valuable from practical point of view but still forms the backbone of the physical procedures in reality. In practice, the time horizon does not reach infinity as the discount factor diminishes the weight of the distant future. The time horizon can be divided into different lengths. The strategic horizon

[4]At the time of the writing the web-page of the market place can be found at `http://www.nordpoolspot.com`. For the web-pages of the Finnish TSO Fingrid, see `http://www.fingrid.fi/EN/`.

is counted in decades, the tactical horizon in months or years and the shortest time-span is days or hours, and it is the actual operation. This thesis focuses on the short-term problem.

The feasible region, $\mathcal{S}$, takes into account the dynamics of the cascade, all the operational restrictions, such as the water regulation limits of the reservoirs, and all the technical restrictions, including the maximum production outputs of the generating units. The tactical objectives, such as the volume of the water in the regulated lakes, must be respected in the short-term optimization. Such an objective can be implemented, for instance, as one value stream component by applying negative prices for the deviation from the objective.

Harpman [2012] separates the short-term problem into the dynamic economic dispatch problem and the unit dispatch problem. The unit dispatch problem is computationally complex binary problem where the operating units are set on or off. However, in the end it is actually tightly related to the economic problem. In general, the short-term problem belongs to the class of non-linear constrained optimization problems. Non-linearities arise from the generation and head relationships and from delay models.

The problem can be solved in many ways, depending on the model designed. One common way is to model the problem as a linear model that can be solved using an algorithm called Linear Programming. Examples of it are presented, for instance, in Zheng et al. [2013] and Kerttula [2012]. A linear model can be solved also by the control-theory approach. An example of a linear state-space model is shown in Pursimo and Lautala [1993][5]. As the phenomenon introduces stochastic characteristics, such as the tributary waters and the prices in the future, the stochastic optimization could be considered. A long-term stochastic optimization application is provided, for instance, by Flach et al. [2010].

Another family of solving methodologies are heuristic methods that – in theory – do not set limits on the modelling level. Thus the model can be non-linear and non-convex. However, their performance cannot be guaranteed as the solution found is not necessarily the optimum. On the other hand, if the method does not find any solution, it does not prove that such would

[5]The model is designed for KEJO and is as such worth mentioning here, even though the formulation of the optimization problem does not reflect the current needs.

not exist. However, for example [Zanakis and Evans, 1981, pp.85–86] gives many reasons for using the heuristic methods. The reasons are that they are easier to implement and the model might be more relevant. They claim that "a fast near-optimal solution makes more sense than a time-consuming exact solution to an inexact problem." In the context of hydro-power optimization, Harpman [2012] discusses for example Particle Swarm Optimization and Differential Evolution.

2.1.4 Ex post optimization of the hydro-power production

An *ex ante* choice refers to a decision made under uncertainties and expectations whereas *ex post* is looking backwards to the occurred outcome. In reality, the hydro-power production is optimized ex ante under expectations regarding to the different markets and hydrological factors. Starr [1973] discusses the optimality of production and allocation under uncertainty in terms of general economy, but the theme can be applied to electricity markets as well. He shows that ex ante optimal allocation of resources over time implies an ex post optimal outcome if and only if the subjective probability distributions of the future states are identical. If the universal similarity of the distributions does not hold, then the result will not be ex post optimal. The reason is that if the subjective probabilities are different ex ante, their marginal rates of substitution are different ex post. He shows that a necessary condition for Pareto optimum is the equality of those rates. Harris [1978] discusses the results of that article and proposes that the conflict between the ex ante and ex post optima is caused by "changing tastes". Also, Conrad and Unger [1987] develops different ex post performance tests on both the short and the long term optimization, based on the historical data of real industries. The figure 6 illustrates the distributions of prices of some relevant electricity markets in Finland. All the distributions are skewed, of which the balancing power market is an extreme example, as the maximum prices exceeded 2000 €/MWh in that year. The volume of the energy market (Spot) dominates the other two.

Zheng et al. [2013] proposes "the post-evaluation" of hydro-power production based on the actual operation data. The setup is the Three Gorges dam in China with its 34 generating units and the two surrounding plants. The idea is to develop "potential hydro-power output" that is the difference of

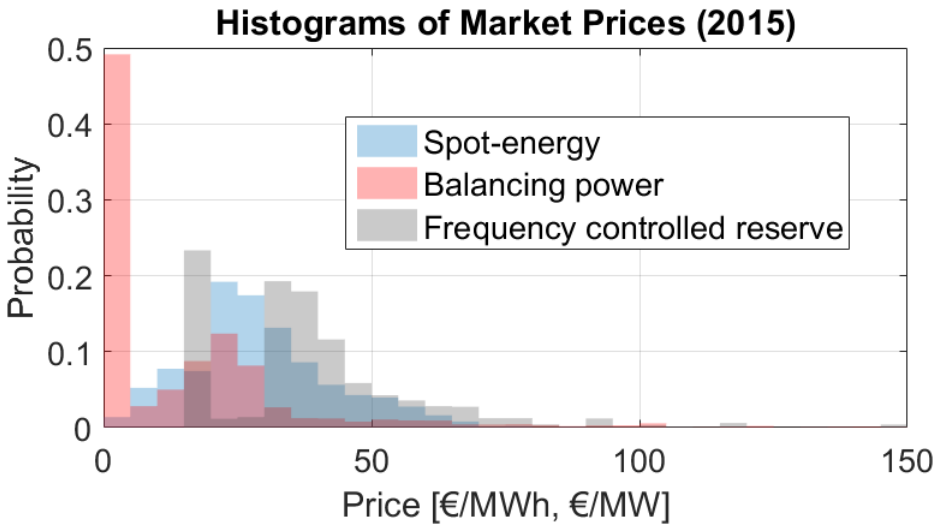

Figure 6: Distributions of hourly prices of certain relevant market products in 2015. The volumes of the markets differ significantly from each other. The units of frequency controlled reserve (FCR-N) is €/MW but the others are €/MWh. For visual reasons, prices over €150/unit are stacked at 150.

the actual and the afterwards, or ex post, optimized results. He proposes a dual methodology consisting of "the integrated evaluation model" and "the sensitivity analysis". The former takes into account the river as a whole, whereas the latter studies the sensitivity of power outputs to the expected external inflow to reservoirs or controlled flood level. The river is modelled linearly and problem is solved using Linear Programming for a period of one year at a time. The results propose that the energy gained is $3-10$ % ($1-6$ TWh) less than the potential maximum, depending on the hydrological year. The results of the sensitivity analysis are as significant. For example, if the forecast accuracy factor (defined on the page 1170 of the article) increased from 0.80 to 0.95, the output was 3.8% larger than the actual output. The results of the flood control limits are also of that magnitude.

The methodology proposed by Zheng et al. [2013] is not directly applicable to the Kemijoki river. Kemijoki consists of 16 remarkably smaller plants and therefore the unit dispatch problem does not play that significant role in the optimization. Furthermore, the dynamics of the water flowing in the long cascade differs significantly from the short cascade. In the following, an ex

post optimization methodology is defined for longer cascades to determine the optimality of the actual operation.[6] For a case example, see Section 2.1.6.

First, the measured hourly turbine-specific discharges, *the schedule*, is simulated using a river model. Next, the schedule is optimized with respect to all real operational constraints that includes the realistic decision space. The objective of the optimization is to minimize the amount of water value expended. The water in the reservoirs, and in between them, have value based on the production potential of the rest of the river. The difference between the simulated and optimized results in the model world is *assumed* to reflect the difference in reality as well. This deviation from the optimal efficiency is then the quality metrics whose development over time could be studied. To sum up, the procedure consists of the following phases:

1. Choose the optimization period from the history
2. Simulate the model with the actual (measured) discharges, including sufficient length of data before and after the horizon. Adjust the model world to correspond the real world in the beginning of the optimization horizon.
3. Minimize the use of resources during the ex post horizon subject to the real operational constraints:
    - The total simulated production of the cascade must be satisfied.
    - The total amount of water discharged from the strategic reservoirs in reality must remain the same during the optimization horizon.
    - Only such turbines that were available in reality can be used in the optimization. Likewise, the turbines that must not be stopped cannot be stopped in the optimization.
    - The tributary water flows into the reservoirs as observed.
    - The sold spinning reserves must be maintained
    - The water discharged before the beginning of the optimization horizon is taken into account during the horizon, thus reflecting

[6]Similar methodology was applied *ad-hoc* during the transfer process of operation KEJO in 2013. The credit for the idea belongs to the team formed by experts representing the shareholders of the company.

the real decision-making situation.

- The optimized schedule should enable the production after the horizon ends.
- The optimized schedule must respect all the real environmental constraints, such as the levels of the reservoirs and the changes of flows.

4. Comparison of the simulated and the optimized results.

The methodology remains the same if the ex post optimization is conducted through every time point. Then the period of the first phase is a moving window and the other phases remain the same for each case.

2.1.5 The river and production model applied

The ex post optimization can be implemented with any sufficiently good river model. To provide a concrete example, this section shortly presents a river model suitable for that purpose.

The modelled period consists of timesteps $t_k, k = 1, \ldots, N$ with even discretation, or $d = t_k + 1 - t_k$ is constant for all $k = 1, \ldots, N-1$. Then, the length of the period is $L = N \cdot d$. The most important sub-models are the plant interconnection model, the tail-water model and the generating unit model.

There are I hydro-power plants in the cascade and the water from the plant i flows to the single plant $Next(i)$. Incoming water to the plant i may origin from several plants. Also, tributary water from the catchment area flows into the reservoirs, adding stochasticity to the system. The water volumes in the reservoirs are transformed into water levels using a reservoir function $C_i(\cdot)$. The water flow from i arrives to the next plant $Next(i)$ faster if the river course is full of water, because the cross-sectional area of the stream increases. The arriving flow is depicted by the delay distribution $D_i(t_k) = \lambda_i(t_k) \cdot P_i^F + (1 - \lambda_i(t_k)) \cdot P_i^E$, where $P_i^F \in \mathbb{R}^{b_i}$ is the distribution in case the river is "full" and $P_i^E \in \mathbb{R}^{b_i}$ in case the river is "empty". Use of the two distributions is computationally quite expensive and in cases use of only one distribution is recommended. The weighing factor is defined as $\lambda_i(t_k) := \frac{1}{b_i} \sum_{s=1}^{b_i} u_i(t_{k-b_i+s})$ and scaled into $[0, 1]$. In other words, the delay distribution of a moment

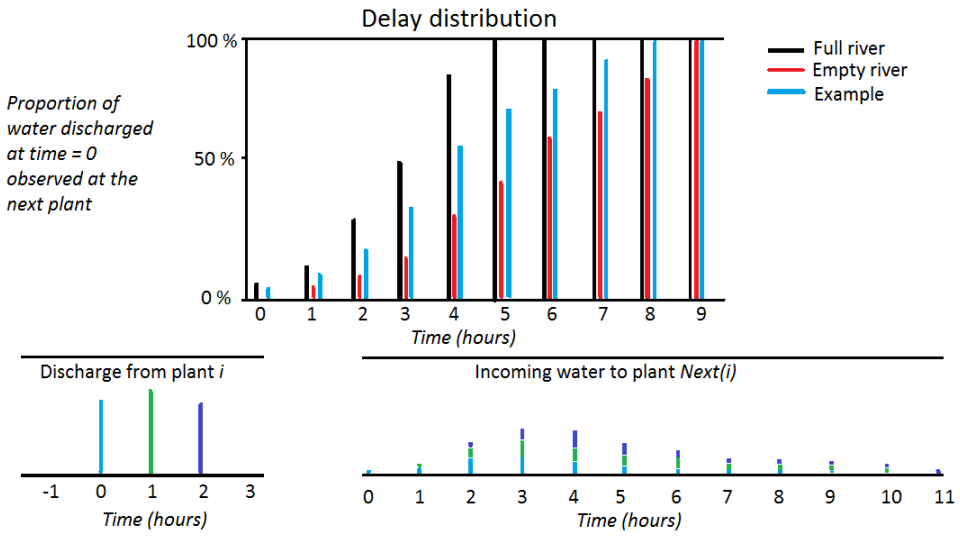

Figure 7: The water discharged flows to the next plant according to the delay distribution.

depends on the past discharges whose length is determined by the length of the distributions, b_i. In Figure 7, one example is presented with $b_i = 10$. In the upper sub-figure, $\lambda_i \approx 0.5$ and in the lower one it is slightly increasing from the initial level of about zero.

The tail-water level of the plant i depends, with exceptions, on the next plant's upper reservoir level, and on the occurred discharges of the plant. One way is to model the tail-water using an autoregressive model with external variables (ARX). One external parameter could be the plant total discharge. The main advantage of ARX is that the parameter values can be estimated effectively from the close past using the standard least square fit method. Hence, the model automatically adapts to the current situation. It is possible to choose between different ARX-models that have different numbers of AR- and X-parameters so that the model minimizes the error between modelled and measured tail-water levels with minimum number of parameters. In the very model, the selection is based on the Akaike's information criteria. Another way is to model the tail-water with partially linear fit as a function of plant discharge. Both methods encounter problems with tail-waters that incorporate so called hysteresis-phenomenon, as illustrated

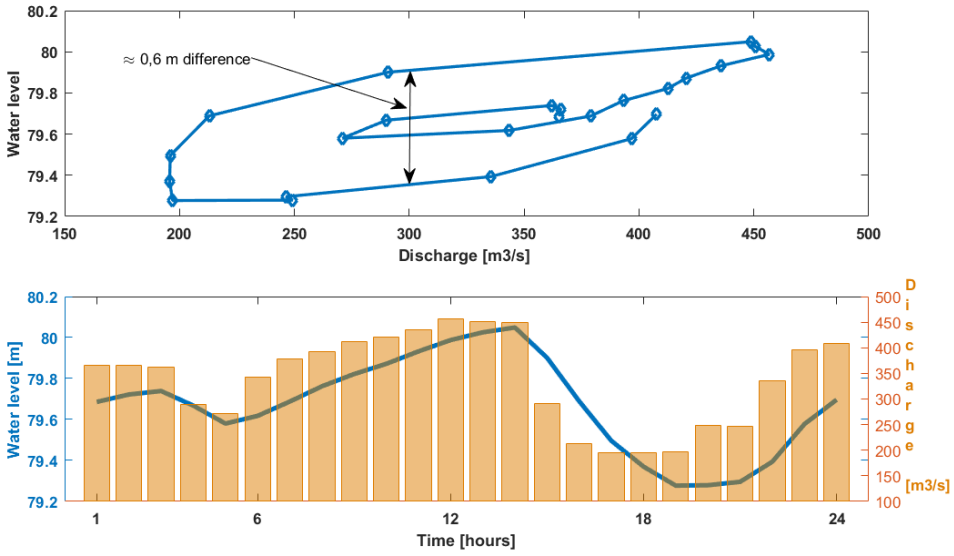

Figure 8: An example of the hysteresis-phenomenon in the tail-water of a plant. In the lower figure, the bars represent the discharge and their scale is the right y-axis.

in Figure 8. Because of the phenomenon, the tail-water levels with the same discharge remarkably differs during the day. The deviation of $0,6$ meters in the head of about 20 meters causes approximately 3% error in the production rate if the model does not take into account the hysteresis.

Each plant i has J_i turbines and each turbine has its individual technical properties such as maximum discharge, waterway losses, efficiency table η_{ij} (as represented in Figure 4), maximum rated production and amounts of different technical reserves.

The decision variables are discharges and the number of units on. At time t_k, the discharge through plant i is sum of the turbine discharges and spillage, $\mathbf{q}_i(t_k)$. The turbine j of the plant i is being utilized at time t_k if $S_{ij}(t_k) = 1, j = 1, \ldots, J_i$. In this model, this decision is made manually due to the original purpose of the model. The decision variable matrix, or the schedule, is formulated in the equation 7.

$$\begin{aligned} \mathbf{U}(\mathbf{Q}(\mathbf{t}), \mathbf{S}(\mathbf{t})) &= [\mathbf{u}(\mathbf{q}(t_0), \mathbf{S}(t_0)), \ldots, \mathbf{u}(\mathbf{q}(t_N), \mathbf{S}(t_N)] \in \mathbb{R}^{I \times N} \\ \mathbf{u}(\mathbf{q}(t), \mathbf{S}(t)) &= [u_1(q_1(t), S_1(t)), \ldots, u_I(q_I(t), S_I(t))]^T \in \mathbb{R}^I \end{aligned} \tag{7}$$

The explicit outputs are the production rate and the frequency controlled reserves over the horizon. Furthermore, the information of the regulating capacity rate is valuable for trading and is thus an implicit output of the model. Obviously, the outputs include also the water level estimates.

The optimization is based on an evolutionary heuristic algorithm, inspired by bee-algorithms presented by Pham et al. [2006] and Gavrilas [2010]. The main idea of the algorithm is to have numerous solutions that evolve based on mutual comparisons. The comparison is based on the Pareto-optimality of the objective functions $f_a(\mathbf{U})$.The value function $v_i(f_i(x))$ transforms the different objective functions into a commensurable value scale. The objective of the optimization is then to minimize the weighted sum of the different value functions, expressed in the equation 8. The best solutions, including Pareto-efficient solutions, continue to the next iteration loop with variations. The objective functions must be strictly decreasing so that the additive objective function results Pareto-optimal solutions, as discussed in Chapter 2.1.2.

$$V : \mathbb{R}^m \to \mathbb{R}, V(f(x)) = \sum_{i=1}^{m} w_i v_i(f_i(x)) \in [0, 1] \tag{8}$$

In practice, the objective functions are the energy value of water in the reservoir, the energy of moving water and some other functions that correspond the values of the producer. In addition, a constraint can be implemented as an objective. For example, the deviations from the required spinning reserves can be modelled as "fines".

The model can deal with multiple objectives but the problem is that as the number of objectives increases, the number of Pareto-optimal solutions increases as well. Then, it is possible that the heuristic method does not know how to evolve. Also, as the dynamic decision space is not necessarily convex, there may exist local optima where the population gathers. Evading the local optima requires good global evolution rules, but their suitability or performance cannot be guaranteed.

2.1.6 Comparison of the results: A case example

This section demonstrates the ex post methodology with a concrete example using the river model presented in Section 2.1.5. The period is selected from August 2014. Wider statistics will not be provided, and any conclusions regarding the optimality must not be made based on this single case study. First, the optimized and simulated results are compared shortly. Next, the model is validated. The feasibility of the results is evaluated in the end.

The results of the case study are presented in Table 1. The optimization method discharged the same amount of water from the strategic reservoirs as in reality, but it did not satisfy perfectly the simulated total production. The error (-13 MWh) was taken into account in the benefit comparison. Compared to the simulated schedule, the ex post optimized one spared about 374 + (-60) + (-13) = 301 MWh energy, stored in the reservoirs or flowing between them. That made about 1,4 % of the total simulated production.

Table 1: Results of the case example, partly shown as differences.

	Simulated	Optimized
Energy Produced [MWh]	20944	20931 (-13)
Energy left (in reservoirs) [MWh]	0	374
Energy left (flowing water) [MWh]	0	-60
Strategic reservoir discharge [10^6 m^3]	45,187	45,187

The simulated and measured production of the whole cascade is illustrated in Figure 9. The error of the sum production varied hourly between $-5 \cdots +25$ MW. The measured production during the horizon was 20426 MWh, implying that the model produced 518 MWh (2,5%) more than measured. The sum of the hourly plant-wise errors (absolute value) varied between 10 – 25 MW, meaning that the relative error is more than 5% of the total production during times of lower production. During the times of higher production, there are times when the plant-wise errors summed up to the total error. In other words, all the plant-wise errors were systematically into the same direction. The absolute errors of the simulated and measured water levels, both the upper reservoir and the tail-water, for few plants are presented in Figure 10. The absolute errors of the reservoirs were less than 10 cm and those of the tail-waters only partly exceeded 20 cm.

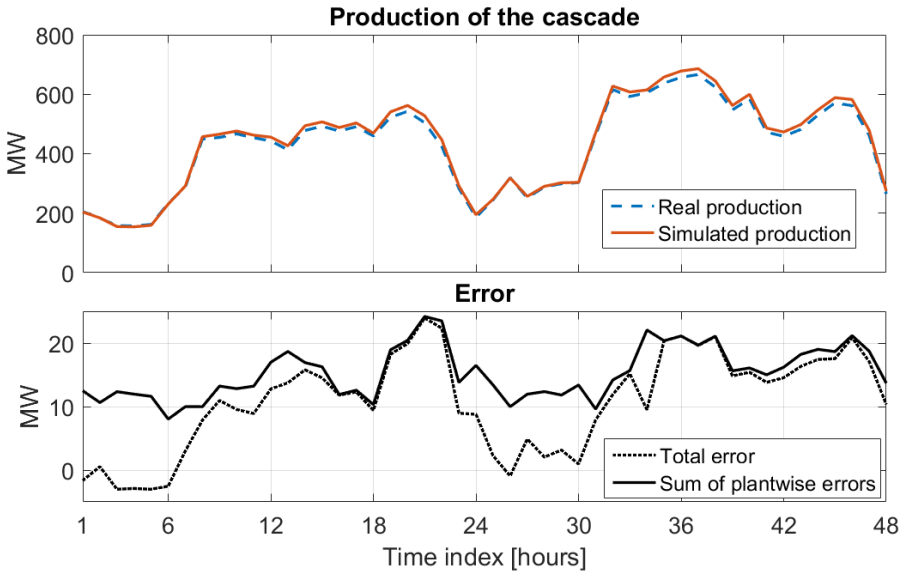

Figure 9: The total production of the cascade for the period of the case example. The dashed error line represents difference between cascade sums (simulated - reality) and the solid error line the sum of absolute values of plant-wise errors.

The optimized schedule, along with the simulated one, for three plants is presented in Figure 11. The first plant controls the strategic reservoir and therefore the water levels, or storage volume, must equal in the end of the period. The *simulated* discharges broke the regulation rule in the third plant, which did not happen in reality. The minima water levels of the reservoirs were less in the optimized results; in the second about 12 cm and in the third about 20 cm. Also, the high water levels of the last two hours might cause problems just after the optimization period. On the other hand, the optimized schedule respected all the operational constraints and requirements and hence the schedule could not be justified as non-feasible. The number of the turbine start-ups was bit smaller in the optimized schedule, which is a positive aspect for it.

To sum up, the optimized results spared energy about 150 MWh per day in the model world. With the energy prices of that time, the loss would have been worth of roughly €6000 per day.

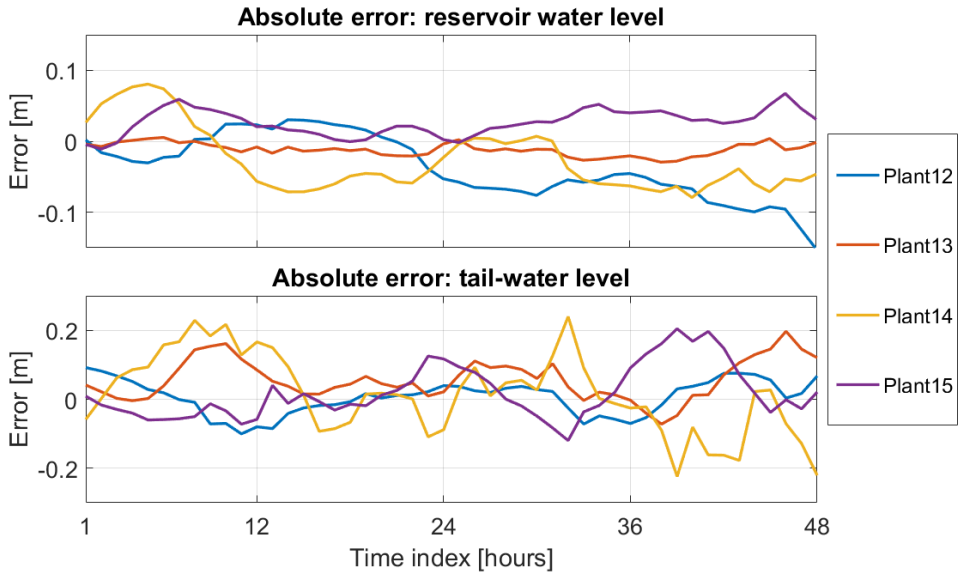

Figure 10: Absolute errors of the water level models for four plants. The nominal geodetic heads of the plants are between 11 – 21 meters.

2.2 Operative buffer

A new measure of operational quality, *operative buffer*, originates from the discussions related to the risk of the flowing water, and how different operators consider the suitable water levels in different ways. The operative buffer is defined as *time it takes to fill the reservoir in case the plant discharge would go to zero*. The buffer measures fullness of the reservoir with respect to the incoming water and its unit is time. Too long a buffer means that the reservoir is too empty and, respectively, too short a buffer that it is too full. A long buffer implies a loss of energy as the head of the plant could be higher, and a restricted capacity of up-regulation (increasing power rate output). On the other hand, a short buffer implies a restricted down-regulation and a higher risk of undesired spill discharge. The suitable range can be defined based on the relevant plant properties and, for example, on the distance from the nearest maintenance unit.

The calculation of the buffer is straightforward: until the reservoir of a plant is full, the consequent discharges of that plant are returned to the reservoir. The buffer is the time required to violate the reservoir constraint. The

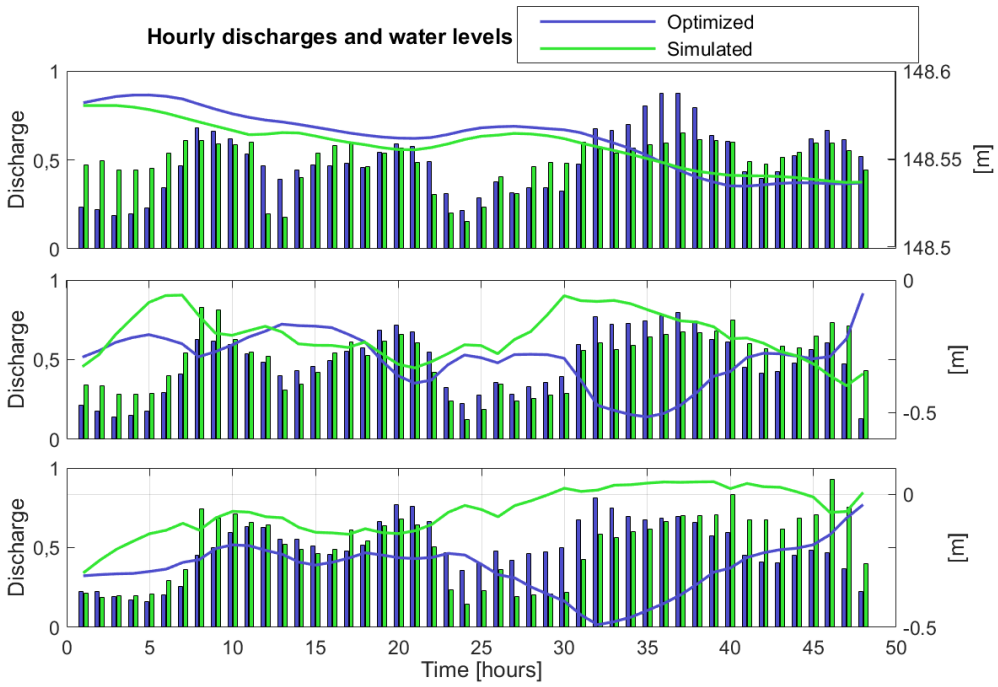

Figure 11: The simulated and optimized schedules and modelled water levels for three plants, presenting key findings on the optimized results. The uppermost figure depicts the strategic plant and therefore the water levels must equal in the end of the horizon. The water level of the plant in the middle increases rapidly in the end of the horizon. The simulated production in the lowest figure violates the regulation limit, which did not happen in reality.

calculation is repeated for each sampling time step and for each plant.

More formally, let $B_O^i(t_k)$ denote operative buffer of the plant i, $W_U^i(t_k)$ observed water level of the corresponding reservoir, $V^i(t_k)$ current volume of water and $Q^i(t_k)$ corresponding plant discharge at time point t_k, $k = 1, 2, \ldots, N_T$. The sampling is uniform, namely $t_{k+1} = t_k + \Delta t \forall k = 1, 2, \ldots, N_T$. Also, the water level constraint $C_U^i(t)$, the function $F_{WU}^i(W_Y, \Delta V)$ for estimating water level as function of initial water level W_Y and additional volume V are known. Then, the algorithm for calculating the operative buffer at time t_k is presented below. Its complexity for one period of length N_T is of order $\mathcal{O}(N_T)$, assuming N_T is much greater than s, where s denotes the maximum buffer length (in consequent time steps).

```
s = 0 % Number of consequent time step
v = 0 % volume of the returned water

while Ŵ_U^i(t_{k+s}) < C_U^i(t_{k+s})
    s = s + 1 % time step increment
    v = v + ∫_{t=t_{k+s-1}}^{t_{k+s}} Q^i(t)dt % return water
    Ŵ_U^i(t_{k+s}) = F_WU^i(W_U^i(t_{k+s}), v) % estimated level

end

B_O^i(t_k) = s · Δt % Buffer length
        % is the number of time steps
        % multiplied by the length of one step

% Then, continue to the next time step t_{k+1}...
```

Figure 12 demonstrates the operative buffers for a plant data covering three summer periods. The length of the buffer is mostly over 50 minutes and most often is more than two hours. The shape of the buffers reflects the dynamics of the operation: the discharge of the plant is increased when the discharged water from the upper plant comes available. Waiting for that causes the shorter buffer in the morning.

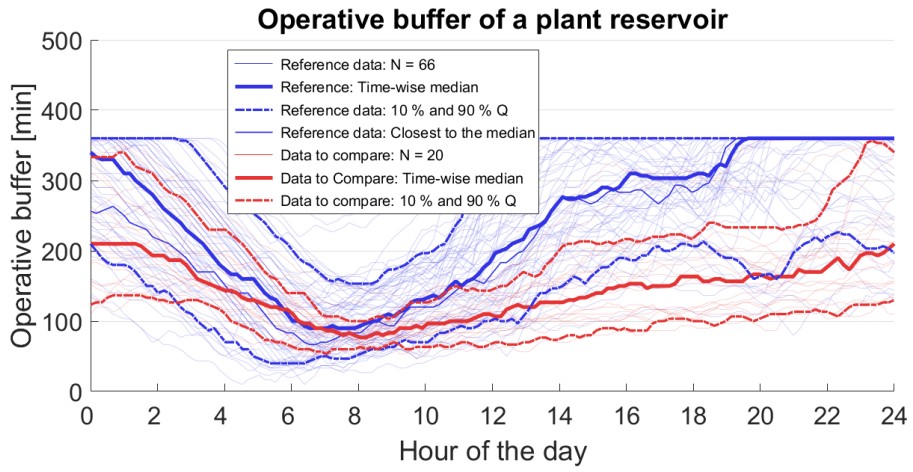

Figure 12: The operative buffers visualized and layered day-on-day for a plant reservoir. The reference data is from the years 2012 – 2014 and the red data is the corresponding period of 2015.

2.3 Evaluation

The *operative buffer* has many straight advantages. Firstly, this could facilitate the operative decision-making as one operator typically controls multiple reservoirs. Secondly, the managers of the operation can determine the suitable buffer limits so that the risk of the spillage is independent of the operators' risk profiles. Thirdly, calculated from the historical values, the buffer can be utilized for revealing systematic errors of the model used in optimization.

In practice, the reservoir is often modelled simply based on the maximum volume and the corresponding area but the bottom shape is not evaluated. The effect of the bottom shape can often be neglected as the relative change of reservoir areas with different water levels is small. More significantly, this method does not take into account the waving effects in the reservoir. Therefore, the modelled reservoir level might differ remarkably from the reality, implying wrong timing of the constraint violation. On the other hand, one remarkable source of error is eliminated because the calculation for historical values does not need the delay model. The water from the upper plant(s) is already reflected in the reservoir level. Calculating the buffer of the modelled results, for instance of the planned schedule, incorporates the

errors of the delay model as well.

The *ex post optimized* schedule expended 150 MWh (1,4 %) less net potential energy than in the reality. The same amount of the electricity was produced, thus implying improvement in the efficiency. The order of relative magnitude was less than in the results presented by Zheng et al. [2013]. The main reason for that difference might be that the hydrological and price-related uncertainties regarding the whole year are much more significant than those of the two following days. One of the main conclusions of Zheng et al. [2013] was that the accuracy of different forecasts are of high economic importance to hydro-power companies.

The weaknesses of the heuristic method were revealed clearly in this case study. The simulated total production was not exactly satisfied, though the error is negligible (less than 0,1%) and could be taken into account in the comparison. Furthermore, the optimality or Pareto-optimality of the schedule cannot be guaranteed.

The key assumption of the methodology is that the deviation from the optimal efficiency in the model world would be preserved in the real world. But, to what extent does the key assumption hold? If the model and the methodology were perfect, then the assumption would hold. There are, however, many sources of errors in the methodology.

The primary sources of systematic errors are the production model and the river model. The production model includes for example the efficiency tables and all the relevant technical parameters of the generating units. The river model consisted of the plant interconnection model, tail-water models and the plant interconnection model. In the case example presented, the total errors of the model (2,5%) were greater than the gap to the optimal efficiency (1,4%).

The secondary sources of systematic errors relate to the ex post methodology itself. In the procedure proposed, it was ensured that the decision space in the ex post-situation must correspond the real situation. It, of course, includes realistic turbine availability but also that the optimization period is only a sub-period of the simulated one. Then, the water from the simulated past is rolling downwards during the optimization horizon. So, the decisions made before the optimization period restrict the decision space during it. Also, as the water is already rolling downwards, the initial water levels of

the optimization horizon are aligned with the measurements – thus reducing model errors.

The point being is whether the errors of the model might favor, or alternatively disfavor, the optimization. In Figure 13 the reservoir level is higher in the model world, thus enabling more flexible use during the optimization. If the optimized results could be transformed to reality, it might appear that the optimal schedule is not feasible, or the gap is of different size. Based on that simple and provocative example, the gap is not necessarily preservable in the real world.

The problem illustrated in Figure 13 is similar as in our case example (the lowest sub-figure in Figure 11), where the *simulated* water level violated the regulation limits (did not happen in reality). In the optimized results, the water level reaches the minimum at the same time. Did that enable the improvement?

In theory, it might be possible to cluster all the identified errors into two groups: "favoring the optimization" and "disfavoring the optimization". If the groups are not in balance, then the results should be given less weight. In practice, can we be sure that all the errors are identified or their magnitudes evaluated correctly? Also, the dynamics of the river makes that question even more difficult. This question is not discussed further in this thesis. Until that problem is *reliably* solved, the procedure cannot be automatized and the feasibility of the results must be studied separately case by case. This declines the practical value of the methodology. For instance, is the selection of the samples justified?

The methodology can be used to reveal weaknesses or strengths of the river model or the actions conducted. They must be taken into account by both the operators and the developers of the model. In all, the methodology is suitable to quality analysis, though the results must be considered in accordance with the weaknesses of the methodology.

More broadly, the optimality as a whole is more relevant from the standing point of the shareholders. All the value streams must be considered. The same resource is part of larger portfolios and the energy is offered to many markets, or consumed. The value distributions of the electricity markets are quite skew, and therefore the assumption of the equal probability distributions does not hold (see page 13 and the figure 6), which was the necessary

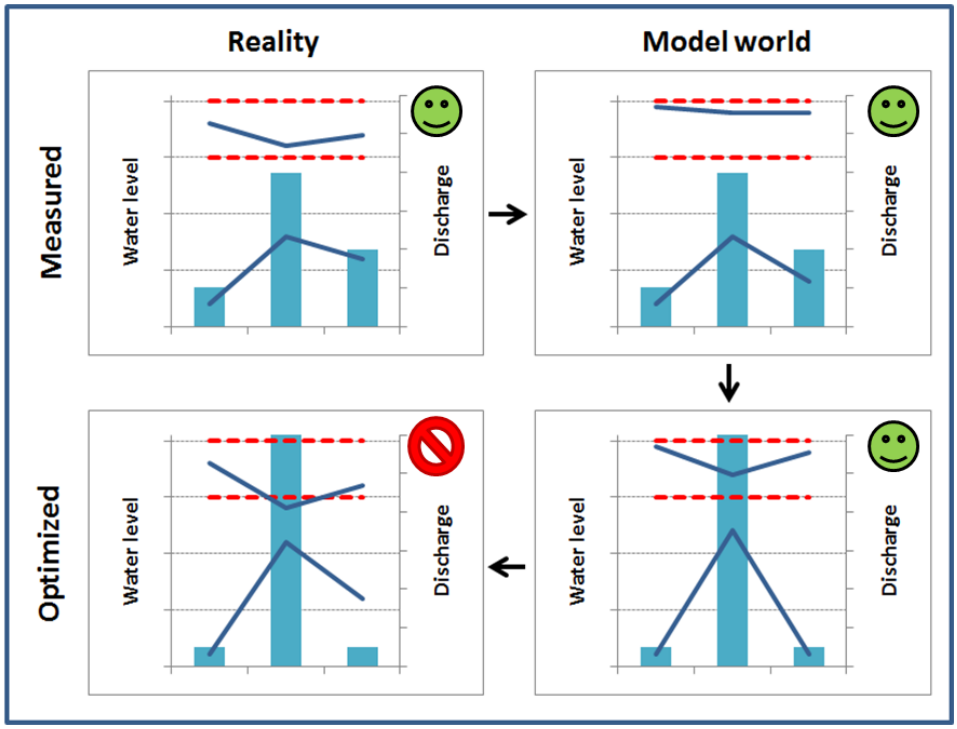

Figure 13: A conceptual example of a problem with the ex post methodology. The lines are water levels (left y-axis) and the blue bars discharges (right y-axis). The red lines are regulating limits of the reservoir water level, the upper blue lines depict the reservoir water level and the lower ones the tail-water level. The x-axis depicts consecutive time steps. The columns depict real and model world and the rows measured and optimized results. The optimized schedule is not feasible in reality as it would cause a violation of the regulation limit.

condition for the ex ante optimum to be the ex post optimum. As the assumption does not hold, should the ex post optimum be even expected? Is it relevant to separate this ex post methodology, measuring only the efficiency of the production, from the optimality of market actions? Those actions, or, preparing to them might, firstly, explain the deviations from the optimal efficiency and, most importantly, be more valuable than the losses in the efficiency. If so, the ex post optimal efficiency does not imply ex post optimal economic value, and thus it would not be a holistic measure of optimality. As the counter-factual outcome does not exist, that problem is difficult to tackle. However, it is always desirable to produce the *actual* output with higher efficiency as long as it adds value to the shareholders but it should not direct the actions in the first place. In addition, the existence of this kind of methodology forms a persisting incentive for the operators to focus on their performance and for the developers to continuously improve the river and the optimization model.

3 In search for typical

Before the river Kemijoki was harnessed, its flow changed significantly between seasons [Linkola, 1967]. Even though the river basin and the catchment area have remained the same, the river is now constructed and utilized for power production. As a result, the river course is maintained full whole the year. However, the within-day changes of the flows are larger than at its natural state. Due to its technical properties, the hydro-power production is suitable for balancing the electricity system – the production and the consumption must equal in the grid all the time. Up to these days, the daily production profile has followed the rhythm of natural life – electricity consumption is higher during day-time as shown in Figure 2. The behaviour of the river has possessed a clear daily rhythm, as exemplified in Figure 14 with a sample of daily reservoir water level curves. Also, the annual flood determines the annual rhythm. In all, these clear and natural rhythms leads us study the data that is splitted into days or years.

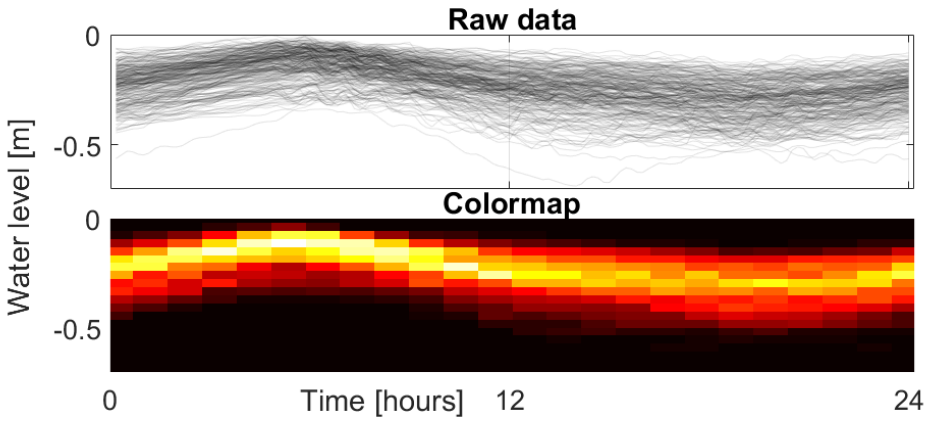

Figure 14: Daily water level observations (N = 266) presented. The $1 \times \infty$-dimensional curves have clear daily structure that leads us to study it daily-wise.

An underlying assumption behind the approach is that to the other users, both the professional and the recreational ones, it is essential that the behaviour of the river is predictable, or typical. For example, unpredictable changes of the flows might transport the nets of the fishermen. Also, the changes of the flows are reflected directly in the water levels that might af-

fect the recreational users.

In future, the typical rhythm may change. Firstly, the transition of the energy production towards the renewables may raise the amount of intermittent production, which requires more flexibility from the system.[7] That, on the contrary, might add value for the actors that can balance the system. The production of hydro-power is among them, and its timing may diverge from the current typical. From the wider perspective, the annual rhythm may change as well. On that subject, Veijalainen [2012] concludes that the winter discharges may increase, the spring floods distribute more in time and the variation between dry and wet periods may become more remarkable.

Functional Data Analysis (FDA) refers to the field of mathematical statistics in which observations are considered as functions instead of points. Ramsay and Silverman [2002] provides the basic principles related to functional fitting and statistical analysis. Also, they present a wide range of application areas from weather curves to human growth curves and from economics to recognition of hand-writing. In general, the functions are multivariate and the time does not have any special role among other dimensions. However, quite often the time is the baseline dimension in the applications. For example, daily weather can be studied by observing many variables such as temperature, humidity and wind during the day. Then, the sample consists of multiple days. Hubert et al. [2012] and Claeskens et al. [2014] capsulize clearly the fundamental questions of interest of FDA: the estimation of the central tendency of the curves, the estimation of the variability among the observations, the detection of outlying curves and classification or clustering of the observations. This thesis focuses on the first two objectives. The findings form a basement for later studies on the last two objectives.

How to define the most typical functional datum? A conceptual example of this problem is shown in Figure 15. Is the curve in the middle "the most typical curve"? It is the closest to both the total and the time-wise mean, and is the median observation at each time point. However, clearly it is an exceptional datum as it has an exceptional characteristic, the rapid variation. Common measures of centrality, or typicality, are average or median calculated coordinate-wise. More sophisticated version is the functional median discussed in the chapter 3.3.4. Another family of solutions is provided

[7]For example the Finnish TSO Fingrid discusses this topic in its publication "Electricity markets needs fixing - what can we do?", published 17th May 2016.

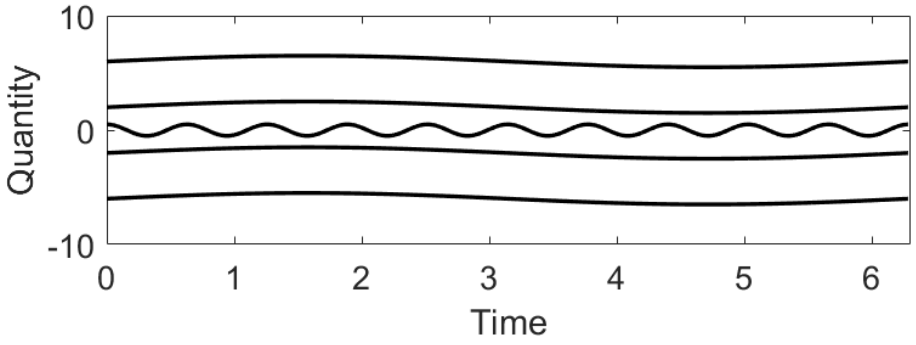

Figure 15: Five simulated functional univariate datum. Note that this is a conceptual example and a statistical analysis cannot be based on this few observations in theory or in reality.

by depth functions that measure centrality of multivariate observations and provide ordering of the curves based on that measure.

For consistency, the notations are agreed upon here. Let $C(I)$ be a complete normed space of continuous functions defined on the compact interval I. Let $\mathcal{Y}(t)$ denote a s-variate stochastic process with distribution $P(t)$, or with the cumulative distribution function $\mathcal{F}_Y$. In general terms, the data is $s \times \infty$-dimensional, as the time is a continuous dimension. Realizations of the stochastic process are paths in $C(I)$. Let $\mathbf{Y}_i(t_j)$ denote the i^{th} s-dimensional observation at time $t_j \in I$ of the sample. Observing the process T times at I produces one datum (curve, observation):

$$\mathbf{Y}_i = [\mathbf{Y}_i(t_1), \mathbf{Y}_i(t_2), \ldots, \mathbf{Y}_i(t_T)] \tag{9}$$

Then, we will work on a sample of N such curves:

$$\mathbf{Y} := \{\mathbf{Y}_i\}_{i=1}^{N} \tag{10}$$

Using this notation, the time-wise calculated average is calculated as below (11). It is a common measure of functional centrality. It can be used for example for functional *ANOVA*, discussed by Cuevas et al. [2004], in which the test statistic is based on the group means.

$$\hat{\mathbf{Y}} := [\frac{1}{N}\sum_{i=1}^{N}\mathbf{Y}_i(t_1), \frac{1}{N}\sum_{i=1}^{N}\mathbf{Y}_i(t_2), \ldots, \frac{1}{N}\sum_{i=1}^{N}\mathbf{Y}_i(t_T)] \tag{11}$$

The aim of this chapter is to define measures of normality for different time-varying quantities. In this context, the data consists of day-layered water level or discharge measurements. An example of such data is given in Figure 14. The problem of typicality is approached from the following three directions. The first one converts the curves into points. The second applies Functional Data Analysis (FDA) to study the problem. In the end, we combine the two with the concept of Pareto-optimality and propose a new methodology for finding the most typical observations in multi-variate setting.

3.1 Analyzing functional data using point values

3.1.1 Descriptive statistics

One can get an overview of a process or functional data via descriptive statistics. For example trends, averages, ranges, extreme values or variation measures can be used. Two real-data samples of water level measurements are depicted in Figure 16 using histograms and box-whiskers plots of daily ranges, daily means and daily extreme values. The data set *Summer* consists of workdays from three summers (266 days) and the set *Winter* of one winter time (65 days). According to the visualization, the water level minimum is lower and the range higher in winter time. Also, daily range of discharge seems to be slightly larger in winter time as well. The daily mean discharges are approximately the same during the both seasons but in winter time they distribute more densely.

3.1.2 Depth functions for point-wise data

Typical is not unambiguous to agree about even in simple one-dimensional cases. For instance, salary distributions may be strongly skewed. So, in practical applications the distributions do not necessarily satisfy (multi)normality assumptions. In such cases, even a few outliers may significantly influence

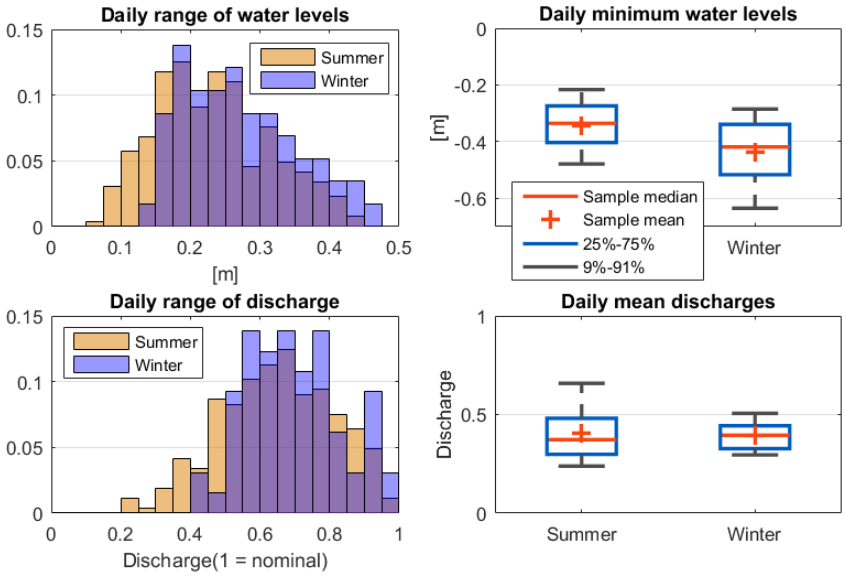

Figure 16: Comparison of summer (N = 266) and winter (N = 65) measurements with histograms and boxplots. Water level 0 m corresponds to the maximum level of the reservoir and the discharge is scaled into interval $[0, 1]$. Box and whiskers plots are generated using the Matlab-code provided by Jonathan Lansey.

the centre, or the location, of the sample. That phenomenon is underlined if the sample sizes are small. For instance Oja [2010] considers *several sample location problem* using standard MANOVA-method but also developing new robust non-parametric test statistics. One of them is based on a spatial rank function that ranks d-variate points inside the p-sphere. Even though the test statistics are not affine invariant, they are location and scalar invariant, which is often sufficient in practice.

Another way to address the problem is the depth functions that provide an ordering from the most central object, or the deepest observation, to the least central one. Historically, the concept of depth was first discussed by Tukey [1975] to generalize *statistical ranks* to multivariate setting. The proposed *half-space depth*, or *Tukey depth* of a multivariate point $\mathbf{x} \in \mathbf{R}^d$ is the minimum probability mass P of any closed halfspace H containing $\mathbf{x}$, or, $\inf\{P(H)|\mathbf{x} \in H\}$ [Zuo and Serfling, 2000, p. 461]. The equivalent formulation of [Claeskens et al., 2014, p. 8], given in the equation 12 is straightforward to transform to sample version. An example of a bivariate data set and its Tukey depth values is shown in Figure 17. Half-space depth has been studied widely and it plays, or its variations play, crucial role in many functional depth definitions as well. Rousseeuw and Ruts [1996] provide an efficient ($\mathcal{O}(nlogn)$) algorithm for calculating Tukey-depth.

$$HD(\mathbf{x}, F_{\mathcal{X}}) := \inf_{\mathbf{u}\in\mathbb{R}^k, ||\mathbf{u}||=1} P(\mathbf{u}^T\mathcal{X} \geq \mathbf{u}^T\mathbf{x}) \tag{12}$$

Zuo and Serfling [2000] proposes general prerequisites for any depth function, provided in the definition 3.1. The first desired property means that the coordinate system used or scaling must not affect depth ordering. The second means that if the centre can be uniquely defined, then the depth should get its maximum value at that centre. The third in turn means that the depth should decrease monotonically when moving away from the deepest point to any direction and the fourth that the depth should approach zero when approaching infinite distance from the deepest point. These all are important properties but a depth function can be feasible even if some of these are not satisfied. More desired properties, including symmetry around the deepest point, are discussed for example in Serfling [2006].

Definition 3.1. *Statistical depth function in* $\mathbb{R}^d$*:*
Given a point $\mathbf{x} \in \mathbb{R}^d$*, a class of distributions* $\mathcal{F}$ *of* $\mathbb{R}^d$ *and a distribution of a*

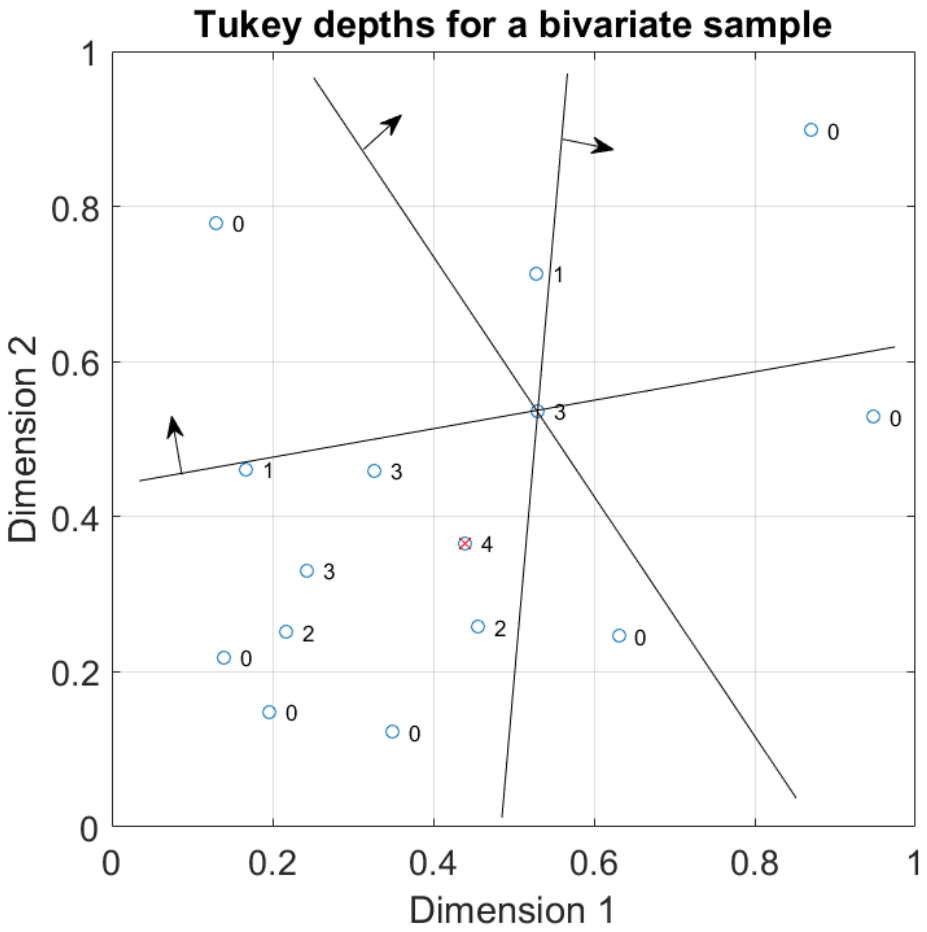

Figure 17: Tukey depths of a random two-dimensional data set, N = 15, evaluated. Next to each data point is given the depth rank and the depth value is the rank divided by N to give the probability mass. In this case only one point has the maximum depth, but in general that does not hold. For one point the evaluation is illustrated with lines and arrows. For that point it is not possible to make half-space division such that there would be less than 3 data points in the other half-space, yielding a half-space depth value 3.

general random vector $\mathbf{X}$ $\mathbf{F_X}$ *whose deepest point is the centre* θ*. Then, the bounded and non-negative mapping* $D : \mathbb{R}^d \times \mathcal{F} \to \mathbb{R}^1$ *is a statistical depth function if it satisfies the following properties:*

1. *Affine invariance:* $D(\mathbf{Ax} + \mathbf{b}; F_{\mathbf{AX}+\mathbf{b}}) = D(\mathbf{x}; F_{\mathbf{X}})$ *holds for any random vector* $\mathbf{X} \in \mathbb{R}^d$*, for any non-singular* $d \times d$*-matrix* $\mathbf{A}$ *and for any* d*-vector* $\mathbf{b}$*.*
2. *Maximality at the centre:* $D(\theta; F) = sup_{\mathbf{x}\in\mathbb{R}^d} D(\mathbf{x}; F)$*.*
3. *Monotonicity relative to the deepest point:* $D(\mathbf{x}; F) < D(\theta+\alpha(\mathbf{x}-\theta); F)$ *for* $\alpha \in (0, 1]$*.*
4. *Vanishing at infinity:* $D(\mathbf{x}; F) \to \infty$ *as* $||\mathbf{x}|| \to \infty$*.*

3.1.3 Bagplot

A bagplot is a bivariate extension of the boxplot. It consists of a convex hull, or a *bag*, that contains 50% of the observations and of *a fence* that separates outliers. Rousseeuw et al. [1999] define the bagplot for bivariate point-wise data using the half-space depth (see the equation 12). An example of half-space depth based bagplot on the historical flood data is shown in Figure 18. The selected variables for the bagplot are the magnitude and timing of the flood peak. Other important variables could be the starting time and the volume of the flood. The figure also illustrates the intersection of the boxplots of the marginal distributions. Hyndman and Shang [2012] extends the bivariate bagplot to the functional setup.

3.2 Time-warping

In the context of time based data, there may exist variation in amplitude or phase of the data. Time-warping, or registration, is the phase of making the observations comparable by manipulating dimension scales. Ramsay and Silverman [2002][p.102] define the time-warping formally by assuming a sample of N observation functions $x_i(t)$, $t \in [0, T_i]$, $i = 1, 2, \ldots, N$. These observations are transformed into a common interval $[0, T_0]$ using *time warping functions*:

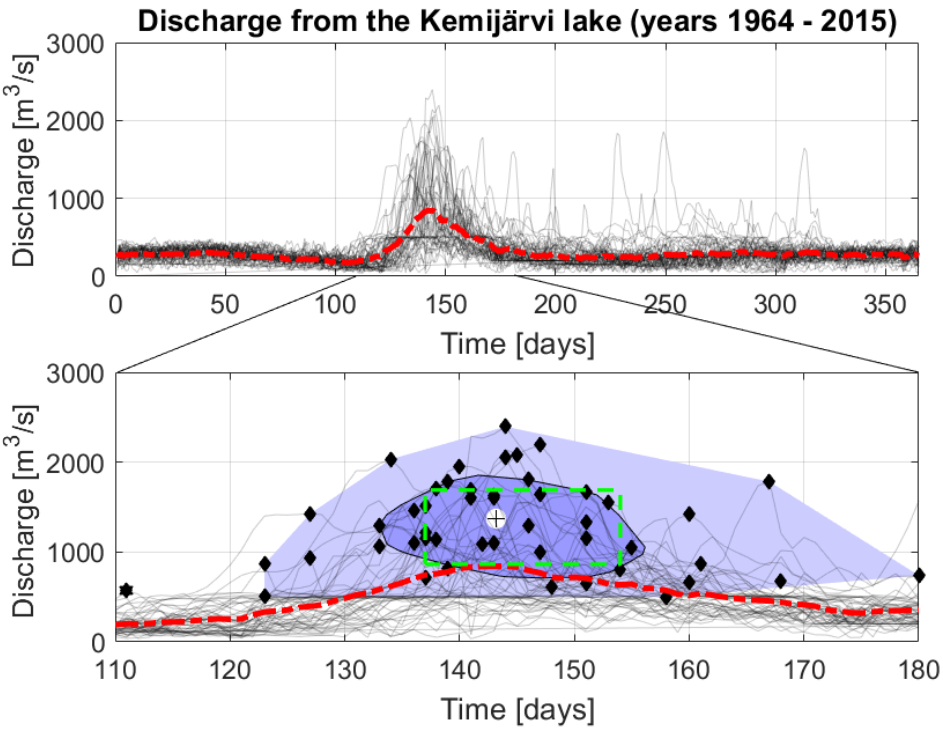

Figure 18: The historical discharge data from the Kemijärvi lake observed in 1964 – 2015. The lake is regulated and therefore the discharge is condensed at the plant nominal discharge of about 500 m^3/s. The x-axis is the ordinal number of the day (1st May = 120, leap years 1st May = 121). The red dashed line is the time-wise mean (eq. 11) that does not represent average flood. The diamonds in the lower figure represent the peak discharge of that year. Both the bag and the fence of the bagplot are drawn. The intersection of the marginal distributions boxplots is shown as dashed black box. The data is openly accessible at `http://www.syke.fi/en-us/Open_information` (SYKE, Finnish Environmental Institution). The Matlab-library for calculating robust statistics (including the *bagplot* applied here) is maintained at `http://wis.kuleuven.be/stat/robust/LIBRA/LIBRA-home` by the research group on robust statistics of the *KU Leuven*. The computation is described by Rousseeuw et al. [1999] and its variation by Hubert and Van der Veeken [2008].

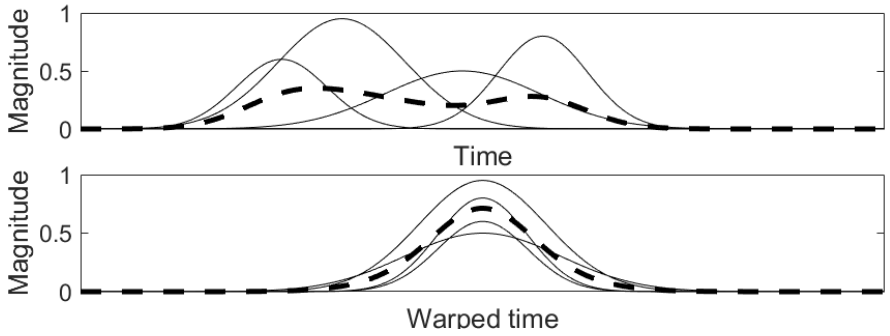

Figure 19: A conceptual example of time-warping. The time-wise mean is shown with the dashed line. The upper figure shows the original data as function of time. The time is warped in the lower one to align the peak magnitudes of the observations.

$$h_i(t) \in [0, T_i], \quad t \in [0, T_0]$$

For any time warping function holds strict monotonicity $h(t_1) > h(t_2) \Leftrightarrow t_1 > t_2$, which leads to *invertibility* $h_i^{-1}(h(t)) = t$, or, one-to-one correspondence between the time scales. Mathematically $h(t) > t$ and $h(t) < t$ correspond to the process "running slow" and "running fast", respectively. In the methodology described above, only the starting and ending points were fixed. Some more "landmarks" could be defined, and the registration could be implemented even continuously. This, however, requires a satisfactory measure of similarity with respect to the context. A conceptual example is shown in Figure 19 where the non-aligned observations are aligned according to the peak amplitude. This makes the time-wise calculated mean more representative of the data. Ramsay and Silverman [2002] also propose that these time-warped functions themselves are interesting subjects of study. For example, see Figure 20 presenting some time warping functions of the Berkeley growth data set[8]. Individuals seem to mature at different ages, which is normal among human beings. Note that the time t can be plotted against the difference $h(t) - t$ (*time deformation function*) to demonstrate the amount of warping [Ramsay and Silverman, 2002, p.108].

[8]Berkeley growth data set is open data and has become somewhat standard in FDA as the methodologies are often illustrated and tested with it.

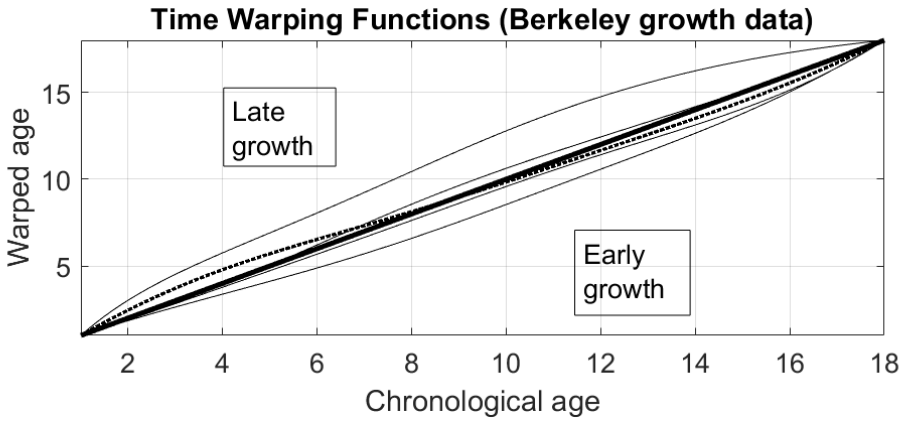

Figure 20: Two "normal" time warping functions that are close to the diagonal and two presenting early and late maturer of the Berkeley growth data. One observation (dashed line) turn from late maturer to early maturer before his/her early teenage years. The data is provided by Ramsay and Silverman [2002] as a web-appendix to the book.

3.3 Statistical functional depth functions

Nieto-Reyes and Battey [2016] provides the formal definition of statistical functional depth, given in Definition 3.2. The three first properties remain somewhat similar to those of the point-wise case (see Definition 3.1 for reference) but the two last properties are inevitable due to the functional setting. The first desired property is distance invariance which ensures that the depth ordering is maintained in any scaling that preserves distance up to non-zero scaling factor. The second property requires that the deepest observation should locate at a unique centre of symmetry. The problem with it is that existence of such *unique* centre is not clear. Therefore, the second requirement is altered so that the depth function should satisfy that property for well-known Gaussian process that is a zero-mean, stationary and almost surely continuous. The third and the fourth properties are the same as for a cumulative distribution function. The fifth property is more in-depth discussed in the referred source. Briefly, it aims to take into account such sub-sets of the interval I that exhibit only little variability within a sample giving less weight to those sub-intervals in the final depth value.

Definition 3.2. *Statistical functional depth function:*

Let $(\Omega, \mathcal{A}, P)$ denote probability space, $\mathcal{P}$ the space of all probability measures on Ω and $|\cdot|_p$ a p-norm. Then, the (bounded and non-negative) mapping $D : \Omega \times \mathcal{F} \to \mathbf{R}^1$ is a statistical functional depth function if it satisfies the following properties:

1. *Distance invariance: $D(f(\mathbf{Y}); P_{f(\mathcal{Y})}) = D(\mathbf{Y}; P_{\mathcal{Y}})$ holds for any random vector $\mathbf{Y} \in \Omega$ and for any $f : \Omega \to \Omega$ with the condition of $|f(\mathbf{Y}^*) - f(\mathbf{Y})|_p = a_f \cdot |\mathbf{Y}^* - \mathbf{Y}|_p$, given that $a_f \neq 0$.*
2. *Maximality at the centre: For any $P \in \mathcal{P}$ having a unique centre of symmetry $\theta \in \Omega$ with respect to some notion of functional symmetry, $D(\theta, P) = \sup_{\mathbf{Y} \in \Omega} D(\mathbf{Y}, P)$.*
3. *Strictly decreasing with respect to the deepest point*
4. *Upper semi-continuity in $\mathbf{Y}$: For all $\mathbf{Y} \in \Omega$ and for all $\epsilon > 0$ there exists a $\delta > 0$ such that $\sup_{\mathbf{Y}^* : |\mathbf{Y}^* - \mathbf{Y}|_p < \delta} D(\mathbf{Y}^*, P) <= D(\mathbf{Y}, P) + \epsilon$*
5. *Receptivity to convex hull width across the domain*
6. *Continuity in P.*

One important property of any estimator is its robustness against outliers. Breakdown point is the maximum fraction of arbitrary contaminated observations that the estimator tolerates without biases. For sample mean, for instance, the breakdown point is zero as one infinitely large observation biases the estimate. For point-wise L_p-depth functions Zuo [2004] introduces a finite sample breakdown point in order to study their global robustness. Functional breakdown point could be defined, for example, based on time-wise mean added by a big number. If that breaks the centrality estimator, then the breakdown point would be zero.

3.3.1 Minimum within-curves-distance depth function

Zuo and Serfling [2000] and Zuo [2004] discuss depth functions that are based on L_p-norms. In the equation 3.3.1, the deepest curve minimizes the sum of time-wise L_p-distances to all the other curves. Let $|\cdot|_p$ denote L_p-norm of d-vector.

$$LD_p(\mathbf{Y}) = \arg\min_i \sum_{s=1}^{N} \sum_{j=1}^{T} |\mathbf{Y}_i^j - \mathbf{Y}_s^j|_p \tag{13}$$

3.3.2 Multivariate functional depth function

Hubert et al. [2012] proposes a depth function called *multivariate functional halfspace depth function* (MFHD), given in the equation 14. An important block of that equation is the sample version of the half-space depth of the equation 15, where $\#(\cdot)$ counts the number of set members. Note that the extra time point in the weight function must be defined somehow, for example as $t_{T+1} = t_T + 0.5(t_T - t_{T-1})$. Claeskens et al. [2014] provide proofs that MFHD satisfies similar properties than those of Definition 3.1. All the proofs are based on that Tukey-depth do satisfy those properties. The authors do not discuss explicitly the properties of Definition 3.2.

$$MFHD_{N,T}(\mathbf{Y}_i, \alpha) = \sum_{j=1}^{T} w_{\alpha,N}(t_j) HD_N^j(\mathbf{Y}_i^j) \tag{14}$$

$$HD_N^j(\mathbf{y}) = \frac{1}{N} \min_{\mathbf{u}, ||\mathbf{u}||=1} \#\{\mathcal{Y}_N : \mathbf{u}^T \mathbf{Y}_i^j \geq \mathbf{u}^T \mathbf{y}\} \tag{15}$$

MFHD gives more weight to time points where the variation is more significant. The weight function $w_{\alpha,N}$ (eq. 16) takes into account the amplitude variability at each time by considering volume of the convex hull of the depth region (vol). The depth region at level α for any depth function $D_{\mathcal{F}_X}$ is defined in the equation 17.

$$w_{\alpha,N}(t_j) = \frac{(t_{j+1} - t_j) vol(D_\alpha(\mathcal{F}_Y^j)}{\sum_{j=1}^{T} (t_{j+1} - t_j) vol(D_\alpha(\mathcal{F}_Y^j)} \tag{16}$$

$$D_\alpha(F_\mathcal{X}) = \{\mathbf{x} \in \mathbb{R}^k : D(\mathbf{x}; F_\mathcal{X}) \geq \alpha\} \tag{17}$$

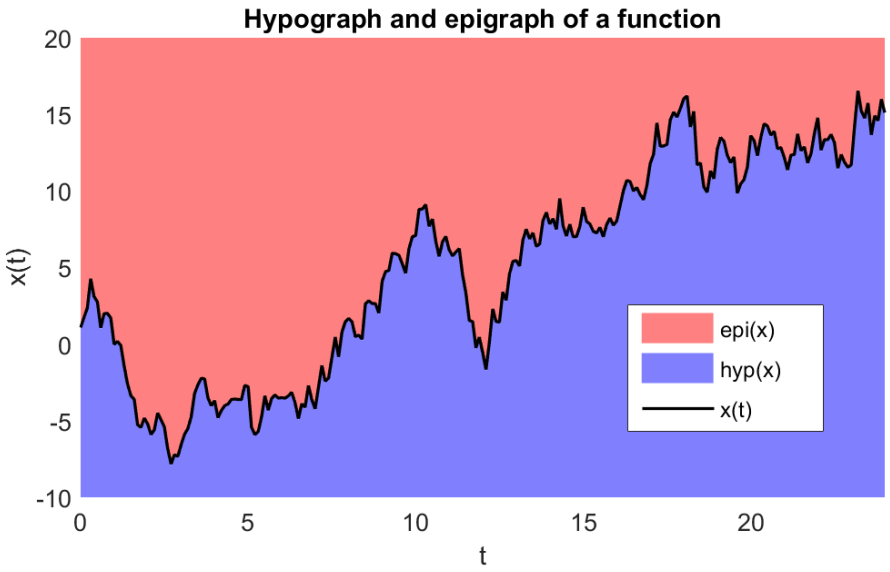

Figure 21: Hypograph and epigraph of a univariate function $x(t)$ (in this case Brownian motion).

3.3.3 Half-region depth

López-Pintado and Romo [2011] proposes computationally fast depth function that can be easily applied to multivariate data. The *half-region depth* HRD is based on the concepts *hypograph* and *epigraph* of a function $\mathbf{y}(t)$. The concepts are defined in the equation 18 and visualized in Figure 21. Note that their intersection is the function itself. The authors of the article discuss the general consistency requirements of Definition 3.1 but not those of Definition 3.2.

$$hyp(\mathbf{x}) = \{(t, y) : y \leq \mathbf{x}(t), t \in I\} \tag{18}$$
$$epi(\mathbf{x}) = \{(t, y) : y \geq \mathbf{x}(t), t \in I\} \tag{19}$$

The sample half-region depth HRD of $\mathbf{x} \in C(I)$ with respect to the whole sample $\mathbf{Y}$ is then defined below:

$$HRD_i(\mathbf{x}) = \frac{1}{N} \min\{\sum_{i=1}^{N} IND(\mathbf{Y}_i \leq \mathbf{x}), \sum_{i=1}^{N} IND(\mathbf{Y}_i \geq \mathbf{x})\} \tag{20}$$

where $IND(\cdot)$ denotes indicator function. $IND(\mathbf{Y}_i \geq \mathbf{x})$ gets value one if and only if $\mathbf{Y}_i$ belongs to the epigraph of $\mathbf{x} \in C(I)$ and zero in other cases. Kuelbs and Zinn [2015] discuss problems related to the half-region depth, namely that the depth might be zero-degenerate for sample whose all elements cross-over each other. Problems arise if the observations cross-over, namely belonging at times to epigraph and at times to hypograph because for such observation both $IND()$-functions will give value zero. The HRD can be modified so that it takes into account the proportion of time that the stochastic process is greater (superior) or smaller (inferior) than $\mathbf{x}$ [López-Pintado and Romo, 2011, p. 1689]. Its sample version is given in the equation 21 where $SL(\mathbf{Y}, \mathbf{x})$ and $IL(\mathbf{Y}, \mathbf{x})$ are the superior and inferior lengths, respectively and $\lambda\{\cdot\}$ denotes Lebesgue-measure[9] on $\mathbb{R}^s$. Somewhat similar and graphically grounded depth functions are band depth and modified band depth proposed by López-Pintado and Romo [2009] but those are not discussed further in this thesis.

$$MHRD(\mathbf{Y}, \mathbf{x}) = \min\{SL(\mathbf{Y}, \mathbf{x}), IL(\mathbf{Y}, \mathbf{x})\} \tag{21}$$

$$\begin{aligned} SL(\mathbf{Y}, \mathbf{x}) = &\quad \frac{1}{N\lambda\{I\}} \sum_{i=1}^{N} \lambda\{t \in I : \mathbf{Y}_i(t) \leq \mathbf{x}(\mathbf{t})\} \\ IL(\mathbf{Y}, \mathbf{x}) = &\quad \frac{1}{N\lambda\{I\}} \sum_{i=1}^{N} \lambda\{t \in I : \mathbf{Y}_i(t) \geq \mathbf{x}(\mathbf{t})\} \end{aligned}$$

3.3.4 Functional median

Gervini [2008] discusses a mean estimator called *functional median* that is more robust than the common sample mean. Its sample version for univariate

[9]Lebesgue-measure is the normal way of measuring subsets of s-dimensional Euclidean space, like *length*, *area* and *volume* for 1-, 2- and 3-dimensional cases, respectively.

$\{\mathbf{Y}_i\}_{i=1}^N$ is given in the equation 22, where the norm $||\cdot||$ is common L_2-norm $||\mathbf{x}|| = (\mathbf{x}^T\mathbf{x})^{1/2}$.

$$\hat{\mu} = \arg\min_{\mu \in \mathbb{R}^{1\times T}} \sum_{i=1}^{N} ||\mathbf{Y}_i - \mu|| \tag{22}$$

The sample estimator $\hat{\mu}$ must satisfy the equation 23 with the condition $\mathbf{Y}_i \neq \hat{\mu} \quad \forall\, i = 1, \ldots, N$.

$$\sum_{i=1}^{N} \frac{\mathbf{Y}_i - \hat{\mu}}{||\mathbf{Y}_i - \hat{\mu}||} = 0 \tag{23}$$

That prerequisite leads to the fact that the median is a weighted sum of the observations, provided that $w_i \geq 0\, \forall\, i$ and $w'w = 1$. The equation 24 holds even if the (23) is not defined.

$$\hat{\mu} = \sum_{i=1}^{N} w_i \mathbf{Y}_i \tag{24}$$

So, the functional median can be calculated if the weight vector is known. This and the original minimization problem of the equation 22 yields the new minimization problem (25) with respect to the weight vector. It can be formulated as the common convex minimization problem (eq. 26) with the aid of canonical vectors $\mathbf{e}_i$ (whose only non-zero element i equals one).

$$\min_{\mathbf{w}^T\mathbf{w}=1, w_i \geq 0} \sum_{i=1}^{N} |\mathbf{Y}_i - \sum_{j=1}^{N} w_j \mathbf{Y}_j|_2 \tag{25}$$

$$\min_{\mathbf{w}^T\mathbf{w}=1, w_i \geq 0} \sum_{i=1}^{N} ((\mathbf{e}_i - \mathbf{w})^T \mathbf{G} (\mathbf{e}_i - \mathbf{w}))^{1/2} \tag{26}$$

Gervini [2008] provides relevant proofs and the algorithm to solve the optimization problem of the equation (26).

3.3.5 Pareto-efficient functional depth

Pareto-optimality (discussed in chapter 2.1.2, page 8) can be applied in FDA as well. In this section, we propose a new method for finding Pareto-efficiently typical observation among a N-sized sample of $s \times \infty$-curves. The three phases of the method are the following:

1. Define the criterion of typicality and their measures.
2. Measure the data and order (rank) the results.
3. According to the rank scores, find Pareto-efficient observations. They are, according to the given criteria, the most typical observations.

The first phase is defining typical properties and their measures. In general, these are identified with the experts of the field. Applicable measures could be be for example mean, range, changes or variance of the data itself or of its derivatives. The second phase is ordering the data within all these measures individually so that the most typical observation gets the extreme value. For the ordering, one can use for example the spatial rank scores (mentioned in Section 3.1.2). If so, the ranks from the negative half-sphere must be transformed into positive one using for example absolute function. The third phase is finding Pareto-efficient points using these rank scores. The Pareto-points then pose as typical.

A similar technique could be applied in the functional quality analysis in general – there are no reasons for the criterion to measure only typicality. If the criterion measures extreme values, such as "minimize the range of the acceleration", then the ranking can be omitted. However, if the results are strongly skewed, some robust measure, such as spatial rank function, could be used.

3.3.6 Comparison of the depth functions

Nieto-Reyes and Battey [2016] discusses widely certain functional depth functions that have appeared lately in the literature. An interesting point stated is that the half-region depth (eq. 20) does not satisfy the second (even altered) property whereas the modified version (eq. 21) do satisfy it. The properties three and five are not satisfied by any of the depth functions of the

article. The article, however, does not consider the multivariate half-space depth (eq. 14) that especially intents to give more weight to sub-sections of wider convex hull volume. The requirements specified recently by [Nieto-Reyes and Battey, 2016, p. 64] are not discussed in the former articles on the subject, though discussion on consistency is provided in each article. That discussion is excluded from this thesis. We will illustrate the properties and compare the different depth functions using the real data.

Figures 22 and 24 present raw data and depth functions for cases A and B. The data of case A is water level measurements of a reservoir and the case B then controlled variable, discharge, of a generating unit. The depth functions are multivariate functional half-space depth (MFHD) function, L_2-depth function (LP) and modified half-region depth (MHRD) function. In the same axes it is drawn the area covered by the T_p^f-set of the corresponding depth function. The T_p^f-set is defined as the space below the time-wise supremum (maximum) and above the time-wise infimum (minimum) of the first $p\%$ deepest observations with respect to the depth function f. The histograms on the right side of the axis present the distribution of the corresponding depth values.

In Figure 22, the deepest observation with respect to the modified half-region depth function exceeds the typical daily range of water level (mean 0,2356 m and median 0,2263 m) with its range of about 0,38 m. The mean itself seems quite smooth and is affected by non-aligned events in time. Its shape or range underestimates the variation in reality. Also, the modified half-region depth has the largest variation as the thirdly deepest observation has a range of only about 0,08 m. The distribution of the L_2-depth function is the most skewed as the few outlying observations affect the measure a lot.

The property of variation within a depth function is illustrated with the *capture ratio* in Figures 23 and 25. The capture ratio at p is defined as the fraction of data mass belonging to the T_p^f. In Figures L1 refers to L_1- and L2 to L_2-depth functions. MHRD captures the measured values fastest in this case and MFHD differs only slightly from the L_p-depth functions. Interestingly, there are no notable differences between the capture ratio curves of L_1 and L_2.

In Figure 24 the depth functions do not capture the fundamental rhythm of the generating unit. The generating unit is shut down in more than 77% of curves, which makes that property quite typical. However, only L_2-depth

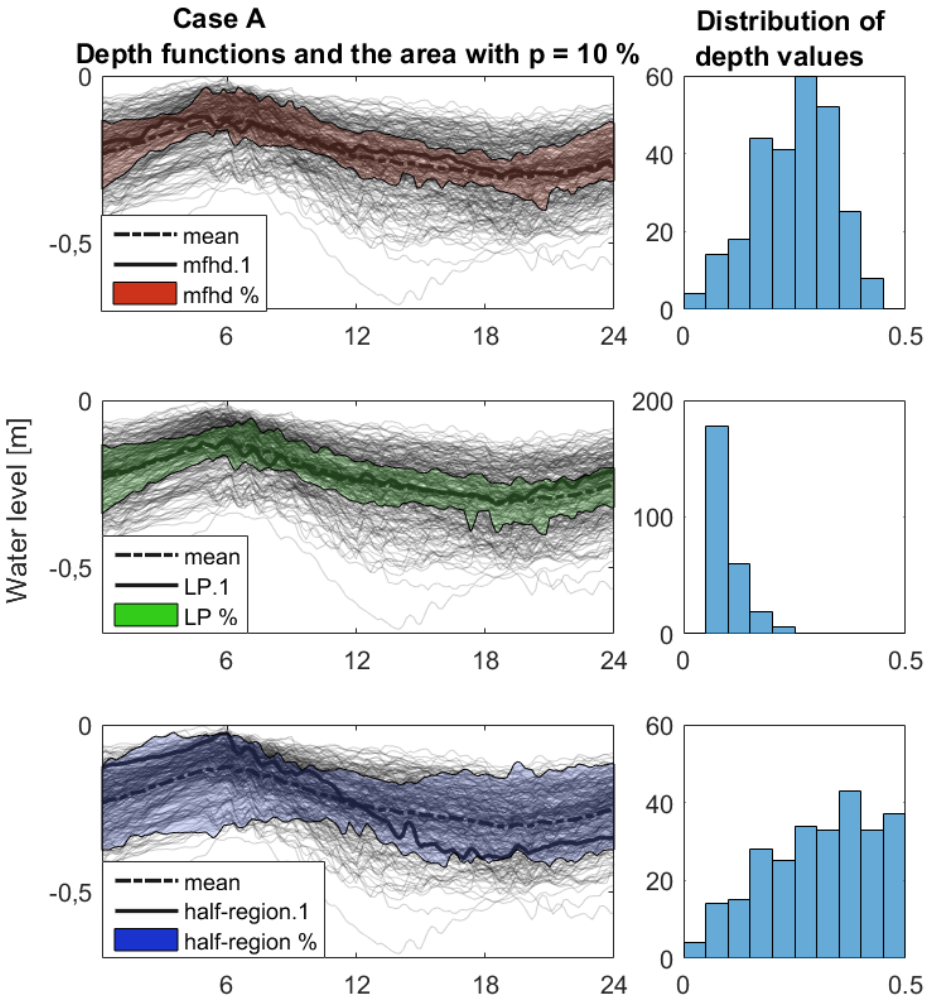

Figure 22: Summer-time water level measurements of a reservoir and three different depth functions day-layered (N = 266). The histograms depict the distribution of the depth functions.

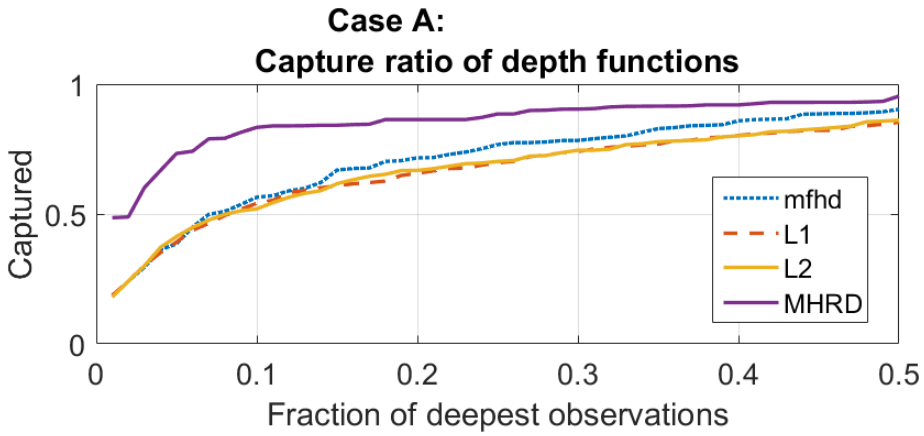

Figure 23: Capture ratios of different depth functions in case A. L1 refers to L_1- and L2 to L_2-depth function.

has one shutdown datum among the 5% deepest observations. MFHD gives more weight to the intervals of wide convex hull. In this case, the value range does not change remarkably but condensations clearly exist. Furthermore, none of the depth functions capture the discharge rate of about $150m^3/s$ even though there are densely observations and could be considered as typical. The steep slopes in the capture ratios (figure 25) mean that the depth functions widen their view on typicality at different times, or at different number of the deepest observations. So, if 10% (the same percentage as in the case A) of the deepest observations had been taken into account, all the depth functions would have recognized at least one night-time shutdown and thus would have captured more of the data mass. The 5% was selected because then the number of the deepest observations (N = 8) will compare to the Pareto-efficiently deepest observations, below.

Pareto-efficient depth seems to capture well the fundamental rhythm discussed above, as illustrated in Figure 26. The normality measures in the Pareto-methodology are central-ordered spatial ranks of the following criteria. "Typical volume" measures the amount of discharge during the day, "Typical sum of changes" refers to the total sum of the absolute changes during the day and "Typical range" the daily range of the discharge (max - min). The number of Pareto-efficient curves is eight (8) and they might include some curves that are not among the ten most central ones per criteria.

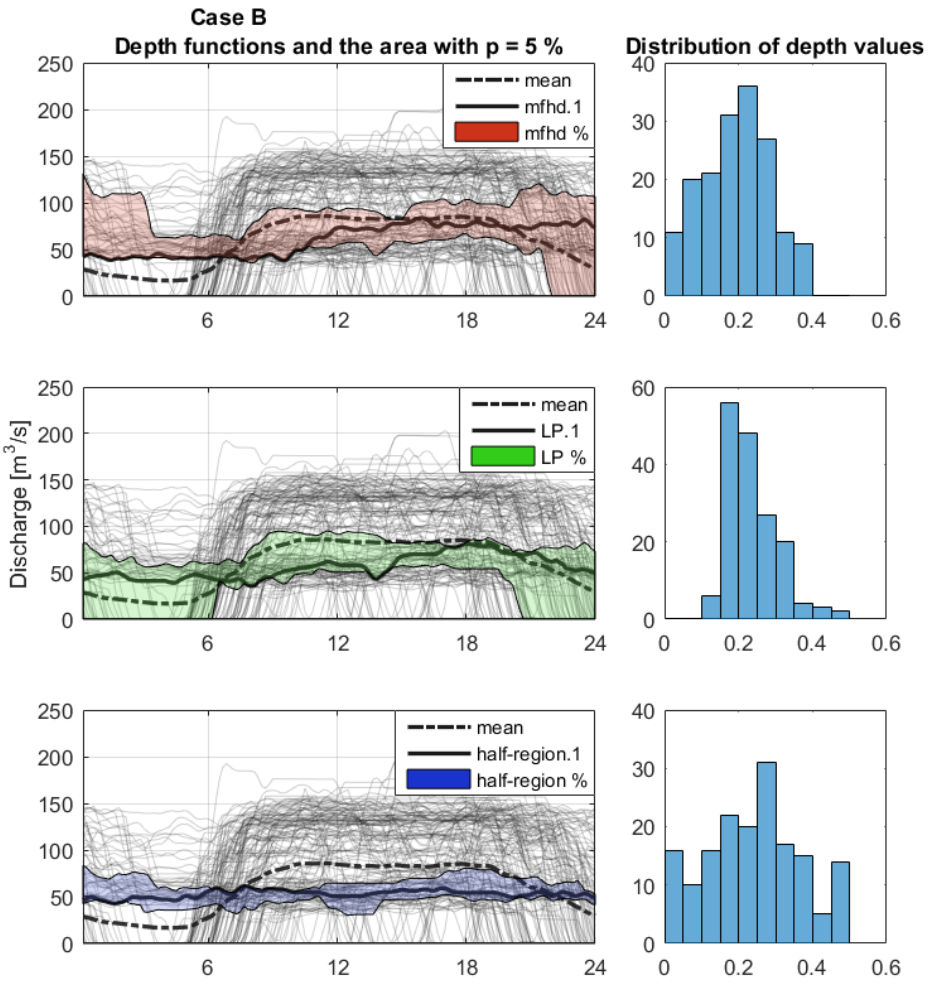

Figure 24: Discharges of a generating unit day-layered (N = 166). The histograms depict the distribution of the depth functions.

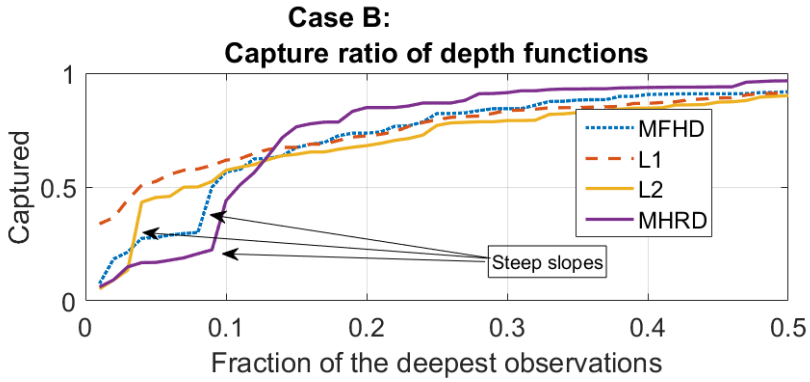

Figure 25: Capture ratios of different depth functions in case B. L1 refers to L_1- and L2 to L_2-depth function.

The same number of the deepest observations of the best depth function (L_2) of Case B is illustrated alongside, but its range is not typical and, except for one curve, the curves seems quite constant.

All the methods can be applied to multi-dimensional setting. In Figure 27, a bivariate setting with the data of the case A extended with the corresponding discharge data is presented from four perspectives. The applied depth function is the modified half-region depth (MHRD) function. The deepest observations evaluated using only univariate data do not exactly match with the deepest observations that were found using bivariate data. The bivariate deepest observations form a hull, as visualised in the two lower sub-figures. One observation seems an outlier, which is actually visible in the upper water level figure as well.

Pareto-methodology encounters a new problem in multi-dimensional setting where the typical properties and their measures is even more difficult to define. As a result the number of criteria easily increases. As a consequence, the number of the Pareto-efficient points often increases. For example, the same data as in Figure 27 resulted in 19 Pareto-efficient curves with two criteria, and 41 with four criteria.

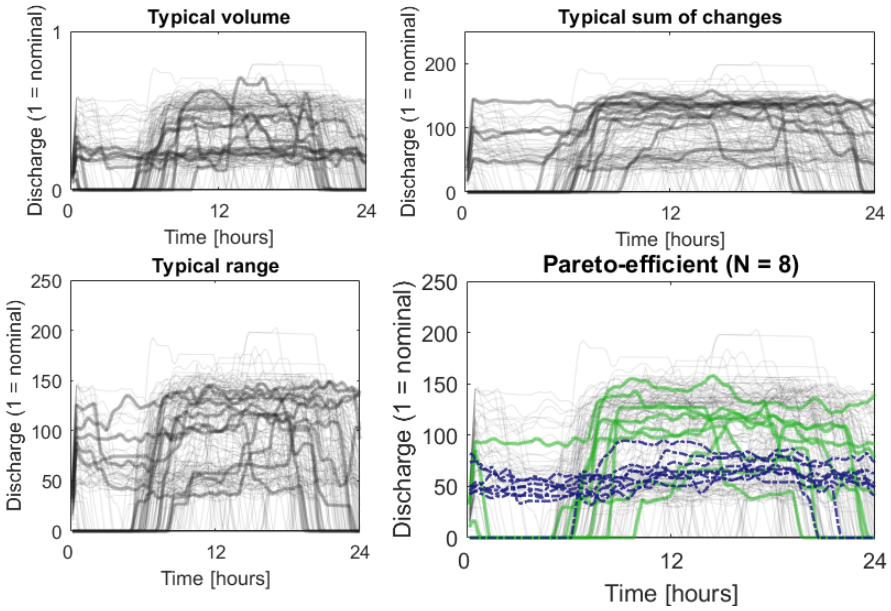

Figure 26: Three measures of centrality are ranked using spatial rank function. The ten most typical points per criteria are shown. Pareto-efficient curves (N = 8, green solid thick lines) may contain observations that are not highlighted in the other sub-figures. The 5% most L_2-deepest curves (N = 8) from Figure 24 are shown (blue thick dashed lines) alongside. The Pareto-efficiently deepest observations capture the rhythm of the phenomenon.

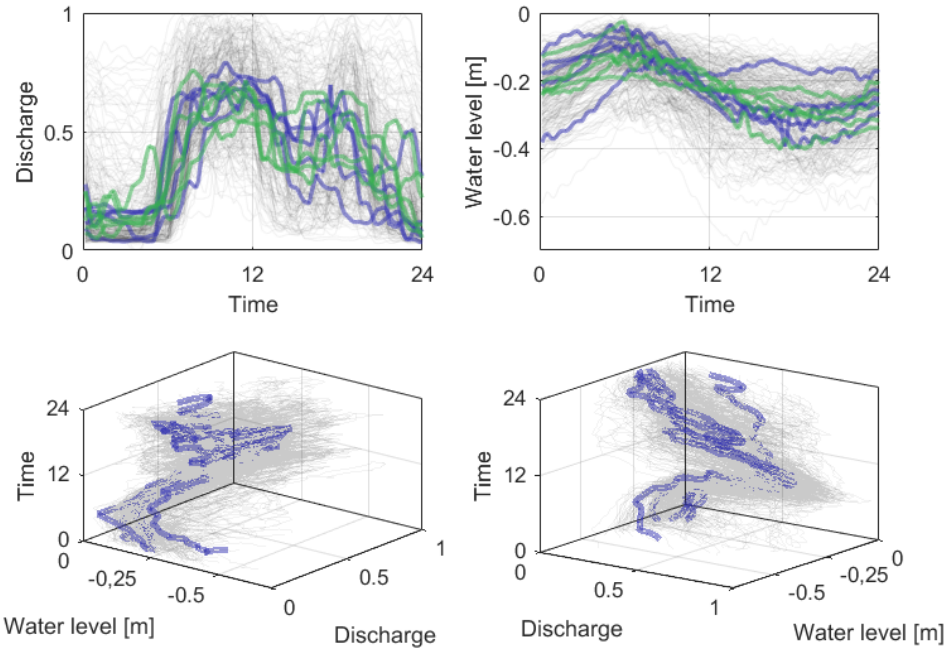

Figure 27: Depth in bivariate setting. In the two uppermost figures the green lines are the univariately evaluated deepest observations and the blue lines correspond to multivariate depths.

3.4 Evaluation

Exploring the phenomenon and its *raw-data* is essential. If a characteristic is not observed at this stage, it is possible that it is not taken into account in the subsequent considerations either. The first image could then affect, for example, the quality control process. The descriptive statistics offer a wide range of tools for that, including histograms and such. For example, the colourmap (see figure 14) clearly visualizes the overall rhythm, even though it loses the within-day progression of the individual curves. Therefore it is important to thoroughly study the raw data and its different properties – with the field experts who could explain the interesting characteristics.

The advantages of the Tukey-depth based bagplot are that it is robust against outliers, and its shape shows the centre and shape of the non-normal distribution as well, including correlation between the two variates. For example, the flood peak data of Figure 18 interestingly proposes that the flood peaks are higher if it occurs between days 140 – 150. The advantage of the intersection of the marginal boxplots is that it is easier to implement. As a future research proposal, the bagplot could be extended to three dimensions, where its geometric interpretation is still possible, or to N-dimensions.

The *time-warping*, or lack thereof, might lead to misinterpreting the results. For example, Figure 18 shows functional progressions of historical flood development of the Kemijoki river. The mean is represented with thick red dashed line but it does not depict correctly typical magnitude. It could be interpreted that the expected peak occurs at the end of May and has the maximum magnitude of about 900 m^3/s. However, the lower figure shows the peaks of the individual simulations with the diamonds. The majority of the individual simulations peaks up significantly over the peak shown in the upper figure. So, the averaging might lead to misinterpretation of the results, which might affect the preparation for the flood. Time-warping before averaging would produce more reliable peak magnitude in the mean scenario. However, the alignment rule requires contextual understanding of the problem and, even worse, cannot be unambiguously determined. Furthermore, presenting the data along registered time might bring about unwanted misunderstandings. In all, time-warping sounds intriguing but might be misleading in practice. The time-shifting functions themselves might be interesting area of studies in future.

All the depth functions discussed had weaknesses. The multivariate half-space depth (MFHD) was to pay attention at local variability but it clearly lost condensations of the data mass. It, along with the functional median, underestimated the daily range as they are stuck to the time-wise mean. On the other hand, the multivariate half-region depth had suspiciously large variation in it. In the few cases represented in this thesis, the Pareto-efficient depth function seemed the most acceptable depth function. In theory, the Pareto-methodology seems interesting as it inherently fulfills the *a priori* criteria that is not restricted by any assumptions. More broadly, the methodology brings together two independent fields of mathematics, Functional Data Analysis and Decision theory[10]. However, the "typicality" itself might be difficult to define a priori in multi-dimensional settings, which reduces the theoretical and practical value of the methodology. In practice, two or three dimensions might be the relevant maximum for the method. The measures of typicality might have dependencies and could be combined in some way. That would reduce the number of different criteria.

The fundamental idea of the Pareto-efficient depth methodology is to involve decision makers in the process of defining relevant typical characteristic and properties and their measures. As a consequence, the results may feel familiar. In case of depth functions, the complex and black-box measure (depth function) might be blamed if the results seem distant if compared to one's view on normality. Also, if the results do not seem good, one starts seeking faults in the a-priori definitions instead of the black-box.

[10]Decision theory widely defined. Pareto-optimality is applied in many fields that incorporates decision-making.

4 Conclusion

This thesis discussed different mathematical methods that could be utilized in the quality assessment of the hydro-power cascade operation. The first aspect aimed to assess operational optimality. For it, an ex post optimization methodology and an operative buffer were proposed and discussed. The aim of the second aspect was to find typical observation among functional data observations.

The ex post methodology assesses the deviation from the optimal efficiency in case there are no uncertainties. The methodology was considered conditionally suitable for the quality assessment. However, the errors of both the model and the methodology must be taken into account in the assessment. Due to the errors of the mathematical model, the modelled optimum might not be feasible in reality. The evaluation of the errors requires profound qualitative analysis of the results, which may be fruitful as such, but restricts usability of the methodology as a continuous performance indicator. The operative buffer is an interesting measure as it combines the efficiency of the production with the risk aspect. If the operative buffer is maintained at a relevant level, then the water management is efficient and does not cause redundant risks. In all, the methodologies form an incentive for the continuous improvement for both the operators and the developers of the model.

Altogether, the objective of the production is to maximize the net present value of *all* the value streams. As the distributions of the different benefits diverge significantly, the ex ante and the ex post optima are not equal. Hence, the deviation from the optimal efficiency, revealed by the ex post methodology, might have been well compensated by the benefits gained from the unpredictable markets. The most significant weakness of the methodology is that it only considers the efficiency of the production, not the whole economic value of it.

The second part of the thesis considered numerous depth functions that measure typicality of functional data. We proposed a new approach, *Pareto-efficient functional depth*, that combines Pareto-optimality and FDA. The fundamental idea is that the experts of the field, or the decision-makers themselves, may define a-priori criteria of the typicality. Then, the typical observations are those that are Pareto-efficient with respect to the given criteria. The method does not valuate the different criteria, whereas all the

other depth functions discussed assumed some characteristics and gave them implicitly different weights. Therefore, its results might be more convincing or persuasive. Still, determining the criteria form a problem in the new methodology. To sum up, the new approach might guide the discussion into the typical properties of the phenomenon instead of those of the method, which may be, in the end, more essential. The method could be studied further both from the theoretical and the practical point of views.

General problems of the quality assessment are its subjectivity and its quality. The former may lead to ambiguous measures of the quality and to eternal discussions around the trade-offs between the criteria. In the multi-objective situations, such as in the context of this thesis, the problem is highlighted. The quality of the assessment is poor, if it does not guide the actions into the desired direction. If the measures are misleading, they might lose their desired impact and thus the relevance. This problem must be considered in both the design and the implementation of any quality measure.

References

Jürgen Branken, Kalyanmoy Deb, Kaisa Miettinen, and Roman Słowinski, editors. *Multiobjective optimization: Interactive and evolutionary approaches*, volume 5252. Springer Science & Business Media, 2008.

Gerda Claeskens, Mia Hubert, Leen Slaets, and Kaveh Vakili. Multivariate functional halfspace depth. *Journal of the American Statistical Association*, 109(505):411–423, 2014. doi: 10.1080/01621459.2013.856795. URL `http://dx.doi.org/10.1080/01621459.2013.856795`.

Klaus Conrad and Ralph Unger. Ex post tests for short-and long-run optimization. *Journal of Econometrics*, 36(3):339–358, 1987.

Antonio Cuevas, Manuel Febrero, and Ricardo Fraiman. An ANOVA test for functional data. *Computational statistics & data analysis*, 47(1):111–122, 2004.

B.C. Flach, L.A. Barroso, and M.V.F. Pereira. Long-term optimal allocation of hydro generation for a price-maker company in a competitive market: latest developments and a stochastic dual dynamic programming approach. *IET generation, transmission & distribution*, 4(2):299–314, 2010.

Mihai Gavrilas. Heuristic and metaheuristic optimization techniques with application to power systems. In *Proceedings of the 12th WSEAS international conference on Mathematical methods and computational techniques in electrical engineering*, pages 95–103. World Scientific and Engineering Academy and Society (WSEAS), 2010.

Daniel Gervini. Robust functional estimation using the median and spherical principal components. *Biometrika 95*, pages 587–600, 2008.

David A. Harpman. Advanced algorithms for hydropower optimization. Technical report, U.S. Department of the Interior, Bureau of Reclamation, Technical Service Center, Denver, Colorado, 2012. URL `https://www.usbr.gov/research/projects/download_product.cfm?id=402`. Peer-reviewed technical report, attained 8th January 2016.

Richard Harris. Ex-post efficiency and resource allocation under uncertainty. *The Review of Economic Studies*, pages 427–436, 1978.

Mia Hubert and Stephan Van der Veeken. Outlier detection for skewed data. *Journal of Chemometrics*, 22(3-4):235–246, 2008.

Mia Hubert, Gerda Claeskens, Bart De Ketelaere, and Kaveh Vakili. A new depth-based approach for detecting outlying curves. In *Proceedings of COMPSTAT*, pages 329–340, 2012.

Rob J. Hyndman and Han Lin Shang. Rainbow plots, bagplots, and boxplots for functional data. *Journal of Computational and Graphical Statistics*, 2012.

Insinöörijärjestöjen koulutuskeskus INSKO. *Hydro-power plants (Vesivoimalaitokset)*. Insinööritieto Oy, Helsinki, 1978. In Finnish.

Kemijoki Oy. Vuosikertomus (annual report) 2015, 2016.

Teemu Kerttula. Short-term planning and bidding of hydropower production in the electricity market. Master's thesis, Tampere University of Technology, 2012.

Pekka Korhonen and Jyrki Wallenius. *Visualization in the multiple Objective Desicion-Making Framework*, chapter 8, pages 195 – 212. Volume 5252 of Branken et al. [2008], 2008.

Tapio Kovanen. *Voimatalous (Power engineering)*, chapter 21, pages 439 – 449. Volume 1 of Äijö et al. [1992], 1992. In Finnish.

James Kuelbs and Joel Zinn. Half-region depth for stochastic processes. *Journal of Multivariate Analysis*, 142:86–105, 2015.

Martti Linkola, editor. *Entinen Kemijoki (The Bygone River Kemijoki)*. Kemijoki Oy / Weilin & Göös, 1967. In Finnish.

Sara López-Pintado and Juan Romo. On the concept of depth for functional data. *Journal of the American Statistical Association*, 104(486):718–734, 2009.

Sara López-Pintado and Juan Romo. A half-region depth for functional data. *Computational Statistics & Data Analysis*, 55(4):1679–1695, 2011.

Mika Marttunen. *Interactive multi-criteria decision analysis in the collaborative management of watercourses*. PhD thesis, Aalto University, 2011.

Kaisa Miettinen. *Introduction to Multiobjective Optimization: Noninteractive approaches*, chapter 1, pages 1 – 26. Volume 5252 of Branken et al. [2008], 2008.

Kaisa Miettinen. *Nonlinear multiobjective optimization*, volume 12. Springer Science & Business Media, 2012.

Alicia Nieto-Reyes and Heather Battey. A topologically valid definition of depth for functional data. *Statistical Science [ISSN 0883-4237 (print); ISSN 2168-8745 (online)]*, 31(1), 2016.

Teemu Nurmi. Use of water level fluctuation indicators in assessment and operational use of water course regulation. Master's thesis, Aalto University, 2010. In Finnish.

Hannu Oja. *Multivariate Nonparametric Methods With R*. Lecture Notes in Statistics. Springer, 2010.

Oxford Dictionaries. Oxford University Press, 2016. URL `http://www.oxforddictionaries.com/definition/english/optimizationER`. The definitions of *optimization*, *ex post*, and *ex ante*. (accessed May 24, 2016).

Tapio Palonen. The impacts of Larsmo lake regulation on recreational use and the value of the water front. Master's thesis, Helsinki University of Technology (TKK), 2006. In Finnish.

DT Pham, A Ghanbarzadeh, E Koc, S Otri, S Rahim, and M Zaidi. The bees algorithm–a novel tool for complex optimisation problems. In *Proceedings of the 2nd Virtual International Conference on Intelligent Production Machines and Systems (IPROMS 2006)*, pages 454–459, 2006.

Juha Pursimo and Pentti Lautala. *Sähkön tuotannon lyhyen aikavälin optimointi - loppuraportti (Short-term optimization of hydro-power production - final report)*. 1993. In Finnish.

J.O. Ramsay and B.W. Silverman. *Applied Functional Data Analysis*. Springer Series in Statistics. Springer-Verlag New York, 2002.

Peter J. Rousseeuw and Ida Ruts. Algorithm AS 307: Bivariate location depth. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 45(4):516–526, 1996.

Peter J. Rousseeuw, Ida Ruts, and John W. Tukey. The bagplot: a bivariate boxplot. *The American Statistician*, 53(4):382–387, 1999.

Robert Serfling. Depth functions in nonparametric multivariate inference. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 72:1, 2006.

Lea Siivola. *Virkistyskäyttö (Recreational use)*, chapter 22, pages 451 – 468. Volume 1 of Äijö et al. [1992], 1992. In Finnish.

Ross M. Starr. Optimal production and allocation under uncertainty. *The Quarterly Journal of Economics*, pages 81–95, 1973.

John W Tukey. Mathematics and the picturing of data. In *Proceedings of the international congress of mathematicians*, volume 2, pages 523–531, 1975.

Noora Veijalainen. *Estimation of climate change impacts on hydrology and floods in Finland*. PhD thesis, School of Engineering, 2012.

Stelios H Zanakis and James R Evans. Heuristic "optimization": Why, when, and how to use it. *Interfaces*, 11(5):84–91, 1981.

Ying Zheng, Xudong Fu, and Jiahua Wei. Evaluation of power generation efficiency of cascade hydropower plants: A case study. *Energies*, 6(2): 1165–1177, 2013.

Yijun Zuo. Robustness of weighted L_p–depth and L_p–median. *Allgemeines Statistisches Archiv*, 88(2):215–234, 2004.

Yijun Zuo and Robert Serfling. General notions of statistical depth function. *Annals of statistics*, pages 461–482, 2000.

Helena Äijö, Lea Siivola, and Pentti Vakkilainen, editors. *Hyödyn ja haitan arviointi vesitaloudessa (Assessing advantages and disadvantages in water resource management)*, volume 1. Helsinki University of Technology (TKK), 1992. In Finnish.