Photovoltaic Output Modeling

Monitoring, Forecasting, and Applications

Herman Böök
Photovoltaic Output Modeling

Monitoring, Forecasting, and Applications

Herman Böök

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall K203a of the school, and also via remote technology, on 10 December 2021 at 12:00.

Aalto University
School of Science
Department of Applied Physics
New Energy Technologies
Supervising professor
Prof. Peter D. Lund, Aalto University, Finland

Thesis advisor
Prof. Anders V. Lindfors, Finnish Meteorological Institute, Finland

Preliminary examiners
Dr. Marion Schroedter-Homscheidt, German Aerospace Center, Germany
Dr. Rubén Urraca Valle, European Commission (Joint Research Centre), Italy

Opponent
Dr. Tomas Landelius, Swedish Meteorological and Hydrological Institute, Sweden
Photovoltaics (PV) has emerged from a niche market towards becoming a potential mainstream electricity source. Despite the rapidly increasing share of PV systems in electricity production, the widespread adoption of PV still poses a handful of challenges related to the intermittent and weather-dependent nature of its output. Even though certain best practices in PV output modeling have gradually taken shape, no comprehensive solutions have yet been established for handling the entity, as a whole. In this doctoral thesis, modeling methods for monitoring and forecasting the output of specific PV systems were developed and validated, also demonstrating the utilization of these methods in real-world applications for assessing the general viability of PV technology in a Nordic context.

In-situ measurements at two PV sites were used for optimizing and evaluating the performance of the PV output conversion model. The model was shown to perform well in snow-free conditions, demonstrating its value by estimating distinct system losses and its PV system monitoring capability. As a part of this process, the proposed quality control method for treating calculated direct normal irradiance (DNI) values was shown to provide a feasible approach for processing calculated DNI values. In order to attain an accurate and truthful depiction of a set of unique PV systems, the occurring characteristics of the investigated systems are to be taken into account. For sites where no external measurements are available, a novel approach for adjusting the baseline model for forecasting the output of specific PV systems was validated at 23 separate PV sites. The method was shown to capture time-dependent losses, proving it as a feasible approach for providing adjusted site-specific PV output forecasts.

The studied PV output modeling tools were demonstrated in three different applications, covering domestic hot water heating cost optimization with a PV output forecast-based control method, a residential PV profitability study in Finland, coupled with energy storage and optimization, and a virtual power plant concept, required to regulate and aggregate active consumer behaviour in the future markets. In each case, a clear added value of PV output modeling through distinct cost reduction potential and imbalance mitigation, could be demonstrated.

The contributions of this thesis demonstrate the performance of the selected modeling approaches, and provide tangible information about their application and value creation potential within the selected subject areas.


Väitöskirjan tulokset osoittavat valittujen mallinnusmenetelmin suorituskyvyn, tarjoten konkreettista tietoa kyseisten menetelmin soveltuvuudesta ja arvonluontipotentiaalista valittujen aihepiirien osalta.

Avainsanat aurinkosähkö, auringon säteily, laadunvalvonta, mallinnus, kiinteistöautoamatio

ISBN (painettu) 978-952-64-0610-7 ISBN (pd) 978-952-64-0611-4
ISSN (painettu) 1799-4934 ISSN (pd) 1799-4942
Julkaisupaikka Helsinky Painopaikka Helsinky Vuosi 2021
Solar energy and its diverse applicability potential has fascinated me for years. After some years of following the swift solar energy development from the sidelines, while simultaneously working with meteorology-related international development projects, I made a deliberate decision to pursue a multidisciplinary career in solar energy. This doctoral thesis is a culmination of that decision.

The work presented in this thesis was carried out in the Meteorological Research Applications group of the Finnish Meteorological Institute, while being enrolled as a doctoral candidate at the Department of Applied Physics at Aalto University. I would like to express my gratitude for the Nessling Foundation who supported my thesis with a long-term funding, alongside with the funding provided by the Finnish Meteorological Institute, in particular through the projects ICASIF (Academy of Finland) and BCDC Energy (Strategic Research Council). My advisor, Prof. Anders Lindfors from the Finnish Meteorological Institute, deserves my recognition for giving valuable feedback and insights on my research ideas and manuscripts, and for enabling my doctoral pursuit in solar energy. I acknowledge Prof. Peter Lund from Aalto University for the opportunity to enlist as a doctoral candidate under his supervision. All of the multidisciplinary manuscript collaborators in Oulu, Lappeenranta, and the Finnish Meteorological Institute also deserve my recognition. I am grateful to the pre-examiners, Dr. Marion Schroedter-Homscheidt and Dr. Rubén Urraca Valle, for their careful work and feedback on this thesis. I am also thankful to Dr. Tomas Landelius for acting as opponent in the defence of this thesis.

I acknowledge my mother for being a role model of perseverance and resilience, and my father for inspiring me towards an endless curiosity and an insatiable thirst for knowledge. I thank Elmo for being an unlimited source of optimism. Lastly, I recognize my better half, Dr. Essi Karjalainen, for the peer support and pressure, and everything else.

Helsinki, November 8, 2021,

Herman Böök
## Contents

Preface ......................................................... i

Abbreviations ................................................... v

List of Publications .......................................... vii

Author’s Contribution ........................................ ix

1. Introduction ............................................. 1
   1.1 Background ........................................... 2
   1.2 Objectives and scope ................................ 3
   1.3 Thesis outline ......................................... 3

2. Intermittent PV generation in the power grid and markets .... 5
   2.1 Power generation and power grid ....................... 5
   2.2 Electricity markets .................................... 7
   2.3 Local PV generation .................................... 8
   2.4 Solutions for tackling the challenges ................. 8

3. Photovoltaic output modeling ................................ 11
   3.1 Approaches ........................................... 13
      3.1.1 Physical models ................................... 13
      3.1.2 Parametric models ................................ 14
      3.1.3 Machine learning models ......................... 14
   3.2 Input data ............................................. 15
      3.2.1 PV system metadata ................................ 15
      3.2.2 In-situ measurements ............................. 16
      3.2.3 Satellite measurements ............................ 16
      3.2.4 Forecaster selection ............................... 16
      3.2.5 Uncertainty of input data ......................... 18
   3.3 Applications of PV output modeling ..................... 19
      3.3.1 Site planning and output monitoring ............. 19
      3.3.2 Output forecasting ................................ 20
## Contents

### 4. Results

4.1 QC method for handling calculated DNI values (Publication I) ........................................ 21
4.2 PV output modeling with in-situ observations (Publication I) ........................................ 23
4.3 Site-specific adjustment of a NWP-based PV output forecast (Publication II) ...................... 31
4.4 Application to assess a GSHP control method for PV-assisted DHW heating (Publication III) ........................................ 36
4.5 Application to assess small-scale PV profitability with thermal storage and carbon-corrected electricity prices (Publication IV) ........................................ 43
4.6 Application to assess distributed thermal storage utilization in PV output forecast error cost minimization (Publication V) ........................................ 50

### 5. Concluding remarks

5.1 Future prospects ........................................ 57

Bibliography ........................................ 61

Errata ........................................ 67

Publications ........................................ 69
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>AOI</td>
<td>Angle Of Incidence</td>
</tr>
<tr>
<td>BAS</td>
<td>Building Automation System</td>
</tr>
<tr>
<td>BRP</td>
<td>Balance Responsible Party</td>
</tr>
<tr>
<td>c-Si</td>
<td>Crystalline Silicon</td>
</tr>
<tr>
<td>CF</td>
<td>Capacity Factor</td>
</tr>
<tr>
<td>CHP</td>
<td>Combined Heat And Power</td>
</tr>
<tr>
<td>COP</td>
<td>Coefficient Of Performance</td>
</tr>
<tr>
<td>COR</td>
<td>Correlation</td>
</tr>
<tr>
<td>CS</td>
<td>Clear Sky</td>
</tr>
<tr>
<td>DHI</td>
<td>Diffuse Horizontal Irradiance</td>
</tr>
<tr>
<td>DHW</td>
<td>Domestic Hot Water</td>
</tr>
<tr>
<td>DNI</td>
<td>Direct Normal Irradiance</td>
</tr>
<tr>
<td>DR</td>
<td>Demand Response</td>
</tr>
<tr>
<td>EHWH</td>
<td>Electric Hot Water Heater</td>
</tr>
<tr>
<td>GHI</td>
<td>Global Horizontal Irradiance</td>
</tr>
<tr>
<td>GSHP</td>
<td>Ground Source Heat Pump</td>
</tr>
<tr>
<td>HDD</td>
<td>Heating Degree Day</td>
</tr>
<tr>
<td>IRENA</td>
<td>International Renewable Energy Agency</td>
</tr>
<tr>
<td>IRR</td>
<td>Internal Rate Of Return</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>LCOE</td>
<td>Levelized Cost Of Electricity</td>
</tr>
<tr>
<td>LSF</td>
<td>Least Squares Fit</td>
</tr>
<tr>
<td>MAE</td>
<td>Mean Absolute Error</td>
</tr>
<tr>
<td><strong>MetCoOp</strong></td>
<td>Meteorological Cooperation On Operational NWP</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>MPPT</td>
<td>Maximum Power Point Tracker</td>
</tr>
<tr>
<td>NLSF</td>
<td>Non-Linear Least Squares Fit</td>
</tr>
<tr>
<td>NPV</td>
<td>Net Present Value</td>
</tr>
<tr>
<td>NREL</td>
<td>National Renewable Energy Laboratory</td>
</tr>
<tr>
<td>NWP</td>
<td>Numerical Weather Prediction</td>
</tr>
<tr>
<td>POA</td>
<td>Plane-Of-Array</td>
</tr>
<tr>
<td>PSH</td>
<td>Pumped-Storage Hydroelectricity</td>
</tr>
<tr>
<td>PV</td>
<td>Photovoltaics</td>
</tr>
<tr>
<td><strong>PVGIS</strong></td>
<td>Photovoltaic Geographical Information System</td>
</tr>
<tr>
<td>QC</td>
<td>Quality Control</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root-Mean-Square Error</td>
</tr>
<tr>
<td>RTP</td>
<td>Real-Time Pricing</td>
</tr>
<tr>
<td>SAPM</td>
<td>Sandia PV Array Performance Model</td>
</tr>
<tr>
<td>SPA</td>
<td>Solar Position Algorithm</td>
</tr>
<tr>
<td>STC</td>
<td>Standard Test Conditions</td>
</tr>
<tr>
<td>SW&lt;sub&gt;net&lt;/sub&gt;</td>
<td>Shortwave Net-Radiation</td>
</tr>
<tr>
<td>T&lt;sub&gt;air&lt;/sub&gt;</td>
<td>Air Temperature (2 meter height)</td>
</tr>
<tr>
<td>ToU</td>
<td>Time-of-Use</td>
</tr>
<tr>
<td>TSO</td>
<td>Transmission System Operator</td>
</tr>
<tr>
<td>UTC</td>
<td>Coordinated Universal Time</td>
</tr>
<tr>
<td>VAT</td>
<td>Value Added Tax</td>
</tr>
<tr>
<td>VPP</td>
<td>Virtual Power Plant</td>
</tr>
<tr>
<td>WS</td>
<td>Wind Speed (10 meter height)</td>
</tr>
</tbody>
</table>
This thesis consists of an overview and of the following publications which are referred to in the text by their Roman numerals.


Author’s Contribution

Publication I: “Photovoltaic system modeling: A validation study at high latitudes with implementation of a novel DNI quality control method”

The author wrote the manuscript, acquired the data, developed and tested the PV output conversion model, and produced all the calculations, results, and analyses. Antti Poikonen established the data collection for the PV output data and the PV-related weather observations. Antti Aarva installed the PV module temperature and solar radiation sensors, and maintained the radiation observations, helping also in determining the metadata and site horizons for the radiation measurements. Tero Mielonen and Mikko Pitkänen helped with the upkeep of the Kuopio site, and gave feedback on the manuscript. Anders Lindfors supervised the study.

Publication II: “Site-specific adjustment of a NWP-based photovoltaic production forecast”

The author wrote the manuscript, acquired the data, developed and tested the methods, and produced all the calculations, results, and analyses. Anders Lindfors supervised the study.

Publication III: “Ground source heat pump control methods for solar photovoltaic-assisted domestic hot water heating”

The author acquired the NWP ensemble forecast data and produced the actual PV output forecast, also writing the forecast description, results, and analysis. The author also helped in writing and revising the manuscript. Jere Knuutinen had the main responsibility for writing the manuscript.
He developed the control strategies and carried out the main analysis related to the results, implementing the control strategies together with the produced PV output forecast. Other co-authors gave feedback during the development of the study, and helped in revising the manuscript.

**Publication IV: “Residential solar power profitability with thermal energy storage and carbon-corrected electricity prices”**

The author acquired the weather observation data and produced the PV output time series, also writing the description regarding the PV output conversion model. Hannu Huuki and Santtu Karhinen developed the main model, conceived the experiments, and analyzed the main results, also writing the manuscript with the input of the author and other co-authors.

**Publication V: “Utilizing the flexibility of distributed thermal storage in solar power forecast error cost minimization”**

The author acquired the NWP forecast data and produced the PV output forecast, also writing the description regarding the the PV output forecast. Hannu Huuki and Santtu Karhinen developed the main model, conceived the experiments, and analyzed the main results, also writing the manuscript with the input of the author and other co-authors.
1. Introduction

A solar cell is a device that utilize the photoelectric effect to convert light into electricity. Although the first solar cell was constructed by Charles Fritts already in the 1880s [61], development of solar technologies stagnated in the early 20th century due to an increasing availability of coal and petroleum. Climate change predictions have, however, accelerated the development of new approaches for carbon intensive sectors such as electricity production, as 25% of the global greenhouse gas emissions are generated by electricity and heat production [57]. Advancements in solar and wind energy technologies during the two past decades, partly motivated by the climate targets, have created a premise for these technologies to become worthy options for electricity production. Leaps in solar cell technology and more efficient cell production methods, combined with large-scale production volumes, have lead to vastly reduced manufacturing costs, resulting in the wide-spread adoption of photovoltaic (PV) technology. This has enabled PV to emerge from a niche market of small-scale applications towards becoming a potential mainstream electricity source (Fig. 1.1).

Figure 1.1. Electricity generation by source in Europe (1990–2018). Figure from [28].
The global installed PV capacity has experienced nearly exponential growth in recent years, increasing from 40 GWp to 580 GWp during 2010–2019, the total average installation costs dropping from 4702 $/kWp to 995 $/kWp, respectively [30]. In Finland, a more or less exponential growth in the total installed PV capacity has also been present during the recent years, the total installed PV capacity reaching 214 MWp in the end of 2019 [3]. Accordingly, the Levelized Cost of Electricity (LCOE) of PV has dropped dramatically in recent years. For newly commissioned projects, the global weighted-average LCOE of utility-scale PV fell by 82% from 378 $/MWh to 68 $/MWh during 2010–2019. This reduction has primarily been driven by the decline in PV module prices, which have fallen by around 90% since 2010 [30]. Utility scale PV systems are projected to dictate also in the future, however, the total installed capacity of distributed PV is expected to grow faster, driven by supportive measures and policies, together with further consumer commitment [29].

1.1 Background

Despite the swiftly increasing share of PV systems in electricity production, the widespread adoption of PV still pose a handful of challenges to be overcome. In order for an efficient large-scale PV integration to take place, the intermittent nature of PV power output needs to be addressed. Information on the intermittency is relevant for various interfaces where accurate information regarding the power output is crucial for their optimal operation. Such interfaces are, e.g., the power grids and markets, smart grids, and building automation systems (BASs). In order to tackle this challenge, different approaches for mitigating the PV output intermittency effects have to be developed. Short-term PV output is however affected by a multitude of aspects, the most essential ones being the diurnal cycle of the sun, as well as the ever-changing weather conditions, thereby making skillful modeling and forecasting of PV system output easily a complex exercise.

As the intermittent PV production gains a considerable proportion of the electricity generation within a considered area, an apparent fiscal incentive for implementing a skillful large-scale PV output forecast becomes obvious [78]. Incentives regarding forecasting and monitoring of small-scale PV systems are, on the other hand, induced from the benefits of local electricity usage optimization [69, 83]. Several solutions for acquiring a more detailed picture regarding the real-time and future output of a PV system can be found from the literature, comprising of a diverse selection of monitoring, modeling, and forecasting approaches. These methods vary in required input and serve distinct functions, covering spatially from regional to local, and time-wise from real-time monitoring and nowcasting to days-ahead
forecasting, and climatological resource assessments [4, 48].

1.2 Objectives and scope

In recent years, large-scale PV output modeling has been the dominant trend, motivated partly by its apparent contribution potential on mitigating the described challenges associated with wide-spread PV integration. The focus of this doctoral thesis, however, is mainly targeted towards providing new information regarding site-specific aspects of PV output modeling. As clearly shown in some of the thesis results, this point of view is an important aspect of the PV output modeling entity. The Nordic context of this thesis also provides a viewpoint, not commonly covered within this research area.

The objective of this thesis is to develop and verify modeling methods for monitoring and forecasting the output of specific PV systems, in addition to demonstrating their utilization in a Nordic context. The added value of the developed tools are demonstrated by utilizing them in several real-world applications, such as PV output monitoring, economical viability assessments, electricity system balancing, and BASs.

This thesis focuses only on the crystalline silicon (c-Si) PV module technology. Although several different PV module technologies are available, the c-Si technology continues to dominate the markets. The production share of c-Si technology in terms of manufactured solar modules was 95% in 2019, its total cumulative share of produced PV modules being 93% [62].

1.3 Thesis outline

The thesis is organized, as follows. Chapter 2 introduces the large-scale implementation challenges and implications of PV on the functions of the society, also justifying the demand for the research in this field. Chapter 3 is a literature review of PV output modeling, demonstrating the different approaches, applicable datasets, and applications, utilized within this domain. The work and results of this thesis are covered in Chapter 4. Within the literature review and results, the main focus is set on the themes closely related to this thesis. Finally, the thesis is concluded in Chapter 5.
2. Intermittent PV generation in the power grid and markets

The output of solar-powered energy sources are intermittent and weather-dependent by nature, each affected by a somewhat unique set of production factors. The diurnal cycle of the sun, together with the local solar radiation climate, are the two main characteristics defining the average output potential of solar energy. The diurnal cycle induces its own challenges and boundary conditions for solar energy, however, its behaviour is very predictable. Solar radiation, instead, has a high and difficult to predict short-term variability, comprising the potential for rapidly changing conditions for solar energy production. This variability, together with the boundary conditions determined by the diurnal cycle, are the two main challenges for large-scale solar energy integration, whether the challenges are related to balancing the real-time power supply with demand regarding the whole electricity system or a small part of it, or predicting the intermittent future output of the considered entity.

2.1 Power generation and power grid

Electricity supply must be equal to electricity consumption at all times. Base load is the minimum level of demand over a specific span of time in a power grid. The traditional base load generation consists of thermal power plants and hydro power. The remaining fluctuating demand on top of the base load is covered by dispatchable, quickly responding generation sources. The order of utilization (merit order) for different generation types is generally determined by the marginal cost of each electricity source [54]. Although apparently dispatchable, the traditional coal and nuclear plants are usually designed to run as base load power plants, and it may take hours to days for them to cycle off and on again [22].

The dispatchable sources, nowadays commonly utilized for balancing the output of the intermittent renewable electricity sources, have distinct power modulation and response time properties. These, combined with different ramp-up costs, results them being variably suited for acting as
Intermittent PV generation in the power grid and markets

a reserve or balancing power source [23]. In addition to the utilization of dispatchable sources, also curtailment of dispatchable power plants, i.e., the deliberate reduction of their power output, is utilized in order to meet the demand at each given time (Fig. 2.1).

Figure 2.1. Estimated power demand in Germany over a week in May 2012 (left) and May 2020 (right). Figure from [11].

The capabilities of dispatchable generation allow for load and peak matching, lead-in times, frequency regulation, and backup for base-load generators [23]. Dispatchable sources can be divided between load following power plants, peaking power plants, and energy storage. Load following plants are generally located between base load and peaking power plants by means of efficiency, start-up and shut-down speed, construction cost, cost of electricity, and capacity factor (CF). Peaking power plants generally run only when there is a high demand for electricity. Due to the fact that they supply power only occasionally, the supplied power calls for a much higher price per energy unit when compared to base load power.

The increasing penetration of intermittent electricity sources and the bi-directional nature of local electricity production induce increased flexibility requirements for the power generation and the power grid. In general, the total integration costs of intermittent production depend on the current conditions of the power grid and power system in its entirety, making some systems initially more robust and ready for the implementation of intermittent electricity production.
2.2 Electricity markets

Electricity can be considered a commodity. Electricity markets differ however substantially from conventional commodity markets. The difference is in the physical characteristics of electricity, being that large volumes of electricity cannot, as of now, be stored economically, electricity flows cannot be easily and efficiently controlled, and that electricity poses the ability to change in supply and demand at a short notice. These properties induce challenges for balancing the electricity markets with an increasing share of intermittent electricity production (Fig. 2.2).

Figure 2.2. Electricity production and prices in Germany (one week in July 2020). Negative electricity prices occur during the weekend, accompanied by high wind and solar output and a lowered load. The zero price level is being highlighted as a black horizontal line. Figure from [16].

An electricity market is a system enabling purchases, sales, and short-term trading, bids and offers usually utilizing the supply and demand principles to determine the price [27]. The commodities within an electricity market generally consist of two types: power and energy. Power is the metered net electrical transfer rate at any given moment, while energy is the electricity flowing through a metered point over a given period.

Electricity markets are currently transitioning from a one hour imbal-
Intermittent PV generation in the power grid and markets

ance settlement period towards a 15 minute time resolution, partly mot-
vated by the increased penetration of intermittent electricity sources [13].
The 15 minute imbalance settlement period will advocate power producers
to actively participate on the intraday markets.

2.3 Local PV generation

Photovoltaics is one of the most scalable electricity production methods.
This enables PV implementations to range anywhere from single modules
to huge PV array parks, large enough to be seen from space. Due to its
easy scalability, PV has also been widely adopted by households, as well as
industrial and commercial consumers. Local electricity generation and its
bi-directional nature poses, however, its own challenges. Local generation
decrees the demand seen by the power grid and the power markets, ne-
cessitating adequate measures to be taken by regions where a larger local
generation share is present. Production of excess local PV electricity is
however often not economically viable in countries where no feed-in tariffs
are in place for PV. In such cases, the own share of PV output utilization
should be maximized. This can be done either by utilizing a BAS together
with energy storage, or by simply implementing a PV system which does
not produce any excess electricity. Predictability of PV output could there-
fore provide economical benefits also for local generation, whether related
to maximizing the self-utilization of the generated PV output, or for deter-
mining the optimal overall economical end result, whatever the effectual
local electricity market conditions and implemented PV system size may
be.

2.4 Solutions for tackling the challenges

Due to the constant requirement for balancing supply and demand of elec-
tricity in the power grid, countries with an increased share of intermittent
power production have generally been obliged to increase the availability of
dispatchable power generation, e.g., by increasing the gas-based generation
potential, in order to match up with the increased share of intermittent
production (Fig. 2.1). Another potential option to mitigate the effects of
vast intermittent power integration is the possibility to efficiently trade the
imbalances with neighbouring power grids. All of the Nordic Transmission
System Operators (TSOs) are involved in the Nordic balancing energy
market, an aspect that, e.g., Denmark has been able to efficiently exploit
by utilizing the vast hydroelectric capacity of Norway and the nuclear
base power potential of Sweden [5]. This has enabled Denmark to vastly
increase its intermittent generation share, without the need of applying
a similar reserve capacity of dispatchable options, when compared with some other European countries with a similar intermittent penetration share (Fig. 2.3).

![Electricity production share by source between 1990–2019 in Germany (top), Denmark (middle), and Spain (bottom). The figures do not take imported and exported electricity into account. Figures from [28].](image)

The most utilized short-term reserve and balancing power sources for balancing intermittent generation are pumped-storage hydroelectricity (PSH), modern Combined Heat and Power (CHP) plants, utilizing either natural gas, coal, or a range of different biomass options, and other specific
Intermittent PV generation in the power grid and markets

gas power plants [9, 44]. The development in energy storage technologies is also projected as one of the key factors for providing balancing capacity in the future. Although the share of new storage technologies, such as electrochemical and thermal storage, has been steadily increasing during the 2020s, 95% of worldwide energy storage capacity is still covered by PSH [70].

A somewhat easily achievable and obvious way to mitigate the intermittency of renewable energy sources is to diversify the renewable technology types and their geographical distribution. Output forecasting is also seen as one integral part of the solution framework. Combining these different approaches with demand response (DR) and curtailment of intermittent production can provide an energy system that can more reliably match real-time energy demand [68].

As a clear monetary incentive for mastering PV output modeling in order to mitigate the negative effects of wide-spread integration of intermittent PV generation is obvious, a great emphasis has been directed towards undertaking each of the separate challenges related to PV output modeling. This is a feature also seen in the research literature as a multitude of different approaches. Despite this, no specific best practices have yet been established for handling the PV output modeling aspects, as a whole.
3. Photovoltaic output modeling

As described in Chapter 2, PV output modeling has the potential of providing benefits for a range of different scale applications. In general, PV output modeling serves as a useful tool for local site assessments and plant monitoring. PV output forecasting, on the other hand, has been proven to be an effective tool for the power grid and markets, operating on a regional grid level. Potential future benefits can, however, also be seen for microgrids and individual buildings.

The causality between ambient weather conditions and PV output is to be fully understood in order to develop a competent PV output model (Fig. 3.1). The main factors that affect the relative efficiency ($\eta_{rel}$), i.e., the PV module efficiency relative to Standard Test Conditions (STC), are absorbed solar radiation ($G_{POA,absorbed}$) and module temperature ($T_{module}$) [4, 26, 65]. STC is an industry standard to illustrate the performance of PV modules at a cell temperature of 25 °C, when predisposed to an irradiance of 1000 W/m$^2$ with an air mass of 1.5. These conditions correspond approximately to solar noon conditions near the spring and autumn equinoxes in the continental United States, with the module surface aimed directly at the sun [77].

The behavior of relative efficiency is highly nonlinear in low to mid-range solar radiation conditions. The relative efficiency has an increasing trend with increased absorbed radiation, in case the module temperature can be considered a constant (Fig. 3.2(a)). In real life, however, PV module temperature is highly affected by the amount of absorbed radiation, resulting in a reduced efficiency as the module heats up (Fig. 3.2(b)), and the relative efficiency to peak already below mid-range solar radiation, in case ambient temperature and wind conditions are considered as constants (Fig. 3.2(c)).

PV module temperature is generally defined by absorbed radiation, ambient temperature, and ambient wind conditions, as well as the installation method (and type) of the module [38]. The distribution between direct and diffuse radiation as well as the solar angle of incidence (AOI) and module slope angle, together with PV module specific characteristics, affect the total reflection losses of the PV module (Fig. 3.3) [46]. The reflection
Photovoltaic output modeling

losses determine the share of the impinging radiation that is absorbed into the module, the losses generally increasing with an increasing AOI. The amount of ground-reflected diffuse radiation depends highly on the ground albedo, whereas the slope angle of the PV module affect the potential for this ground-reflected radiation to impinge the module, and potentially be absorbed by it. This is due to the fact that the ground-reflected radiation losses decrease with an increasing slope angle. Diffuse radiation reflection losses, on the other hand, do not generally vary much for different orientations.

The impinging radiation amount is determined by the position of the sun, the prevailing weather conditions, and the orientation of the PV module [60]. Site-specific shadowing conditions, affected by the position of the sun, as well as the distribution between direct and diffuse radiation, also play a role. When considering the PV system in its entirety, also the system-specific efficiency aspects influence the output, the inverter efficiency generally behaving similarly to the $\eta_{rel}$ curves in Fig. 3.2(a) as a function of inverter capacity [56].

Figure 3.1. A schematic representation of the main factors affecting PV output. Here, a simplified separation has been made between environmental and system-specific factors. The main factors that determine the relative efficiency of the PV system are highlighted in orange.
3.1 Approaches

A range of energy conversion modeling approaches for estimating PV output can be identified from the literature. These approaches can be generally categorized into physical, parametric, and machine learning (ML) methods.

3.1.1 Physical models

Physical methods are, by definition, derived from the physical principles that govern PV electricity generation, trying to explicitly imitate the physical behaviour of the system [45]. In physical modeling, a basic single...
diode equivalent circuit model is often utilized. The equivalent circuit model defines the I-V curve of a PV cell, module, or array, for a given set of operating conditions. The complexity of these models vary generally between three, five, and seven estimated parameters [24].

The potential benefits of the equivalent circuit model rises from its explicit and thorough capability to describe the physical behaviour of the system. The drawbacks are that some module-specific experimental data is always required, resulting in the minimum prerequisite of at least knowing the specific manufacturer and model of the studied PV module [40].

3.1.2 Parametric models

Parametric models are here defined as simplified estimates of the physical systems, being more implicit approaches when compared with physical models. The nature of these models is to determine and parameterize the relationship between the relevant production factors, such as absorbed solar radiation and module temperature, with the actual power output of the PV system. Examples of such models can be found, e.g., in [26] and [33].

This approach removes the constraints of utilizing module-specific information, and enables a more generalized modeling approach. The potential drawback is that one generic fitting might not cover all the studied PV module setups as well as required.

The Publications presented in this thesis focus solely on the parametric PV output modeling approach.

3.1.3 Machine learning models

The methods under machine learning (ML) constitute a large set of different algorithms, where the training of the model through experience, i.e., training data, is a common denominator [2]. ML methods enable predictions or decisions without being explicitly programmed to do so. Some of the ML methods are closely related to computational statistics, however, not all ML is statistical learning.

The training of these models is usually implemented by utilizing a recent historical dataset of on-site PV output data, preferably together with on-site radiation and ambient temperature measurements, among other potentially relevant measured parameters [2]. The benefits of ML is that no PV site metadata (see Section 3.2.1) is mandatory. The accurate selection of the model architecture is, however, crucial for reaching a satisfactory end result.

One of the major drawbacks of ML is the black box nature of the approach, meaning that explicit diagnostics of the logic and output of the model is often hard or impossible to attain. Sophisticated ML-based PV
output conversion models may gain a modest edge in precision when compared to other simpler methods, however, the difference is usually more or less insignificant. As a trade-off, ML requires the utilization of an implementation-heavy black box approach, by definition [18].

3.2 Input data

As described previously in this Chapter and highlighted in Fig. 3.1, the two most crucial factors that determine the output and relative efficiency of a PV system are absorbed solar radiation and PV module temperature. The module temperature, in addition to module and installation-specific properties, can be determined by absorbed solar radiation, ambient temperature, and wind speed. Altogether, solar radiation can be considered the parameter with the largest interest, when considering the input data of the model.

The input data type for the PV output model is selected based on the application target of the model. Real-time monitoring applications utilize in-situ measurements, while forecasting applications require the selection of a specific forecaster, such as a Numerical Weather Prediction (NWP) model [2, 48]. Climatological timescale estimates, on the other hand, can be attained through utilizing longer satellite or ground measurement time series.

The fluctuating conditions and losses encountered by the PV system are implicitly manifested within the measured PV output data. If available, PV measurements can be generally utilized for configuring the models in order to acquire a more detailed depiction of the investigated PV systems.

3.2.1 PV system metadata

Due to the fact that PV output is location and weather dependent, specific information regarding the assessed PV site is generally required as model input. This metadata usually comprise of at least the geolocation, PV array capacity, and its orientation [36]. In addition, more specific information regarding the PV and installation type as well as inverter capacities can be useful. Depending on the available set of metadata parameters, a part of them, such as the orientation of the PV array, can also be estimated to a certain degree by various methods, such as the one described in [35]. Regarding ML methods, metadata information can be implicitly configured within the model logic without the mandatory requirement of any metadata input [2].
3.2.2 In-situ measurements

In-situ measurements are here considered as PV site-specific weather and PV-dedicated observations. As described above, these measurements are generally used for real-time monitoring or model configuring purposes. Analogous with the two most crucial factors determining PV output, PV array oriented global radiation and module and/or ambient temperature are the two most useful in-situ measurements. Also the distribution between direct and diffuse radiation is important regarding the accurate estimation of the PV output. This ratio can, however, be estimated to a certain degree from global radiation with a selection of different decomposition models [21]. Dedicated, high-quality direct or diffuse radiation measurements can be expensive and require generally systematic maintenance. These measurements are thereby usually not implemented for smaller, generic PV installations. In recent years, sky cameras have also established themselves as a noteworthy device for PV output modeling [59], however, they are mostly utilized for short-term forecasting purposes (see Section 3.2.4).

3.2.3 Satellite measurements

Remote sensing methods, more specifically satellites, can be utilized either for short-term forecasting, or for providing long-term climatological assessments regarding planned or operational PV sites. The Photovoltaic Geographical Information System (PVGIS) web service [12] is one of the most popular satellite data-based services, where different PV site assessment tools are freely available. PVGIS utilizes data from several geostationary satellites and reanalysis datasets. As for short-term forecasting, multiple different remote sensing-based approaches are available [82].

3.2.4 Forecaster selection

In PV output forecasting, the most challenging aspect is not related to the accuracy of the selected energy conversion model or approach, but to the application of a skillful forecaster, to be utilized as input data for the selected approach. The selection of the best suited forecaster usually depends on the application. Four main factors can be established for determining the most suitable forecasting approach: required time-horizon and temporal resolution, geographical location, prevalent weather conditions, and the availability and quality of the data [48]. The available forecasters can be divided, e.g., into the following categories [48, 53]:

- **Physical** (sky imagery, satellite, and NWP)
- **Statistical or probabilistic** (e.g., regression models)
• **Machine learning** (e.g., artificial neural networks; ANN)

• **Hybrid methods**

Combinations from the three first categories are generally considered hybrid methods. One example of a hybrid method is the combination of ML and statistical or physical approaches, which generally result in an enhanced forecasting performance. However, all of the categories have specific and complementary features, and the appropriate method selection depends on the application and preferred outcome [32, 48, 53]:

• **NWP** accuracy depends on a multitude of different factors, one common aspect being the stability of the weather conditions. Historical PV output data is not a prerequisite for utilizing the NWP model. However, their potentially limited spatial availability, in addition to the fact that the first hours of NWP forecasts are usually not especially useful for solar radiation forecasting, can be considered potential drawbacks.

• **Satellite** data can be utilized for short-term forecasting, up to 6 h ahead.

• **Sky-imagery** has forecasting potential in sub-hourly time frames.

• **Statistical and ML methods**, by themselves, are primarily applicable for short-term intraday forecasting. ML approaches usually merge the processes of forecasting and energy conversion into one input/output interface, differing from the physical and parametric approaches where the energy conversion is a distinct and separated process from forecasting.

**Numerical Weather Prediction**

NWP is the selected forecaster type utilized in this thesis, producing forecasts ranging anywhere between nowcasting and week-ahead timescales. Different countries and institutions use different kind of NWP models, generally varying in spatial and temporal coverage, as well as the underlying model physics. Due to their designed purpose and their established position to function as the global primary weather prediction method, NWP models have also obvious potential to be utilized for forecasting the weather-dependent PV output. NWP models are thus commonly used for day-ahead PV output forecast applications [2, 82].

In recent years, emphasis has been put on developing the critical issues regarding solar radiation forecasting, namely cloud and cloud property forecasting. These topics are generally considered also one of the most challenging facets of weather forecasting [1, 20, 58]. NWP forecast enhancements have been targeted by means of updated model physics, by the utilization
of additional cloud information, and with the implementation of additional post-processing. In order to be implemented, the post-processing methods generally require either external satellite or ground-derived data (model adjustment, i.e., bias correction) or a collection of adjacent NWP grid points (spatial averaging over adjacent grid points) [43]. A range of solar radiation post-processing methods related to, e.g., producing regression-based probabilistic radiation forecasts, can be discovered from the literature [6, 47]. Although having a slight trade-off in accuracy when compared with ground-measurements, satellite data has also become a very valuable data source for NWP post-processing, vastly increasing the available spatial coverage of ground measurements [81].

In recent years, preliminary assessments regarding the feasibility of ML methods as a replacement for NWP models have been produced. It can however be concluded that the development of ML methods, potentially substituting or at least adding value to the traditional NWP approaches, are facing similar challenges as the development of NWP modeling has faced over the years [10].

3.2.5 Uncertainty of input data

Whether measurement or forecast data is used, specific uncertainties are always present. The most relevant uncertainties, divided by the utilized input data sources are:

- **PV site metadata**: General ambiguities, uncertainties, and missing data is common, especially considering the more specific but potentially as important site details.

- **In-situ measurements**: Sensor accuracy, depending on the calibration and maintenance schedule requirements and implementation, being especially critical regarding the more sophisticated radiation measurements.

- **Satellite data**: Challenges in the definition of cloud column properties within the image pixel area, potentially worsened by image projection challenges [32].

- **Numerical Weather Prediction**: Overall uncertainty of NWP forecasts, induced both by the model’s incomplete physical depiction of the real-world atmospheric processes, as well as by the uncertainties regarding the initial state of the forecast.

*Environmental conditions and geographical influence*

The geographical location is the main factor that determines the boundary conditions for the prevalent solar radiation climate. However, local
geographical effects, such as orography or the land-sea distribution also plays a role [39]. The geographical features and the latitude affect both the predictability of NWP models, as well as the potential data coverage of satellites. High latitudes with snowfall have additional distinct challenges, snow cover being a restrictive factor for PV output generation. As for the local conditions, the local horizon features of each PV site affect the shadowing conditions, imposed onto the PV array.

**Quality control**

In order to mitigate the effect of the described uncertainties and errors, the implementation of varied quality control (QC) approaches are possible. The development of QC methods regarding in-situ weather measurements have already a long development history, and a somewhat general framework has also been established for radiation-specific QC [41]. An initial QC framework for targeting the somewhat common uncertainties and errors found in measured PV output data has also been established, e.g., by [35]. With a similar framework, rough errors in PV metadata can also be detected.

However, due to the generally relaxed boundary conditions usually imposed by QC methods for cleaning only the evidently erroneous values, traditional QC boundaries and approaches might not always fit directly as such for defining whether the investigated data is good enough to be applied as input data for other applications.

### 3.3 Applications of PV output modeling

PV output modeling applications can be divided into separate categories in different ways. Separating factors could be, e.g., the distinction between regional and local approaches, or the utilized time frame of the application, the latter one being selected here as the distinguishing characteristic.

When modeling on a regional level, forecasting the output of every single investigated PV system is usually impossible. Thereby, a typical approach is to implement and upscale a regionally representative subset of PV sites, from which local characteristics have been eliminated or are absent [34, 43]. On the other hand, in order to achieve a detailed depiction of a specific set of PV systems, the occurring properties and losses of the investigated systems are of main interest in the modeling process.

#### 3.3.1 Site planning and output monitoring

The yield assessment of a planned PV site is the first step in PV site implementation where output modeling has potential to provide added value [50, 67]. A selection of commercial and non-commercial tools have been
established for assessing the feasibility and output potential of planned PV arrays. Various site-specific loss factors, depending on the software and available data, can be estimated for the planned investment. These tools can be utilized both for small and utility-scale sites, utility-scale investments naturally being the ones with a more rigorous and emphasized planning phase.

Planning and assessment tools can also be utilized as a research tool, assessing production, feasibility, or economical factors of photovoltaics. The research application potential vary, based on the available input data, as well as the tool in question.

After array commissioning has taken place, monitoring of the site’s performance is important in order to ensure optimal performance [19]. An important factor is also the monitoring or nowcasting of the grid-level PV output. Grid monitoring is especially valuable for the power grid TSOs.

### 3.3.2 Output forecasting

PV output forecasting has a crucial role in enabling an effective integration of intermittent PV production into the energy system, both on a local and a regional level [74]. Regional power grid-level PV output forecasting is already a common practice in many countries where PV has emerged as a significant electricity production source [63]. PV output forecasting is also a relevant research topic on a local level regarding smart grid and BAS applications, commercial solutions being tested for different interfaces and developed by several companies. The significance of such applications may increase in the future when further mitigative measures, related to topics such as real-time DR or energy storage, will be implemented.
4. Results

This Chapter describes the main work and results, carried out within the publications included in this doctoral thesis. The main focus of this Chapter is set on the topics closest related to the subject of this thesis, i.e., PV output modeling and forecasting, together with PV profitability and relevant application possibilities. Publication I focuses on the description, validation, and utilization of the selected PV output conversion model, together with a novel QC method for handling calculated Direct Normal Irradiance (DNI) values, deduced from Global Horizontal Irradiance (GHI) and Diffuse Horizontal Irradiance (DHI) measurements. Publication II describes a straightforward site-specific adjustment approach for the PV output conversion model, implemented with NWP forecast data as input. Publication III investigates the utilization of such a PV output forecast as a tool for optimizing domestic hot water (DHW) heating costs in a household equipped with a PV system and a ground source heat pump (GSHP). Publication IV investigates small-scale PV profitability in Finland, together with thermal storage and carbon-corrected electricity prices. Lastly, Publication V investigates the utilization of distributed thermal storage in PV output forecast error cost minimization.

4.1 QC method for handling calculated DNI values (Publication I)

DNI requires the most sophisticated measurement instruments of the three main solar radiation components. These measurements are therefore not always carried out at solar radiation sites. This was also true for one of two studied PV sites in Publication I, and for two of three radiation measurement sites, utilized for Publication IV. DNI is, however, an essential parameter in PV output modeling, and thus, a novel QC method was introduced for handling the volatile and uncertain nature of calculated DNI. DNI is defined as:

\[
DNI = \frac{GHI - DHI}{\sin(\alpha)}
\]  

(4.1)
where GHI is Global Horizontal Irradiance, DHI is Diffuse Horizontal Irradiance, and $\alpha$ is the solar elevation angle. This shows that the calculated DNI is very sensitive to measurement inconsistencies, especially at low solar elevation. Another factor to point out is the fact that calculated hourly DNI generally has a rather large positive bias in low solar elevation, when compared to observed DNI [8].

For establishing and evaluating the QC approach for calculated DNI values, five years of hourly data from three Finnish high quality solar radiation sites, having virtually obstruction-free horizons, were utilized. For solar position calculations, the so-called "half-hour" timestamp, i.e., the middle of the hour ($t \pm \frac{1}{2}h$), together with the National Renewable Energy Laboratory (NREL) Solar Position Algorithm (SPA) [66], was used. This approach was utilized for solar position calculations throughout this thesis.

The positive bias for the calculated hourly DNI regarding the three test sites (with $\alpha \geq 0^\circ$) was 27.1 W/m$^2$, corresponding to approximately 120 kWh/m$^2$ on an annual level. This is equivalent to a 2-axis tracking PV system in Southern Finland, gaining roughly a half summer month of plane-of-array (POA) irradiation [12].

Observed DNI tend to create an exponential envelope shape, when plotted as a function of solar elevation angles (Fig. 4.1). An exponential fit was therefore utilized for determining the maximum elevation angle dependent limit for calculated DNI. The resulted function was achieved by binning the data according to the solar elevation, and imposing a Non-linear Least Squares Fit (NLSF) for the bins’ maximum values. Two separate bin widths were implemented, due to the bin trade-off between an adequate amount of observations per bin, and an adequate restriction of values, especially at lower elevations. The site with the lowest maximum solar elevation angle determined the maximum elevation angle included in the QC fit, resulting in a limit of 46 degrees. In order to estimate the QC performance regarding one minute measurements, the resulting fit was also validated with one minute DNI observations from the southernmost site, spanning over two years.

The established DNI QC limit check was determined as:

$$DNI_{QC_{limit}} = A \times \exp(B \times \alpha) + C \quad [\text{W/m}^2]$$

where:

$$A = -838$$

$$B = -0.112$$

$$C = 951$$

$$DNI_{QC} = \min(DNI_{calculated}, DNI_{QC_{limit}})$$
where $DNI_{QC\text{limit}}$ is the maximum limit for DNI, $DNI_{QC}$ is the value chosen by the QC, and $\alpha$ is the solar elevation in degrees. By implementing the DNI QC to the hourly five year radiation dataset from the three radiation sites, the initial bias was reduced by almost 50%. The root-mean-square error (RMSE) was reduced over 80%, whereas mean absolute error (MAE) was reduced by 1/3, and correlation was increased from 0.80 to 0.99. The suggested approach was thereby shown to provide a viable generic solution for limiting unrealistic calculated DNI values in the data, as long as representative DNI measurements are available for defining the maximum DNI limits. This is due to the fact that maximum DNI values are latitude and climatology dependent.

![Figure 4.1](image1.png)

**Figure 4.1.** Five years (2013–2017) of hourly DNI observations from three radiation sites in Finland. DNI observations are divided into bins with a resolution of 0.05 ($\alpha \leq 10^\circ$) and 0.2 degrees ($\alpha > 10^\circ$), with bin maximum values plotted in black. The fitted QC limit function is shown in green. Bin sample sizes and fit errors vs. bin maximum values are shown below. Figure from Publication I.

### 4.2 PV output modeling with in-situ observations (Publication I)

The function of a PV output conversion model is to define the causality between PV output and the ambient weather conditions. In this Publication, this relationship was evaluated at two Finnish PV sites (located in Helsinki and Kuopio) with the help of an extensive set of in-situ measurements. The study focused on model validation, by investigating the performance of the PV output conversion model and its sub-components. In addition to validating the standard version of the model, the conversion model was also separately site-optimized for both PV sites by utilizing the available
Results

In-situ measurements.

The studied PV systems were approximately 20 kWp in nominal capacity, both also having 15 degree slope angles (Table 4.1). Both PV systems were equipped with a set of close-by in-situ weather and solar radiation measurements (Fig. 4.2).

Table 4.1. Basic specifications of the PV systems. Table information collected from Publication I.

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Kuopio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal capacity (kWp)</td>
<td>21.00</td>
<td>20.28</td>
</tr>
<tr>
<td>Slope angle (from horizontal)</td>
<td>15°</td>
<td>15°</td>
</tr>
<tr>
<td>Azimuth angle (clockwise from north)</td>
<td>135°</td>
<td>217°</td>
</tr>
</tbody>
</table>

Figure 4.2. Layout of Helsinki (left) and Kuopio (right) sites, where PV arrays are shown as black-red boxes. Nearby in-situ measurements are marked as coloured crosses. Figure from Publication I.

Roughly one year of one minute averaged datasets were used for evaluating the PV output conversion model at both sites. In-situ measurements of air temperature and wind speed, together with traditional solar radiation components, were used as model input. PV output measurements were utilized both for optimizing and validating the model. Although the dedicated POA irradiance and module temperature measurements were utilized for validating the implemented POA transposition and module temperature schemes, these measurements were not used as an input for the actual PV output conversion model.

Regarding site horizons, the POA irradiance measurements were carried out right next to the PV arrays, separating them more or less from the dedicated solar radiation measurements, which were carried out at both sites at approximately 6–8 m higher elevations. The solar radiation measurements in Kuopio were in some ways influenced by a cluttered horizon...
especially regarding the POA radiation instrument at the lower PV module
level, where the southern half of the horizon was obstructed at an elevation
between 5 and 15 degrees, averaging around 8 degrees. As for the Helsinki
PV site, a virtually clean horizon was present, with some minor obstacles
up to an elevation of 3–4 degrees, between west and southwest.

The schematic description of the baseline PV output conversion model,
implemented in this Publication, is displayed with green color in Fig. 4.3.
The baseline modeling process can be separated into four steps:

1. Estimation of POA solar irradiance ($G_{POA}$)

2. Estimation of reflection losses, i.e., determining the solar radiation
   amount absorbed by the module surface ($G_{POA,absorbed}$)

3. Estimation of module temperature ($T_{module}$)

4. Conversion of effective solar radiation ($G_{POA,absorbed}$) and module tem-
   perature ($T_{module}$) into PV output

For step 1, the Perez et al. transposition model [60], generally acknowl-
èged as one of the leading models for estimating diffuse radiation on
inclined surfaces [55, 79], was used. As for estimating the reflection losses
in step 2, the method introduced by Martin and Ruiz [46], was utilized.
PV module temperature was determined by the Sandia PV Array Perfor-
mance Model (SAPM) scheme [38], whereas the final PV output conversion
step was based on the parametric methods presented by Huld et al. [26].
Exhaustive descriptions of these methods can be found from Publication I.

The performance of $G_{POA}$ and $T_{module}$ schemes were evaluated at both
sites by comparing calculated estimates to measured values at instances
with non-zero irradiance. Due to the fact that continuous visual informa-
tion regarding module snow cover was not available, distinction between
snow-free and snow cover periods were estimated by utilizing nearby ob-
served snow depths. The radiation instrumentation was heated at both
locations, shown as a good overall $G_{POA}$ performance throughout the year.
Conversely, the fact that the $T_{module}$ scheme does not account for snow
cover, was clearly seen in the model performance. In Helsinki, the MAE
for $T_{module}$ was 1.8 °C when compared to module average temperatures in
snow-free conditions, with a correlation of 0.981 and a bias of 1.0 °C. As
for all $G_{POA}$ observations in Helsinki, MAE was 10 W/m² or less, with a
correlation of 0.998 and a bias around −7 W/m². Due to Kuopio’s missing
DNI measurements and a somewhat cluttered horizon, together with the
horizon heterogeneity between the PV modules and the designated radia-
tion measurements, the Helsinki site could in many ways be considered
a superior reference site regarding modeling component validation. This
was supported by the fact that, in Kuopio, an evident positive bias in the
Results

estimated $G_{POA}$, not explicitly shown in the figures, was identified from the data, extending to a solar elevation of approximately 16 degrees.

**Figure 4.3.** Modeling chain of the PV output conversion model. The baseline model part, presented in Publication I with in-situ observations as input, is colored in green, whereas the model adjustment part with NWP data as input (Publication II), is shown in blue.
The utilized parametric output conversion model determines a time-dependent relative efficiency $\eta_{\text{rel}}$, which is based on the technology-dependent coefficients $(k_1 - k_6)$ of the studied PV module type. By utilizing $\eta_{\text{rel}}$, the PV system output is determined as the output relative to nominal Standard Test Conditions (STC):

$$PV_{\text{out}} = PV_{\text{out,STC}} \times G_{\text{norm}} \times \eta_{\text{rel}}$$

where $PV_{\text{out,STC}}$ is the output in STC conditions, $G_{\text{norm}}$ is the normalized absorbed radiation, i.e., $G_{\text{POA,absorbed}} / 1000 \text{ W/m}^2$, and

$$\eta_{\text{rel}} = 1 + (k_1 \times \ln(G)) + (k_2 \times \ln^2(G)) + T_{\text{diff}} \times \left( k_3 + (k_4 \times \ln(G)) + (k_5 \times \ln^2(G)) \right) + (k_6 \times T_{\text{diff}}^2)$$

where $\ln(G)$ is the natural logarithm of $G_{\text{norm}}$, and $T_{\text{diff}}$ is $T_{\text{module}} - T_{\text{STC}}$, i.e., $T_{\text{module}} - 25 \degree\text{C}$.

System-specific values for the coefficients $k_1 - k_6$ can always be determined as long as information about $G_{\text{POA,absorbed}}$, $T_{\text{module}}$, and $PV_{\text{out}}$ are available. If the system-specific coefficients are defined, the inverter conversion losses can also be included in the process, due to the fact that inverter efficiency is essentially a function of its power output. The utilization of high quality weather observations as input is, however, crucial in order to acquire a valid calibration. In order for the compared measurements to observe the same physical conditions, it is also important that an identical averaging and temporal resolution is used for all of the utilized input datasets.

The goal here was to optimize the model itself. Therefore, one minute $G_{\text{POA,absorbed}}$ and $T_{\text{module}}$ estimates, provided by the respective model schemes, together with one minute measured $PV_{\text{out}}$, were utilized as fit inputs. The goal for the fit was to imitate optimal system performance. Vague data, based on the following conditions, was therefore cleaned from the input:

1. PV output only between 1% and 99% of inverter capacity was utilized
2. Freezing wintry snow conditions were excluded from the input data
3. For minimizing power output anomalies and uneven shadowing effects, a 10% power output maximum discrepancy limit between inverters’ two Maximum Power Point Trackers (MPPTs), was allowed
4. For removing the most inconclusive events at low solar elevations, in addition to homogenizing the distinct solar radiation and PV output datasets by means of experienced horizons, an apparent solar elevation angle above two degrees, was required

The Levenberg-Marquardt algorithm [49] without any coefficient boundaries was used for the optimization task. Data from both sites were
randomized into training and testing datasets with a ratio of 3:1. The values found from [25] were here used as reference coefficients \((k_1 - k_6)\). The resulting fit was shown to increase model performance at both sites (Table 4.2 and Fig. 4.4) when compared with the reference coefficients, especially reducing the bias. It should, however, be stated that unlike the reference coefficients, the new fit did also take inverters’ conversion losses into account. This might partly explain the larger positive bias, found in the reference coefficient results.

Table 4.2. PV output conversion model performance and new coefficients, fitted for cleaned one minute data. Table information collected from Publication I.

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Kuopio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training set size</td>
<td>123 502</td>
<td>63 240</td>
</tr>
<tr>
<td>Testing set size</td>
<td>41 168</td>
<td>21 080</td>
</tr>
</tbody>
</table>

**Reference [W/kWp] (fitted for PV\(_{DC}\)) [25]**

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Kuopio</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>25</td>
<td>36</td>
</tr>
<tr>
<td>MAE</td>
<td>18</td>
<td>25</td>
</tr>
<tr>
<td>COR</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>BIAS</td>
<td>11</td>
<td>21</td>
</tr>
</tbody>
</table>

**New [W/kWp] (fitted for PV\(_{AC}\))**

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Kuopio</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>18</td>
<td>20</td>
</tr>
<tr>
<td>MAE</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>COR</td>
<td>0.997</td>
<td>0.997</td>
</tr>
<tr>
<td>BIAS</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

**New coefficients**

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Kuopio</th>
</tr>
</thead>
<tbody>
<tr>
<td>(k_1)</td>
<td>0.045866</td>
<td>0.112983</td>
</tr>
<tr>
<td>(k_2)</td>
<td>-0.002035</td>
<td>0.020414</td>
</tr>
<tr>
<td>(k_3)</td>
<td>-0.006095</td>
<td>-0.010906</td>
</tr>
<tr>
<td>(k_4)</td>
<td>-0.000120</td>
<td>-0.003004</td>
</tr>
<tr>
<td>(k_5)</td>
<td>0.000372</td>
<td>0.000809</td>
</tr>
<tr>
<td>(k_6)</td>
<td>0.000004</td>
<td>0.000200</td>
</tr>
</tbody>
</table>

An illustration of the optimized model’s minute-to-minute performance during two specific days is shown in Fig. 4.5. It can be seen that the overall accuracy of the model was very good, however, some percentage-wise larger errors were still found in clear-sky conditions during sunrise and sunset hours. The afternoon errors in Kuopio could be mostly contributed to known shadowing effects. In order to more specifically identify the error sources, modeling error regarding inclined global radiation, PV module temperature, and AOI, was determined by utilizing 15 min averaged data...
(Fig. 4.6). This inspection confirmed that a positive bias at large AOI values, approximately between 50 and 90 degrees, existed at both locations, indicating that the calculated reflection losses might have been partly underestimated at high AOI angles. This was also supported by a separately concluded fact which pointed out that, in Helsinki, the performance of estimated POA radiation was not particularly affected by larger AOI angles, whereas the performance of estimated PV output was.

![Figure 4.4. PV output conversion model performance with one minute data. Respective LSFs for reference and new coefficients in snow-free conditions are shown in red and blue. Data with observed snow cover are shown in gray. Figure from Publication I.](image)

The optimized conversion model did not account for shadowing effects or module snow cover, but gave an estimation of the optimal PV output, without the presence of such performance degrading factors. With the site-specifically optimized models, the general performance of these sites could next be evaluated. By utilizing the model, several days of sub-optimal inverter behaviour was identified during morning hours, especially in Helsinki. This consistent behaviour was most likely due to the lacking performance of the inverter’s MPPT algorithm, which could not find the global power optimum in instances when a clear-sky steady incline in solar radiation was present. It should be noted that, during the period under review, the Helsinki inverter had an older firmware version compared to Kuopio, which could easily explain this behaviour difference between these sites.

The daily and monthly performance of the PV systems was evaluated next. As some breaks were found from the PV output data, a filter was applied before the daily aggregates were generated. This filter required that day by day PV output data availability should account for at least 95% of the time when the sun was above the horizon, in case no snow cover was detected. The days that did not fill this condition were discarded from the comparison.
As expected, winter months with snow cover caused significant differences between observed and modeled production, while snow-free months indicated a somewhat small bias. This result was supported by the fact that no specific data collection issues had been noticed during the inspected period. A large part of these winter losses could therefore be contributed to snow-covered PV modules. The results were in line with general visual observations, where the Kuopio site had a much larger snow cover impact on PV output, affecting into late spring of 2018. Also shadowing and DNI overestimation influenced to some extent for Kuopio, further expanding the discrepancies between the model and the observations. These two factors, however, did not seem to cause substantial bias between the model and the observations. All and all, the approximate snow cover losses for the winter period between 2017 and 2018 were 110 Wh/Wp (Helsinki) and 210 Wh/Wp (Kuopio), corresponding to PV output of August 2017 (Helsinki) and July–Mid-August 2018 (Kuopio). This gives some insight on the potential snow-induced production losses for similar installations in comparable climatological conditions. No generalizable answers can, however, be given regarding the snow-induced production loss potential. The installation type and slope angle affect the extent to which snow can run off from the module surface during the winter period. A higher slope angle also enables a greater share of the ground-reflected radiation to be absorbed by a snow-free module, a positive effect that is accentuated when the ground is snow-covered. Despite the variable impacts of these individual factors, a common denominator for all snow-induced effects is the fact that they
Results
generally revolve around the darkest and least productive part of the year.

4.3 Site-specific adjustment of a NWP-based PV output forecast (Publication II)

In order to accurately depict the behaviour of a specific PV system, the prevalent system-specific losses and features must be taken into account. A practical way to enable this is to incline the model to the influence of measured PV output data, the system-specific features being implicitly manifested within the data. PV power output is often the sole measured variable available from the PV system, and an efficient approach for utilizing it is therefore essential.

In this Publication, a straightforward method for utilizing PV power output for adjusting a site-specific PV output forecast, was presented. The
The dataset consisted of NWP data, PV site metadata, and measured PV power output. Although NWP data clearly had the largest uncertainties among these sources, appropriate filtering was implemented for each separate data source. The used PV output data and relevant metadata were obtained from research institutions and solar energy companies. The provided metadata was iterated and validated by exploiting satellite images and initial output modeling results. Site selection was implemented by estimating the quality of the metadata, PV sites being classified based on the approximate uncertainties regarding different metadata parameters. Sites with ambiguous metadata or a data availability of less than 75% during daylight hours ($\alpha > 0^\circ$) were excluded from the study.

Hourly averaged PV output data from the selected 23 PV sites, ranging between 5 and 850 kWp, was used. A majority of these PV systems were located on flat roofs, slope angles of the modules generally being either 10 or 15 degrees. Some of the installations included multiple module orientations. Generally, the slope angles of pitched roof installations had larger uncertainties, when compared to flat roof installations.

Hourly day-ahead forecast time series (00 UTC model runs, time steps +23 h to +46 h) from the NWP model’s main (control) ensemble member regarding GHI, DNI, Shortwave Net-radiation ($SW_{net}$), 2 meter air temperature ($T_{air}$), and 10 meter wind speed (WS), were utilized. DHI was
Results

determined from GHI and DNI, and albedo was calculated from GHI and SW_{net}. No limitations were implemented for DNI at low solar elevation angles.

The described weather data from site-specific MetCoOp NWP model grid points were fed to the baseline version of the PV output conversion model, the baseline model producing the expected PV power as the output. Before any site-specific adjustments could be implemented, a method for mitigating the effects of NWP data uncertainty was required, in order to assure that measured and modeled PV data represented somewhat indistinguishable meteorological conditions. One way this could be accomplished was by studying only clear sky (CS) conditions. For determining such CS events, a straightforward quantile approach [35, 42], was applied. The adjustment method compared the normalized forecasted and observed PV output quantiles, hour-by-hour, produced from an hourly dataset that covered a defined range of successive days. Here, a time window of 30 days prior to the forecast initialization was chosen for modifying each day-ahead forecast. The quantile approach relied on the assumption that a CS event could always be found for each used time step, when investigating a set of 30 successive days. The 90^{th} percentile was here selected to mitigate the effects of extreme values, caused, e.g., by cloud-enhancement conditions. The selected method relied on the assumption that the 90^{th} percentile of each dataset depicted more or less identical meteorological conditions, as long as each time step had a sufficient amount of data points available. To minimize the possible bias gained from utilizing two separate sources of data, only the occasions when both data sources were available, was used.

Regarding the forecast skill benchmark selection, the most accurate, naive, no-skill forecasting approach should be utilized as a reference [52]. When benchmarking deterministic solar forecasts, climatology and persistence are the two most common naive reference methods [80]. No common standards have however yet been established for PV output forecasts. The persistence forecast, fulfilling the above mentioned definitions, using the measured PV output two days before each point-in-time (PV_{out-48h}), was thereby selected as the reference.

Independent daily adjustment coefficients $C_N$ for each hour $N$ of each day-ahead forecast were established, as follows:

$$C_N = \frac{P{V_{\text{fcst}90}}_N}{P{V_{\text{obs}90}}_N}$$

(4.6)

where $P{V_{\text{fcst}90}}_N$ and $P{V_{\text{obs}90}}_N$ are the interpolated 90^{th} percentiles of the hourly forecasted and observed PV output data, ranging 30 days before the forecast initialization, for each specific hour N. $C_N$ thereby equals the hourly relative PV output bias in CS conditions. The hourly adjusted forecast was acquired simply by dividing each hour of the the baseline
Results

forecast \( (PV_{f_{cst,N}}) \) with the corresponding adjustment coefficient:

\[
PV_{f_{cst,Nadj}} = \frac{PV_{f_{cst,N}}}{C_N}
\]  

The adjustment process, as a part of the baseline model, is shown in Fig. 4.3. Due to the nature of this approach, various external biases and loss factors could be implicitly included to the model configuration. Such factors included shadowing losses, a certain extent of conversion and reflection loss biases, potential NWP biases regarding solar radiation in clear-sky conditions, as well as moderate correction potential of inaccurate metadata. A special emphasis was thereby targeted towards eliminating the errors from the utilized PV site metadata, before using the data as model input.

A handful of obvious PV measurement data outliers were found, resulting in an implementation of a relaxed QC condition for the ratio \( PV_{obs,N}/PV_{f_{cst,100,N}} \), before any calculations took place. All zero value outputs, as well as negative solar elevation output values, were also excluded. To sum up the raw PV measurement data quality, only one of the studied sites had zero values (2% of observations) with \( \alpha > 5^\circ \). The NREL SPA [66] was utilized for all solar position calculations.

The modeled PV output of the baseline model was generally greater when compared to measured output, due to miscellaneous, more or less time-dependent loss factors, not covered by the baseline model (Fig. 4.8). To reduce the probability of implementing completely incorrect corrections, the final adjustment coefficients were thereby limited to a minimum of 0.75. In other words, the adjusted output was allowed to be at most 1/3 greater than the modeled baseline PV output.

The first 31 days of the dataset were left outside the evaluation, and used only for determining the adjustment coefficients (Eq. 4.6). The baseline and adjusted PV output forecasts were compared against measured PV output, the hourly performance being shown in Fig. 4.9. The used aggregate was defined as the aggregated absolute output of all 23 studied sites, divided by their nominal combined output. Regarding the aggregate, only occasions with measured PV output available from all 23 sites, were incorporated. The hourly results indicated that the adjustment clearly reduced the overestimated values, while simultaneously reducing the positive bias. The corresponding error metrics, the persistence \( (PV_{out-48h}) \) included as a reference benchmark, are shown in Fig. 4.10. The persistence was outperformed by the baseline model at most sites by means of RMSE and MAE, whereas the adjusted model clearly outperformed the persistence at all sites. As expected, the baseline model had a clearly larger positive bias when compared to the persistence, however, the adjusted model performed considerably closer to the persistence. The persistence model had by far the worst correlation. Correspondingly for daily performance, a similar
correcting behaviour was present, together with a significantly smaller dispersion both for the baseline and adjusted models. Regarding the hourly and daily aggregates, a small positive bias was still present, nonetheless, it had been largely lowered especially at higher outputs.

Figure 4.8. The 90th percentiles of hourly observed (red) and forecasted (blue) nominal PV outputs (left axis) over the full data period. The respective adjustment ratio $C_N$ is shown in black (right axis). Figure from Publication II.

Figure 4.9. Hourly PV output model performance and respective fits regarding separate and aggregated baseline (red/orange) and adjusted (blue/light blue) forecasts. The black horizontal line represents the maximum hourly observed PV output of individual sites, which could be used as a limiting factor for the modeled inverter behaviour, if needed. Figure from Publication II.
The adjustment coefficient $C_N$ represents CS conditions, while the true shadowing losses are dependent on the distribution between direct and diffuse radiation. Further results, not displayed here, indicated that also the PV model bias was affected by sky clearness. This was consistent with the remark that reflection and shadowing losses are generally decreased as the share of diffuse radiation is increased. A modified proposal of the adjustment method, considering the sky clearness effect, was tested in order to alleviate the CS dependency of the adjustment approach. The approach was however discarded from this study due to the fact that this approach lacked in added value, while adding needless complication to the process. This systematic behaviour was however a noteworthy remark, despite the fact that the persisting uncertainties of forecasting day-ahead solar radiation continue to dictate the performance of the process, altogether.

4.4 Application to assess a GSHP control method for PV-assisted DHW heating (Publication III)

A major problem of PV is the diurnal mismatch between electricity production and demand. One solution to mitigate this effect is to utilize energy storage. Electric hot water heaters (EHWHs) are in many ways ideal for such applications, due to the fact that they are readily available in many detached houses, and that they currently have a much lower LCOE, when
compared with batteries.

In a typical household, DHW heating is one of the most energy-consuming activities, accounting for 15% of total energy consumption [71]. In this Publication, four different DHW heating schemes were studied in a building, equipped with a PV system and a GSHP. The main control method aimed to minimize DHW heating costs, by utilizing day-ahead Nord Pool Spot market electricity price information, together with a PV output forecast, and the measured coefficient of performance (COP) of the modeled GSHP.

A dominant approach for evaluating the value of PV output forecasting in similar studies, e.g., [64, 75], is to utilize historical weather or PV output measurements as "perfect forecasts". However, in order to estimate the true added value of a PV output production forecast, real weather forecasts are required. In addition, in order to define the current financial value gap between the best currently available regional forecast and a perfect theoretical forecast, an ensemble NWP forecast was used. The NWP data from the MetCoOp ensemble system [7, 17], used operationally by several Nordic meteorological institutes, was utilized for estimating the PV output, the PV output forecast being carried out with the methods described in Publication II. The years 2017–2019 were simulated with a perfect PV forecast, while the period between June–September 2020 was simulated with actual NWP data.

The utilized NWP ensemble forecast system consists of 15 distinct members, updating five of the forecast members each hour. From an operational viewpoint, each five member NWP forecast cycle took up to six hours from the model initialization time for the actual forecast to be available. In order to take this operational aspect into account, 30 newest ensemble members were always discarded from this study, and only six hours and older forecasts were utilized. The optimal ensemble size, i.e., the maximum age of the utilized forecasts, was chosen based on the forecasting accuracy of the ensemble mean. Based on this accuracy comparison, a 30-member ensemble was chosen, concluding that 6–12 h old ensemble members were utilized for each forecast.

For the daily updated ensemble mean forecast, two different forecast update schemes were assessed:

1. The forecast is updated daily at midnight (local time)

2. The forecast is updated daily at midnight, with intraday forecast updates at 6 and 12 o’clock (local time)

The generic baseline PV output forecast was determined and adjusted to the specified site (Table 4.3), according to the methods in Publication II. Although the dispersion of the PV output forecast ensemble could have provided some potentially useful information regarding the uncertainty of the forecast, to be somehow incorporated into the operation logic of the
control method, it was deliberately left outside this study, in order to keep the scope of this study under control.

Table 4.3. Basic information regarding the PV systems of the investigated building [37]. Table information collected from Publication III.

<table>
<thead>
<tr>
<th>Orientation</th>
<th>Slope angle (°)</th>
<th>Capacity (kWp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South</td>
<td>26</td>
<td>10.400</td>
</tr>
<tr>
<td>East</td>
<td>26</td>
<td>5.355</td>
</tr>
<tr>
<td>West</td>
<td>26</td>
<td>5.355</td>
</tr>
<tr>
<td>Total</td>
<td>-</td>
<td>21.110</td>
</tr>
</tbody>
</table>

The DHW energy consumption was calculated by assuming a typical four-person household, with a DHW consumption of 200 L per day. The inlet water temperature was set to 10 °C, whereas the outlet temperature was 60 °C. Daily DHW demand was thereby 11.64 kWh. The capacity of the EHWH was 500 L. Electricity consumption of the house was considered a static base load.

DHW heating, based on total costs and the share of PV output in DHW heating, was studied with two different-sized PV systems. The actual 21.1 kWp PV system (Table 4.3) was evaluated as such, in addition to a smaller 5 kWp south-facing PV system. The smaller system was obtained by scaling the actual south-facing 10.4 kWp system. Full details of the system model are given in Table 4.4.

Table 4.4. System model details. Table information collected from Publication III.

<table>
<thead>
<tr>
<th>Object</th>
<th>Quantity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily DHW demand (m³)</td>
<td>0.2</td>
</tr>
<tr>
<td>Daily DHW demand (kWh)</td>
<td>11.64</td>
</tr>
<tr>
<td>EHWH capacity (m³)</td>
<td>0.5</td>
</tr>
<tr>
<td>Nominal supply power of the GSHP (kW)</td>
<td>1.8</td>
</tr>
<tr>
<td>Supply power of the GSHP (energy optimal) (kW)</td>
<td>0.15</td>
</tr>
<tr>
<td>Static base load (kW)</td>
<td>0.5</td>
</tr>
<tr>
<td>PV capacity (kWp)</td>
<td>5.0 and 21.1</td>
</tr>
</tbody>
</table>

Four different methods were evaluated for controlling the GSHP-based DHW heating:

- **Reference control:** DHW is heated daily at 08:00 A.M. and 08:00 P.M. from 45 °C to 55 °C. The reference control was considered the base case for this study.

- **Clock control:** DHW is heated once a day from 45 °C to 65 °C. This control method aimed to shift the DHW heating towards the hours
when PV production is peaking. Thereby, the method potentially maximized PV electricity utilization in DHW heating, while increasing profitability, due to the simple fact that, in Finland, it is uneconomical to sell local PV generation to the grid.

- **Energy optimal control**: Aims to maximize the COP value of the GSHP, and thereby, minimize the electricity consumption of DHW heating. This control method was only theoretical, as it is impossible to run a GSHP at a very small load.

- **Cost optimal control**: Aims to minimize the daily total DHW heating costs. Here, hourly Nord Pool Spot market price, together with PV output forecast data (or in the case of the perfect forecast, measured PV output data), was utilized. The DHW was heated once a day from 45 °C to 65 °C. The heating duration was determined based on the COP curve and DHW demand, whereas the cost optimal time for heating was found daily at midnight by the use of hourly electricity cost, determined by the electricity price and the forecasted PV output.

The following simplifications were made regarding these control methods:

1. The EHWH tank was assumed to have cooled down back to the original temperature of 45 °C, before the next heating cycle was initiated

2. Thermal losses were not taken into account

3. Daily DHW consumption was assumed to be constant

Between different control methods, the results were compared in terms of DHW heating cost and the share of PV production used in DHW heating. As for the three year comparison (2017–2019) with the perfect PV output forecast, the annual average DHW heating costs and the average share of PV production in DHW heating, are shown in Fig. 4.11. Regarding cost optimal control of the 21.1 kWp PV system, average heating costs were reduced by 14%, when compared with clock and energy optimal controls. Regarding the reference control, a reduction of 31% was achieved by the cost optimal control. As for the smaller 5 kWp PV system, the cost optimal provided a 14% cost reduction compared to the clock control, a 7% reduction when compared to the energy optimal control, and a 22% cost reduction when compared with the reference control. As expected, the cost optimal control method reduced the purchased electricity the most. The results also indicated, that the share of PV production in DHW heating decreased, as the PV system size decreased. This simply raised from the fact that the larger PV system has more capacity to generate PV electricity.
Next, the performance of the actual PV output forecast was studied. The two separate forecast update schemes were compared in order to study the added value of updating the forecast during the day. It could be concluded that, on an hourly level, the intraday updates of the forecast improved the forecast accuracy only marginally, if at all. Based on this result, the intraday update scheme (scheme 2) was left out from the control scheme implementation. Although the studied PV systems suffered only from minimal shadowing losses, the forecast error, and particularly the bias, were drastically reduced by implementing the adjustment process described in Publication II for the PV output baseline forecast (Table 4.5). The high overall accuracy of the adjusted PV output forecast can be seen in the hourly (Fig. 4.12) and daily (Fig. 4.13) performance figures.

Table 4.5. Hourly PV output forecast performance (W/kWp, scheme 1) regarding the baseline and adjusted forecast (June–September 2020). Only instances when either measured or forecasted output is larger than zero, are included. Table information collected from Publication III.

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>RMSE</th>
<th>MAE</th>
<th>Bias</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>0.931</td>
<td>96.1</td>
<td>58.3</td>
<td>16.8</td>
<td>2034</td>
</tr>
<tr>
<td>East-West</td>
<td>0.926</td>
<td>69.5</td>
<td>46.8</td>
<td>11.6</td>
<td>2035</td>
</tr>
<tr>
<td>Total</td>
<td>0.930</td>
<td>81.0</td>
<td>51.3</td>
<td>14.2</td>
<td>2039</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Correlation</th>
<th>RMSE</th>
<th>MAE</th>
<th>Bias</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>0.930</td>
<td>89.5</td>
<td>53.2</td>
<td>2.4</td>
<td>2034</td>
</tr>
<tr>
<td>East-West</td>
<td>0.927</td>
<td>66.1</td>
<td>42.5</td>
<td>0.9</td>
<td>2035</td>
</tr>
<tr>
<td>Total</td>
<td>0.930</td>
<td>76.4</td>
<td>46.9</td>
<td>1.7</td>
<td>2039</td>
</tr>
</tbody>
</table>

The adjustment of the baseline model, although increasing the general accuracy, did not seem to offer significant added value for the decision-
making of the control scheme. This was due to the fact that the baseline model already depicted the studied PV site reasonably well, and more importantly, the uncertainty of the NWP forecast remained to dominate the overall accuracy of the forecast. The adjusted forecast was, however, implemented for the control method simply due to the fact that it outperformed the baseline forecast in every aspect.

Figure 4.12. Hourly performance of the adjusted PV output forecast (scheme 1) regarding the two separate PV systems, and their combination (June–September 2020). Respective LSF lines are shown in red. Figure from Publication III.

Figure 4.13. Daily performance of the adjusted PV output forecast (scheme 1) regarding the two separate PV systems, and their combination (June–September 2020). Respective LSF lines are shown in red. Figure from Publication III.

The simulation results of the four theoretical control methods, together with the actual PV forecast approach, were compared with each other in the period between June–September 2020. This comparison period was determined simply based on the availability of the actual ensemble PV forecast. It should be clarified, that the forecast performance during the investigated months should not significantly differ from other snow-free months, excluded from this comparison. In case the winter period was also modeled with the actual PV output forecast, an additional operation logic for the PV output model in sub-zero, snow-covered periods should be implemented, in order to increase the forecast accuracy in such conditions. Nonetheless, the performance of the demonstrated local forecast application during winter months is not as essential, as most of the PV output is always accumulated during the snow-free period.
The total DHW heating cost between June–September 2020 in all five different control cases is presented in Fig. 4.14. The cost optimal control method with the PV output forecast gave the largest DHW heating cost savings, whether the forecast was based on an actual forecast or measured production (Table 4.6). When comparing the 21.1 kWp and 5 kWp systems, improvements in cost savings tended to increase as the PV system size increased. This is also seen in Fig. 4.14, where the significance of the PV forecast decreased as the PV system size increased. Regarding the utilization of the actual PV output forecast, DHW heating cost savings between 36–53%, depending on the PV system size, were achieved. When comparing the actual forecast against the clock control, 15% (5 kWp) and 21% (21.1 kWp) heating cost reductions could be achieved. It could be concluded that the use of an actual PV forecast did not significantly impair the effectiveness of the cost optimal control, when compared with the perfect forecast. The DHW heating cost with the use of the actual forecast was increased only by 9% in the 21.1 kWp PV system case, and by 11% in the 5 kWp system case, when comparing with the perfect forecast. As for the energy optimal control, the results clearly indicated that by focusing only on minimizing energy consumption, an optimal end result by means of total costs, would not have been achieved.

The average share of PV production in DHW heating for each control case is shown in Fig. 4.15. It can be seen that even the simple clock control significantly increased the share, when compared with the energy optimal and reference controls. It should, however, be pointed out that the sole goal of the cost optimal control was to optimize the costs, not to maximize the share of utilized PV production.

![Figure 4.14. DHW heating costs (June–September 2020). Figure from Publication III.](image-url)
Table 4.6. DHW heating cost decrease (June–September 2020), when compared with the reference control case. Table information collected from Publication III.

<table>
<thead>
<tr>
<th></th>
<th>5 kWp</th>
<th>21.1 kWp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cost optimal</td>
<td>43%</td>
<td>57%</td>
</tr>
<tr>
<td>Cost optimal (forecast)</td>
<td>36%</td>
<td>53%</td>
</tr>
<tr>
<td>Clock</td>
<td>24%</td>
<td>40%</td>
</tr>
<tr>
<td>Energy optimal</td>
<td>18%</td>
<td>25%</td>
</tr>
</tbody>
</table>

Figure 4.15. Average share of PV production, utilized in DHW heating (June–September 2020). Figure from Publication III.

4.5 Application to assess small-scale PV profitability with thermal storage and carbon-corrected electricity prices (Publication IV)

The development in economic profitability of PV systems is a crucial factor in the adoption of the technology. In this Publication, the economic profitability of residential PV was studied at three heterogeneous locations in Finland (Helsinki, Jyväskylä, and Sodankylä; Table 4.7), combined with thermal storage and carbon-corrected prices. The study contributed to the current literature by providing new information on residential PV profitability in a country where no subsidy mechanisms are in place, low solar irradiance levels are present, and a low temporal correlation can be found between electricity consumption and PV output. The penetration of PV is still low in Finland, however, the capacity is increasing rapidly, despite the fact that no direct subsidies are in place.
Results

In order to mitigate the mismatch between PV output and electricity consumption, energy storage can be utilized. Opposed to the high cost of traditional battery storage, the utilization of an EHWH does not require major additional investments. Here, the consumption of electricity regarding a typical four-person household [73] in a 145 m² detached house with electric heating, was modeled. The total electricity consumption was divided into space heating, DHW heating, and other consumption. The consumption profile was based on an hourly type load consumption profile [14], distinguishing between months, weekdays, and weekends.

The space heating consumption profile was scaled for Helsinki, Jyväskylä, and Sodankylä, in accordance to the heating degree day (HDD) coefficients of the normal period 1981–2010 [15]. As a representative load for space heating [51], 110 kWh/m² was used, totaling an estimated annual average of 15 950 kWh space heating consumption. A 3 kW EHWH with a volume of 290 L was considered. The input temperature of water was considered 5 °C, while water was heated to a maximum of 67.5 °C. Thus, the maximum energy storage capacity was 21.15 kWh. Daily DHW consumption of 200 L was assumed. The daily energy required to heat DHW was thereby 11.67 kWh, summing up to a total of 4260 kWh annual DHW electricity consumption. Regarding DHW consumption, an estimated hourly profile was determined by using DHWcalc [31]. Other electricity consumption was calculated as a residual value, with space and DHW heating deducted from the total consumption.

In Finnish electric-heated detached houses, 68% of total electricity consumption is from space heating [72]. DHW heating accounts for 18.2% of the total average consumption, while residual consumption is 13.8%, translating to 3246 kWh of annual consumption. Regarding the representative households, the total estimated consumption was 20 665 kWh for Helsinki, 23 936 kWh for Jyväskylä, and 28 481 kWh for Sodankylä.

The parametric PV output conversion model, described in Publication I, was used for assessing the PV output at each location. As input data, five years (2013–2017) of hourly meteorological observations regarding $T_{\text{air}}$, WS, GHI, and DHI, were used. Also DNI, defined from GHI and DHI, was used. The QC method from Publication I was implemented for the calculated DNI. In addition, no DNI was assumed to be present at $\alpha \leq 0.5^\circ$.

The investigated PV systems were oriented southwards with optimal slope angles, defined or extrapolated by the help of PVGIS data [12]. Site information, together with information regarding missing input data, is shown in Table 4.7. To sum up, the missing data accounted for less than 1% of all daytime ($\alpha > 0^\circ$) data.

Two different sized PV systems were considered: 2.70 kWp and 4.68 kWp. When all PV output could be utilized for personal consumption, the 2.70 kWp system met 22.2% of the household’s electricity consumption in Helsinki and 13.3% in Sodankylä. These values were 40.0% and 24.0% for the
4.68 kWP system, accordingly. The unit investment costs were set to 1925 €/kWP for the smaller and 1568 €/kWP for the larger system, based on actual winning bid values of a Finnish public tender.

Table 4.7. PV site locations, used slope angles, expected annual PV outputs, capacity factors, and missing data (hours when $\alpha > 0^\circ$). Table information collected from Publication IV.

<table>
<thead>
<tr>
<th></th>
<th>Helsinki</th>
<th>Jyväskylä</th>
<th>Sodankylä</th>
</tr>
</thead>
<tbody>
<tr>
<td>Latitude</td>
<td>60.3267°</td>
<td>62.3976°</td>
<td>67.3666°</td>
</tr>
<tr>
<td>Longitude</td>
<td>24.9568°</td>
<td>25.6709°</td>
<td>26.6290°</td>
</tr>
<tr>
<td>Slope angle</td>
<td>39°</td>
<td>42°</td>
<td>49°</td>
</tr>
<tr>
<td>Annual output (kWh/kWP)</td>
<td>1093</td>
<td>968</td>
<td>893</td>
</tr>
<tr>
<td>Capacity factor</td>
<td>12.5%</td>
<td>11.1%</td>
<td>10.2%</td>
</tr>
<tr>
<td>Missing hours</td>
<td>30</td>
<td>200</td>
<td>78</td>
</tr>
</tbody>
</table>

For producing the calculations, the inverter lifetime was estimated to be 12 years, with a change cost of 10% of the system investment cost, together with a discount rate of 3%, and a module lifetime of 25 years. By comparing the calculated LCOE estimates to the average electricity cost from the grid (10.45 cent/kWh in 2016), the economic rationale was assessed. The LCOE results indicated that for the 2.70 kWP system, grid parity was achieved only in Southern Finland (Helsinki), while the investments in Central (Jyväskylä) and Northern (Sodankylä) Finland could not be justified with a 3% discount rate. For the larger PV system with a smaller unit investment cost, Central and Northern Finland investment costs were approaching grid parity, however, not yet reaching it. This LCOE-based profitability analysis depends, however, on two crucial assumptions:

- A fixed price is paid for grid electricity by the household
- All of the PV output can be utilized by the household

It should thereby be pointed out that this approach did not take the hourly changing electricity spot prices, nor the possibility to produce excess PV output, into account. The first hypothesis could be considered valid, as many Finnish households still have a fixed rate electricity contract. Nonetheless, in order to estimate the market-based potential of a PV investment, PV output timing is relevant because electricity price varies over the year. The second hypothesis could be considered valid, as long as the PV system was fitted correctly, and household consumption always covered all produced PV output.

In order to take the described limitations into account, and further investigate the profitability of PV systems in Finland, an optimization model was implemented. The aim of the model was to minimize the annual net electricity costs of a household. The net cost is a sum of two components:
the cost of bought electricity, subtracted by the revenue generated from each unit of excess PV output, which is sold to the grid. The electricity costs for a consumer in Finland consists of the energy cost, the transmission and distribution cost, and the electricity tax, all of which are subjected to a value added tax (VAT). Regarding the excess output that is sold to the grid, the hourly day-ahead market price, subtracted by the electricity retailer’s margin, is received by the household. The optimization strategy thereby aimed to maximize on-site PV output utilization. In addition, the utilization of low-price hours were maximized.

The EHWH was utilized here as a thermal storage. The optimization problem was defined as a discrete-time model with one hour time steps and a one year time span. The hourly DHW heating energy source was chosen in a way that minimized the annual sum of the hourly net electricity cost. The model took the day-ahead market prices, hourly total cost of electricity, exogenously determined other electricity consumption, DHW consumption, EHWH energy content, and PV output, into account. Also the heat losses of the EHWH were taken into account, the losses being proportional to the amount of energy in the heater.

Hourly PV output uncertainties were introduced by fitting a PV output distribution for each combination of hour-of-day and month index, resulting in $24 \times 12 = 288$ probability distributions, constructed from five years of modeled hourly PV output. These distributions were then discretized into $N = 6$ points. The resulting probability distributions indicated that PV output uncertainty was low during winter and high during summer. The model was solved as a stochastic dynamic optimization problem, where the amount of PV output utilized for DHW heating defined the PV output utilized for other consumption, as well as the PV output sold to the grid.

Two separate water heating schemes were evaluated:

- **Time-of-Use (ToU):** a passive night-heating strategy (the EHWH is active 9 hours during nighttime, between 22:00 and 07:00)
- **Optimization (Optim):** an active heating strategy, based on hourly day-ahead market prices

Results indicated that the difference between the passive ToU night-heating and the actively optimized heating strategy increased, as the PV output potential of the household increased. With optimization, PV output self-utilization as well as the unit price (€/kWh) for the sold PV output, were increased. This was seen also in the annual electricity cost savings, where the savings of the optimization strategy exceeded the savings for the ToU strategy in all three locations, and for both PV system sizes (Fig. 4.16). Optimization increased the annual savings by 16.4–32.8 € in Helsinki, 10.7–25.6 € in Jyväskylä, and 6.5–19.5 € in Sodankylä. The cost saving potential regarding the optimal usage of the EHWH increased with
the combination of increased annual PV output and smaller electricity consumption.

Due to the fact that annual savings are not an adequate measure to estimate the investment’s lifetime profitability, the net present values (NPVs) for the lifetime savings were calculated. The lifetime of the system was considered to be 25 years, while the discount rate was set to 3%, being the required rate of return. As shown in Fig. 4.17, the investment was not profitable with a 3% discount rate, meaning that the NPV of savings was always below the investment cost, even when DHW heating was optimized. For the Helsinki site with the passive ToU strategy, the investment cost of the 2.70 kWp system should have been 814 € lower in order for the investment to break even. For the 4.86 kWp system, 1 129 € lower investment costs would have been required. By implementing the active DHW heating optimization, the costs should have been 516 € and 531 € lower, respectively. The corresponding maximum allowed unit investment costs (€/kWp) for each of the studied scenarios are presented in Table 4.8).

The NPVs in Fig. 4.17, calculated with the LCOE principle, indicated a better investment profitability, when compared to the more realistic analysis with actual hourly market conditions. This difference was explained by the fact that the LCOE calculation ignored the potential of excess PV output, and that the calculation did not take the hourly varying electricity value, replaced by the household’s own PV generation, into account. To conclude, the actual value of PV output for the households was lower, when measured with hourly prices, rather than with an average price over the hours.

Figure 4.16. Annual extra savings of DHW heating optimization, when compared to the passive nighttime heating (ToU) strategy. Figure from Publication IV.
Figure 4.17. The NPV of savings with PV under passive (ToU) and optimized (Optim) DHW heating. The LCOE results refer to a scenario with a fixed electricity price, where all PV output is used for own consumption. Figure from Publication IV.

Table 4.8. Maximum unit investment costs (€/kWp) of the PV systems in order to meet the required 3% rate of return. Table information calculated from Publication IV.

<table>
<thead>
<tr>
<th>PV capacity</th>
<th>Helsinki ToU</th>
<th>Helsinki Optim</th>
<th>Jyväskylä ToU</th>
<th>Jyväskylä Optim</th>
<th>Sodankylä ToU</th>
<th>Sodankylä Optim</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.70 kWp</td>
<td>1624 €</td>
<td>1523 €</td>
<td>1595 €</td>
<td>1492 €</td>
<td>1536 €</td>
<td></td>
</tr>
<tr>
<td>4.86 kWp</td>
<td>1336 €</td>
<td>1273 €</td>
<td>1369 €</td>
<td>1280 €</td>
<td>1353 €</td>
<td></td>
</tr>
</tbody>
</table>

In order to determine the profitability of these investments, the internal rate of return (IRR), was calculated. The highest IRR of 2.1% was achieved in Helsinki with the 4.86 kWp system, together with DHW heating optimization. This was an 0.7 percentage point increase, when compared to the passive ToU baseline scheme. These results indicated that small-scale PV investments with the selected parametrizations may not, as of now, have been economically viable in order to enhance rapid PV capacity growth in locations with similar solar irradiance conditions. However, this may very well change due to the following factors:

- It is expected that the declining trend in PV system unit investment costs continues into the future
Results

- A PV system is a relatively risk free investment, when compared with many other asset classes
- Households may put weight on other PV-related factors, such as emission reductions
- The future electricity price level is difficult to forecast, combined with the uncertainties regarding the emission allowance prices, set in the future

The studied higher emission price levels of 25.0 €/tCO₂ and 50.0 €/tCO₂ enhanced the profitability of the PV investments. The IRRs with a 50.0 €/tCO₂ carbon price are illustrated in Fig. 4.18. A higher carbon price improved the profitability in all locations, as expected. The 3% IRR requirement was exceeded by the larger PV system in Southern Finland with an IRR of 3.5%, while the smaller PV system reached an IRR of 3%, respectively. When investigating the overall IRR sensitivity to emission allowance price level, a somewhat linear relationship between the IRR and carbon price was found, where a 10 €/tCO₂ increase in carbon price improved the IRR by 0.3 percentage points for both EHWH schemes, combined with the 4.86 kWp system.

![Figure 4.18. The IRR of savings with PV under passive (ToU) and optimized (Optim) EHWH strategies under 50.0 €/tCO₂ carbon-corrected electricity prices. Figure from Publication IV.](image-url)
Results

4.6 Application to assess distributed thermal storage utilization in PV output forecast error cost minimization (Publication V)

Despite the increasing interest towards demand response (DR), consumer incentives for offering their energy flexibility to power markets has not been widely studied. New types of business models and trading mechanisms, enabled by the new technological solutions are, however, required in order to activate the potential of DR. In this Publication, a virtual power plant (VPP) concept was studied, where VPP operation was based on simultaneous optimization of active EHWHs against PV output forecast error imbalances (Fig. 4.19). The uncertainty in the optimization was introduced through day-ahead PV output forecast errors and balancing power market conditions. The EHWHs were here selected as the controllable distributed thermal storage, due to their typical oversized heat-absorption capacities, good insulation properties, and the fact that their implementation can be done without sacrificing the general living comfort level.

![Figure 4.19](image.png)

**Figure 4.19.** The VPP balances any imbalances caused by the day-ahead PV output forecast errors, either by operating with the TSO, or by controlling the consumption and generation units within its control. The optimization of the VPP operator is here represented by the dotted lines. Figure from Publication V.

VPP operation depends on the electricity production and consumption assets that it controls, coupled with the questions of economies of scope and scale, in addition to the design of an appropriate business model. Here, the economics of scale was studied by identifying the average and marginal added values for the individual households, taking part in the VPP operation. The economics of scope was studied by using EHWHs as a thermal storage resource, both regarding the household’s heating costs, and for minimizing the PV output forecast error cost. The market value of the weather forecast accuracy was also demonstrated.
The developed VPP model, including consumption and production resource control, was implemented by combining models and elements from earlier literature. The individual households provided the flexibility, as their electricity DR was utilized to balance the forecast errors of a 1 MWp PV plant. The goal of the VPP operator was to minimize the forecast error costs by utilizing the uncertainties of the balancing power market conditions and the PV output forecasts. This meant that the operator may have bid its production, based on the latest available forecast at the day-ahead market closing, and dealt with the outcome regarding the forecast errors in the energy imbalance market, after the uncertainties had been materialized. The operator may also have compensated for the forecast errors by utilizing the internally controllable flexible consumption resources.

Real-time pricing (RTP) with hourly fluctuating prices can be regarded the most obvious and economical way to implement DR in the power markets. Here, it was assumed that the consumers were under RTP, and that the power balance of the PV producer was managed by a Balance Responsible Party (BRP), which allocated all the PV producer’s generation imbalance costs directly to the producer. In addition, it was assumed that the PV generation capacity (1 MWp) was sufficiently small, and did not significantly affect the balancing of the power market quantity and price equilibrium, even though the forecasting errors were present.

A producer is obligated to submit a production plan to the TSO the day before delivery, in case it has a significant share of generation units. These initial production plans are then updated constantly in order to include any expected imbalances in the TSO’s balancing plans. The TSO buys and sells imbalance power from and to the BRPs. The producer is required to buy imbalance power from the TSO via the BRP in case of a production deficit, and vice versa. The forecast error costs originate from incorrect bidding in the day-ahead market, resulting in imbalance payments in the imbalance power market. Here, it was considered that the VPP operator cannot, as such, improve the forecasts. It can instead minimize the costs by maximizing the imbalance revenue. This is accomplished by correctly allocating the PV output forecast error between the controllable consumption resources and the imbalance power market.

The model was implemented to the Finnish power market, where most of the electricity is currently traded in a day-ahead auction market. The day-ahead market for bids +17...+40 h is closed at 09:00 AM (UTC). After the day-ahead market is closed, power balance is controlled within several markets, e.g., the intraday market, various reserve markets, and the balancing market. For generating the output for the 1 MWp PV system, power output measurements and corresponding forecast errors from a 21 kWp rooftop PV system (Helsinki site in Table 4.1 and Fig. 4.2) were scaled up to 1 MWp. The utilized PV output forecast was based on the output of the HARMONIE NWP model [7], whereas the PV output conversion
model was based on the model, described in Publication I. The generated hourly time series consisted of successive NWP forecasts, daily initialized at 06 UTC. The utilized forecast horizon for each forecast was set between +17...+40 h, covering the day-ahead time frame.

The hourly maximum errors over the snow-free period were approximately 60% of the nominal capacity. Correspondingly, monthly standard deviations ranged between 10% and 14%. The cumulative PV output imbalance at the end of the annual period 2016 was a deficit of 54.7 MWh. Due to the variability in the sun’s trajectory, i.e., the variability in the amplitude and time frame of daily PV output generation, the uncertainty related to the VPP optimization varied considerably over the year. In general, forecast uncertainty correlated hourly and monthly with the PV output potential, the error increasing with a greater PV output potential.

The modeled four-person households were identical to the ones described in Publication IV, having each an EHWH with a storage volume of 290 L and a heating power of 3 kW, combined with an identical hourly DHW consumption profile as in Publication IV, determined by DHWcalc [31]. The annual discount rate was set to 3%.

In order to replicate the balancing power market outcome, the market state probabilities, together with the price distributions, must be formulated. For the balancing power market prices, the probabilities and distributions for each hour-of-day-by-month combinations, were defined. For each corresponding hour, the imbalance cost/revenue was defined by the price discrepancy between the balancing power price and the day-ahead market price. In order to generalize the model, the local taxes and grid fees were excluded from the analysis.

The results consisted of average values over 25 random sample draws from the hourly solar forecast error and imbalance price probability distributions. Without VPP implementation, the average cost for the optimized EHWH was 2.34 cent/kWh, while the average electricity price was 3.24 cent/kWh. The annual forecast error cost for the PV electricity producer was 830 € on average, meaning that the annual revenues could have been increased by 830 € if no forecast errors would have been present. In other words, it was the added value of a perfect PV output forecast, compared to the actual forecast. It corresponded to 2.5% of the PV plant’s total revenue. The bulk of the costs were generated between April and October when the PV output was at its highest.

Without the VPP, the cheaper night-time hours were utilized more often. By including the VPP, more electricity was used during daytime. However, the profile was approaching the single household’s night-time weighted profile without VPP, as the number of households was increasing. The results showed that the VPP had a stabilizing effect on the system. By utilizing the VPP, the forecast error cost could be reduced from 830 € to 674 € with five households, and to 489 € with 50 households (Fig. 4.20).
The net benefit of the VPP increased from 37.6 € with 5 households, to 252.0 € with 50 households. Even though the net benefit increased with a greater household participation, the average and marginal benefits per member were diminished (Fig. 4.21). The average benefit was here determined as the net benefit divided by the number of households, whereas the marginal benefit was the additional benefit related to new households, divided by the number of new households.

These values can be considered as profits to the households, in case no profits were allocated to the VPP operator or PV producer. As such, the annual monetary compensation per household, ranging between 4.0 € and 7.5 €, was not large. A larger amount of PV generation sources could, however, lead to an increased compensation per household. In addition, optimization of the PV generation day-ahead bidding might also increase
Results

the value of the VPP operation.
PV output modeling is a topic which has experienced great research and development emphasis in recent years, resulting in a variety of proposed approaches. Despite the fact that certain best practices have gradually taken shape, no comprehensive solutions have yet been established for handling PV output modeling, as a whole. The main contributions of this doctoral thesis are related to the development and validation of the proposed components of a holistic site-specific PV output conversion model, in addition to demonstrating the use of such a tool in a selection of useful applications. These contributions demonstrate the performance of the selected modeling approach, and provide tangible information about the future application potential within the selected subject areas.

In Publication I, the proposed quality control method for calculated DNI values was shown to efficiently filter unrealistic data, providing a feasible method for treating calculated DNI values, as long as a set of representative DNI measurements are available for determining the case-by-case prevalent maximum DNI limits.

Through the implementation of the optimized parametric PV output conversion model in Publication I, useful information could be obtained regarding the behaviour of the studied PV systems and their components. The model was shown to be suitable for estimating losses due to sub-optimal inverter behaviour, shadowing and snow cover effects, and in monitoring the general health and efficiency of a PV system. The selected well-performing sub-models regarding the transposition of diffuse radiation as well as module temperature were concluded as feasible options for these tasks. It was also shown that snow cover has a clear impact in mitigating PV output potential in higher latitudes, and generates an additional challenge for implementing an accurate all year round PV output model.

The suggested adjustment approach for the PV output conversion model in Publication II was established as an explicit and straightforward method for adjusting a generic PV output forecast to any specific site. As long as a sufficient volume of measured PV output data was accessible, the cloud
Concluding remarks

Forecasting ambiguities could be removed completely from the adjustment process. The utilization of the shifting 30 day time window also enabled the fluctuating shadowing losses, induced by the shifting solar trajectory, to be taken into account. The method captured systematic biases of the NWP forecast, as well as site-specific conditions that affected the PV system output. Modeling accuracy was increased by allowing site information regarding conversion losses and shadowing, in addition to other important loss factors, to be taken into account. The approach can be considered a viable tool for producing adjusted PV output forecasts for sites with no available exogenous measurements, this being a reality for most of the small-scale PV sites. Despite the limited five month time coverage of the investigated period, it can be stated that the used period gives a good overview on the general performance of the approach in snow-free conditions. All and all, the proposed approach can thereby be considered a potential solution for various site-specific PV modeling applications.

Results documented in Publication I and Publication II support the remark that explicit and uncomplicated PV output conversion models can attain good performance in PV output modeling applications, and comprise the potential for reaching an even greater precision when variable site-specific losses can be taken into consideration.

The potential to decrease DHW heating costs with a PV output forecast-based control method for household GSHP operation, coupled with local PV generation, was demonstrated in Publication III. The utilized control method logic was based on the demonstrated hourly PV output forecast (Publication I and Publication II), an experimentally defined COP curve, and hourly day-ahead electricity spot market information, the control method also being verified with a perfect PV output forecast. Results with the actual PV output forecast, ranging over a four month summer period in Finland, showed that the developed cost optimal control was able to reduce DHW heating costs when compared to all other control schemes. The results also indicated that, as the PV system size increased, the accuracy of the PV production forecast lost its significance. It should be pointed out, that some additional implementation steps for the actual PV output forecast would likely be required for months when snow cover and sub-zero temperatures are present. However, it should also be noted that the performance of the forecast during winter months is not as essential, as most of the PV output is always accumulated during the snow-free period.

A residential PV profitability study, modeled with the help of comprehensive in-situ measurements and the methods described in Publication I, together with DHW storage and carbon-corrected prices, was carried out in Publication IV for Finland, where households’ mismatch between electricity consumption and PV output is one of the key obstacles for PV profitability. It was shown that energy storage, coupled with optimization, has means to mitigate this issue, the residential consumers also providing
Concluding remarks

a valuable future source of demand-side flexibility. Results implied that an LCOE-based evaluation can provide unrealistic profitability estimates for residential PV investments, due to the fact that the approach ignores the produced excess PV output and the fluctuating electricity prices. To better assess the profitability of residential PV investments, households should be able to use a PV investment profitability calculator, provided by a neutral party, where the households’ electricity consumption profile, a realistic PV output profile, and an electricity price profile, would be taken into account in order to provide a truthful PV profitability estimation. It should be noted, that the presented results might slightly overestimate the actual PV profitability, due to the fact that snow cover losses, together with potential shadowing effects, caused by partly obstructed real-world PV site horizons, were not taken into account in the modeling process. On the other hand, the declining trend in PV system prices, being the major driver for PV profitability improvement, is projected to continue also in the future, causing PV profitability calculations to expire quickly.

In the future, new operators, such as VPPs, are required to aggregate and coordinate active consumer behaviour within the markets. Regarding the formation of a VPP, it is crucial to couple the resources in such a way that the aspects of economies of scope and scale are included in the business model design. In Publication V, the demonstrated VPP, modeled with NWP data and the methods described in Publication I, was shown to add value both by minimizing the electricity costs of DHW heating, in addition to utilizing the household EHWHs as a resource in mitigating the PV generation imbalance. The VPP operation was shown to decrease PV forecast imbalances and benefiting the participating households. It should be noted here that the observed positive bias, displayed by the utilized PV output forecast results, could have been mitigated to some extent through utilizing, e.g., the methods described in Publication II, which was published only after Publication V had already been finalized. The general accuracy of the selected NWP model also plays a crucial role in the determination of these initial error costs, and it can be argued that the more recent ensemble NWP model version would have provided some additional value to this process. Overall, however, it can be concluded that a VPP can make a meaningful impact in the future electricity markets, even though only a minimal effort is required from the consumer side.

5.1 Future prospects

A common denominator for all snow-induced effects is the fact that they generally revolve around the darkest and least productive part of the year. Due to this, among other various reasons, the more challenging but less relevant winter conditions were deliberately mostly excluded from this
Concluding remarks

Thesis, when considering the tool development scope for forecasting the output of specific PV sites. To briefly summarize, forecasting of snowfall and the variable snow conditions is itself a challenging task. In addition, modeling how those uncertain forecasts translate into the fluctuating snow conditions experienced by individual PV modules or the PV system as a whole, and how that ever-changing situation in turn then translates to actual PV output, is filled with challenges, enough to comprise a stand-alone Thesis. Implementation of a simplistic PV output data-based correction to local PV generation forecasts is always an option, however, it might not comprise enough interest on its own as a meaningful scientific research question. As for the Nordic PV context in general, it can be stated that a holistic approach for handling the PV output estimation in wintry conditions would potentially be of use in cases where the output of PV systems are to be evaluated throughout the year. This need would become particularly evident on the larger power grid level, in case PV acquired a significant share in electricity generation.

Regarding DNI QC, covered in Publication I, further studies for defining representative QC limits for calculated DNI values in different geographical and climatological regions could prove beneficial. Such information would likely be useful also outside the scope of PV output modeling, in fields where solar radiation data is commonly utilized.

As described in Publication II, sky clearness was shown to affect the bias of the PV output forecast. An appropriate solution for taking sky clearness into account when defining the PV model adjustment coefficients is still left open, one of the main hurdles being the fact that NWP uncertainty plays a significant role in accurately defining this relationship. Further research within this topic is however encouraged, in case this relationship is found to play a significant role in the accuracy of the proposed adjustment method. The potential added value of implementing a separate post-processing step for the NWP model’s solar radiation data, before implementing the adjustment method, could also provide an interesting future research question.

Including the NWP ensemble forecast uncertainty in state-of-the-art PV production forecasts is nowadays a common practice. It would be of potential future research interest to see how such uncertainty information could be exploited in applications which incorporate PV output forecasts into their operation logic, one example being the BASs. This viewpoint was left outside the scope of Publication III, due to the need for narrowing down the research topic.

The forecasting aspect can be considered in many ways the more challenging part of PV output modeling. Best practices, ranging anywhere between the development of PV-specific NWP models and a multitude of different post-processing and ML approaches, are slowly taking their place. Although regional modeling still seems to be emphasized over small-scale
approaches, simply due to its higher current demand, future leaps in DR, BAS, and other relevant topics might reallocate this focus in the future.


Bibliography


**Errata**

**Publication I**

- Parameter $c_1$, included in Eq. 2 and Eq. 3, is not correctly formulated. The correct formulation is $c_1 = \frac{4}{3\pi}$.

- Parameter $b$ value, included in Eq. 4, is not correct. The correct value is $b = -0.0594$.

Errors are only present within the text, not affecting the actual results.