UNIVERSITY
OF OSLO

Mari Benedikte Øgaard

Methodology for performance evaluation of PV systems - with a special focus on high latitudes and snow

Thesis submitted for the degree of Philosophiae Doctor

Department of Technology Systems
Faculty of Mathematics and Natural Sciences

Institute for Energy Technology

2022

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2555

ISSN 1501-7710

Print production: Graphics Center, University of Oslo.

Abstract

In order to mitigate dangerous effects of global warming, transitioning to a sustainable, low emission energy system is necessary. The cost of solar photovoltaics (PV) has declined rapidly in recent years, and the technology is expected to play an important role in this transition. To facilitate efficient use of resources, we should pursue further reductions in costs as well as improvements in the performance of PV installations. Accurate methods for predictions and evaluation of PV performance can be used in various ways to reduce costs and improve performance. PV performance predictions and evaluation are for example necessary to ensure optimized design of systems and to evaluate potential for improvements in loss reduction.

To efficiently reduce costs and improve PV performance, the methods for performance evaluation and prediction should be easy to implement and automate. However, PV performance depend on system design and operating conditions, and the type of losses that are most prominent can vary between installations. Additionally, there are factors not related to the performance that can impact the measured output of PV systems and the performance evaluation, for example the quality of the measured data. Consequently, as multiple different parameters should be considered in the evaluation and prediction of PV performance, development of standardized methodologies is challenging.

The main topic of this work is methodologies for PV performance *evaluation*, i.e. identification and quantification of loss mechanisms. The aim of the research is to contribute to the development of standardized methodologies. The main loss mechanisms studied in this work are component faults and snow shading. The analysis uses data collected from PV installations in Norway with a total capacity of ~3.7 MW. Both PV performance and the methodology for performance analysis are less studied for operating conditions found at higher latitudes compared to the operating conditions found closer to the equator. In the first part of this work, established methodology for performance evaluation are assessed on output data from commercial PV monitoring systems for the case of fault detection. First, factors impacting the calculated performance metrics are classified through analysis of the periods with large degree of noise or systematic trends in performance metrics, or estimated large performance gains or losses. Second, methodology to handle the effects of these losses and impact factors on the performance metrics are discussed, and the use of filtering is specifically evaluated. With targeted filtering, improved sensitivity in a fault detection analysis is achieved.

Through this initial evaluation of factors impacting performance evaluation at high latitude locations, it is identified that there is a need for improved methods to identify and predict the potentially large energy losses caused by snow. To contribute to this, we describe the effect of snow on PV monitoring output parameters, evaluate existing snow loss models, and suggest improvements to the commonly used Marion snow loss model. The improved suggestion gives a reduction in modeling error of 23 percentage points for the studied dataset compared to the default implementation of the Marion model.

Preface

This thesis is submitted to the Department of Technology Systems, Faculty of Mathematic and Natural Sciences, University of Oslo, in partial fulfillment of the requirements for the degree of *Philosophiae Doctor* (PhD). The presented research has been conducted at the Department of Technology Systems, University of Oslo, in collaboration with the Institute for Energy Technology, under supervision of associate professors Josefine Selj, Sabrina Sartori, Halvard Haug, and Dr. Heine Riise. The research has been funded by the University of Oslo's innovation project *Autonomous monitoring, control and protection of renewable energy infrastructure* (project number 80381), popularly known as VÅK. In addition to the research, 25 % of the time in the project have been dedicated to innovation work.

This thesis consists of 7 papers, and 6 introductory chapters. The introductory chapters aim to contextualize the papers, relate the papers to each other, and provide relevant background information. All the papers are based on joint work with colleagues. The overall topic is development of methodology for performance evaluation of PV systems. The effect of high latitude conditions and snow on the PV performance evaluations is specifically studied.

Acknowledgements

During my research for this thesis, the PV system research group at IFE and the Department of Technology Systems at UiO have grown rapidly. The PV system research group at IFE has additionally endured multiple reorganizations. The list of people who have contributed to making this work fun and interesting is consequently very long. I am deeply grateful for all the help and support I have received from all the wonderful people I have met.

First, I want to thank my main supervisor, Josefine. Your support has been rock-solid. Thank you for always finding time for me in between establishing the PV system research field in Norway. I also appreciate your helpful feedback and guidance. You have always given me freedom to follow the paths I believed in, but at the same time also made sure I was on track.

I would also like to thank my co-supervisors Heine, Halvard and Sabrina for your encouragement and your excellent questions and feedback. I highly appreciate your positive attitudes.

I am very grateful that I got to do this work together with great people doing their PhDs on similar topics. You have made this fun and interesting. I am lucky to never have felt alone in this work. First, thanks to Bjørn, who not only should be famous for his joyous demeanor and bad jokes, but also for his critical thinking, thorough work and not least for being a kind and caring colleague. Thanks to Åsmund and Eivind at IFE, for the many interesting discussions, work- and not-at-all work related, and great collaborations. Your input has been immensely important for my research, and I would have achieved much less without you. Thanks to Victoria and Tor Atle for the fun time in the VÅK project. I have really enjoyed the long coffee breaks/project meetings, the drone flying practices, trips to Sweden, and maybe our greatest achievement – making the Real Moro movies!

Thanks also to the rest of the collaborators in the VÅK project, for developing the project and enabling this research.

I feel very lucky to have been a part of the PV system research group/department at IFE. It is extraordinary to have a group with people that not only are extremely clever and fun, but also very kind. You give great feedback and brings forward wonderful research that is easy to trust and lean on. I am grateful for the fun and supportive working environment you have created.

Thanks also to all collaborators in the Norwegian PV industry through various IFE projects. It has been great to collaborate with clever people that both can push the research forward and implement it. In particular, thanks to Fusen and Solcellespesialisten for contributing with data, excellent questions and discussions, and endless PV system knowledge.

Thanks also to all the administrative people at ITS and IFE making everything practical easy, and for contributing to great working environments.

Thank you to Gaute, Josef, Isabelle and Vilde for input and feedback on my thesis writing.

I am additionally immensely grateful for the various social working environments I have gotten to be a part of during this work. Lots of love out to the people in the solar corridor at IFE that contributed to making this a place where desserts, parties, intricate coffee break discussions and games were in focus. To the extended PhD community at ITS, thanks for all the extravagantly long lunches, late afterwork beers and fun conversations on the border of almost too nerdy. Thanks to the pre-covid PhDs for building a community and developing traditions, and to the post-covid PhDs for keeping it alive. Thanks to you, ITS has been a perfect place to finish writing this thesis.

Finally, I want to thank my friends and family for never-ending support, and for always undoubtedly believing in me. And most of all, thanks to Lars. For being tremendously supportive, for always lightening up every day, all big and small adventures, and all your input on writing and design. I could not have done this without you. Also quite literally, without you as a steady home office colleague, my success in Pokémon Go would maybe have been greater, but fewer papers would have been written.

List of papers

This thesis consists of the following papers. In the text the papers will be referred to by their roman numerals.

Paper I

Methods for quality control of monitoring data from commercial PV systems
M.B. Øgaard, H. Haug, and J.H. Selj
In: *Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition* (2018), pp. 2083-2088.
DOI: 10.4229/35thEUPVSEC20182018-6DV.1.53

Paper II

Performance evaluation of monitoring algorithms for photovoltaic systems
M.B. Øgaard, Å. Skomedal, and J.H. Selj
In: *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition* (2019), pp. 1632-1636.
DOI: 10.4229/EUPVSEC20192019-5CV.4.30

Paper III

Photovoltaic system monitoring for high latitude locations
M.B. Øgaard, H.N. Riise, H. Haug, S. Sartori, and J.H. Selj
In: *Solar Energy*, Vol. 207 (2020), pp. 1045-1054.
DOI: 10.1016/j.solener.2020.07.043

Paper IV

Identifying snow in photovoltaic monitoring data for improved snow loss modeling and snow detection
M.B. Øgaard, B.L. Aarseth, Å.F. Skomedal, H.N. Riise, S. Sartori, and J.H. Selj
In: *Solar Energy,* Vol. 223 (2021), pp. 238-247.
DOI: 10.1016/j.solener.2021.05.023

Paper V

Modeling snow losses in photovoltaic systems
M.B. Øgaard, H.N. Riise, and J.H. Selj
In: *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)* (2021), pp. 517-521.
DOI: 10.1109/PVSC43889.2021.9518886

Paper VI

Snow loss modeling for roof mounted photovoltaic systems: Improving the Marion snow loss model

M.B. Øgaard, I. Frimannslund, H.N. Riise, and J.H. Selj

Early access in: *IEEE Journal of Photovoltaics* (2022).

DOI: 10.1109/JPHOTOV.2022.3166909

Paper VII

Estimation of snow loss for photovoltaic plants in Norway

M.B. Øgaard, H.N. Riise, and J.H. Selj

In: *Proceedings of the 38th European Photovoltaic Solar Energy Conference and Exhibition* (2021), pp. 1081-1087.

DOI: 10.4229/EUPVSEC20212021-5DO.4.5

Related work

The following papers support the discussions in the main papers or implement suggested solutions but are not included since they thematically are beyond the scope of this thesis.

Mitigating snow on rooftop PV systems for higher energy yield and safer roofs
B.L. Aarseth, M.B. Øgaard, J. Zhu, T. Strömberg, J.A. Tsanakas, J.H. Selj, and E.S. Marstein
In: *Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition* (2018), pp. 1630-1635.
DOI: 10.4229/35thEUPVSEC20182018-6CO.3.5

Detailed loss analysis for wall mounted photovoltaic systems at high latitude; A comparison of multicrystalline Si- to CIGS- modules
G. Otnes, M.B. Øgaard, L.T. Milde, S.E. Foss, and J.H. Selj
In: *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition* (2019), pp. 1621-1626.
DOI: 10.4229/EUPVSEC20192019-5CV.4.25

The performance of a floating PV plant at the west coast of Norway
J.H. Selj, I.H. Lereng, P. De Paoli, M.B. Øgaard, G. Otnes, S. Bragstad, B. Bjørneklett, and E.S. Marstein
In: *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition* (2019), pp. 1763-1767.
DOI: 10.4229/EUPVSEC20192019-6DO.9.2

General, robust and scalable methods for string level monitoring in utility scale PV systems
Å. Skomedal, M.B. Øgaard, J.H. Selj, H. Haug, and E.S. Marstein
In: *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition* (2019), pp. 1283-1287.
DOI: 10.4229/EUPVSEC20192019-5BO.5.4

The influence of thermal signatures observed with infrared thermography on power production in a utility scale power plant
B.L. Aarseth, Å. Skomedal, M.B. Øgaard, and E.S. Marstein
In: *Proceedings of the 37th European Photovoltaic Solar Energy Conference and Exhibition* (2020), pp. 1360-1363.
DOI: 10.4229/EUPVSEC20202020-5DO.3.1

PV system degradation rates in the Nordics
E.B. Sveen, M.B. Øgaard, J.H. Selj, and G. Otnes
In: *Proceedings of the 37th European Photovoltaic Solar Energy Conference and Exhibition* (2020), pp. 1563-1566.
DOI: 10.4229/EUPVSEC20202020-5CV.3.36

Robust and fast detection of small power losses in large-scale PV systems
Å.F. Skomedal, M.B. Øgaard, H. Haug, and E.S. Marstein
In: *IEEE Journal of Photovoltaics*, Vol. 11 (2021), pp. 819-826.
DOI: 10.1109/JPHOTOV.2021.3060732

Performance analysis of a BAPV bifacial system in Norway
H.N. Riise, M.B. Øgaard, J. Zhu, C.C. You, F. Andersson, T. Bønsnæs, J. Young, and S.E. Foss
In: *2021 IEEE 48th Photovoltaic Specialists Conference (PVSC)* (2021), pp. 1304-1308.
DOI: 10.1109/PVSC43889.2021.9518963

Soiling and Snow Impact on a PV Plant at a Farm in Norway
H.N. Riise, M.B. Øgaard, and T.U. Nærland
In: *Proceedings of the 38th European Photovoltaic Solar Energy Conference and Exhibition* (2021), pp. 1241-1244.
DOI: 10.4229/EUPVSEC20212021-5CV.2.26

Hybrid PV-systems for electrification and sector coupling of road transport sector in Norway
J. Fagerström, L. Kvalbein, J. Danebergs, T.U. Nærland, M.B. Øgaard, and K. Espegren
In: *Proceedings of the 38th European Photovoltaic Solar Energy Conference and Exhibition* (2021), pp. 1350-1355.
DOI: 10.4229/EUPVSEC20212021-6CO.11.4

Hybrid PV Systems and Colocalization of Charging and Filling Stations for Electrification of Road Transport Sector
J. Fagerström, L. Kvalbein, J. Danebergs, T.U. Nærland, M.B. Øgaard, and K. Espegren
In: *Solar RRL* 2100461 (2022).
DOI: 10.1002/solr.202100461

Contents

1 Introduction

1.1 Motivation and background

In August 2021, IPCC published their first major review of climate change research since 2013, and it was announced as a “code red for humanity” [1]. The language of the report is dramatic and clear, stating that “it is unequivocal that human influence has warmed the atmosphere, oceans and land” and that this warming is “already affecting many weather and climate extremes in every region across the globe” [2]. This warming is mostly caused by greenhouse gas emissions caused by burning fossil fuels. There is hope that if these emissions are greatly reduced, the climate can stabilize. UN Secretary General António Guterres summarized the findings of the report like this: “If we combine forces now, we can avert climate catastrophe. But, as today's report makes clear, there is no time for delay and no room for excuses.” [3]. Reducing greenhouse gas emissions is, however, challenging. The burning of fossil fuels has not only produced global warming, but it has also provided access to cheap energy which has played an important role in economic development and growth [4]. To maintain our living standard, enable growth in developing countries, and at the same time reduce emissions, the creation of a more sustainable energy system is expected to play a key role [5,6]. Development of a sustainable energy system necessitates replacement of fossil fuels with low emission technologies, and improvements of the energy efficiency, flexibility, and storage abilities in the system [7].

To change the energy system fast enough to avoid serious global warming is, however, no easy task. Nuclear energy requires large investments and long construction times, and development is additionally limited by discussions on safety and waste handling. Energy generation using renewable energy technologies will depend on the availability of the natural resources utilized in the different technologies. Central renewable energy technologies such as wind and solar are additionally highly weather dependent, resulting in intermittent energy generation. This intermittency introduces multiple challenges and a need for increased flexibility in our energy systems, which often is designed around the characteristics of fuel based thermal power [7]. Additionally, construction of new energy infrastructure can introduce conflicts related to land use, for example with food production [8]. New infrastructure projects can in many cases also negatively impact the local environment and nature [9]. Change in land use often leads to habitat loss, which is considered the main reason to biodiversity loss [10,11]. As underlined in [11], the decline in biodiversity is now faster than at any time in history, which ultimately can impact human quality of life. The impact on the local environment is also often a central reason for development of social resistance, which in multiple occasions has canceled or delayed energy infrastructure projects [12]. Developing a new, low emission energy system without destroying too much nature or creating too much conflict, and at the same time ensuring energy security, is consequently a complex task with many considerations to balance. No technology presents a simple, quick solution alone. To avoid infrastructure development on the most vulnerable areas and to ensure a resilient energy system, major investments in a range of different technologies are necessary.

Solar photovoltaics (PV) is one of the technologies that will play an important role in a low emission energy system [13]. While the technology is not new itself, it is quite recent that PV entered the market as a competitive option. The competitiveness of PV has been driven by an extraordinary price reduction primarily caused by increasing production volumes of PV modules [14] and strong industrial competition. From 2010 to 2020, the annual PV module production increased by a factor of seven, giving a total global annual production of PV module capacity of approximately 140 GW in 2020. While it took almost six decades to reach 100 GW of installed PV capacity in 2012, it is expected that 1 TW of PV is installed by 2022. This makes PV the fastest-growing power generation source of the last decade [13]. Rapid growth in PV installations is seen all over the world [15], also in northern regions with lower annual irradiation, as seen in Figure 1 (although with a total capacity significantly lower than in the largest PV markets).

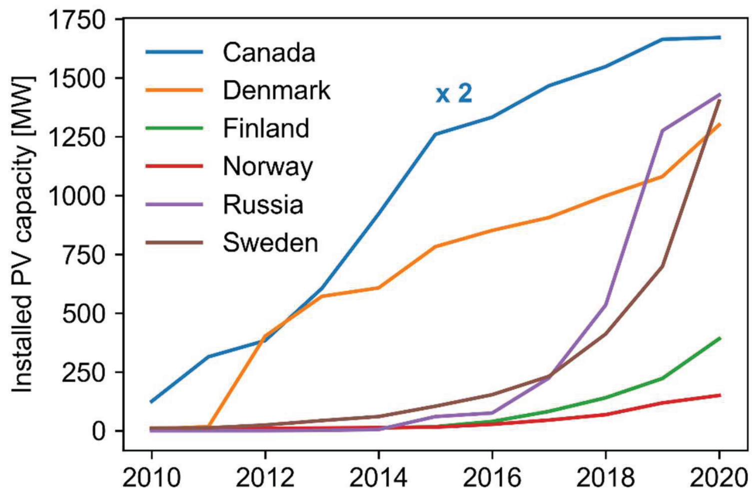


Figure 1: Increase in installed PV capacity from 2010 to 2020 in northern countries. Data from [16]. *For easier comparison of growth, the absolute PV capacity for Canada is divided by two.*

PV do require large areas [17], but PV installations do not require a *specific* type of area. PV can be built on roofs, on water [18], and on land that cannot be used in other ways, such as deserts [19] and landfills [20]. Both the PV system and the solar module itself can additionally be designed in many ways. There are examples of PV system designs adapted to co-localization with farming [21], and solar modules designed for incorporation in building elements (BIPV) [22]. This flexibility enables large-scale installation of PV without building in vulnerable areas. Additionally, because of its modularity and the relatively simple system design, PV can be installed rapidly and in a large range of sizes (giving a relatively low investment threshold) and is not significantly limited by lack of expertise.

Even though PV has become a cheap and competitive technology, further reductions in costs and improvements in the performance of PV installations may contribute to a successful transition to a low emission energy system. Historically, most of the research related to PV performance improvements have been focused on improvements in the solar cell efficiency. With the maturing of the technology, an increasing amount of research on other ways to improve PV performance is performed. Identification and quantification of losses in a PV system can play a role in improving PV performance and reducing costs in multiple ways, and evaluation and prediction of the PV *system performance* are therefore now important research topics [23,24]. Evaluating PV performance through identification and quantification of losses is necessary to identify if there is potential for improvements and to document that the systems are working as intended. Based on performance evaluations of existing systems, we can identify which losses we can expect in a system and determine how (and how long) different technologies work in the field and under different conditions. Identification and quantification of system losses in existing systems is additionally essential input to performance predictions. Predicting the performance of future systems is necessary to find the optimal design and to evaluate cost- and energy-efficiency. Improved understanding of the performance of a system will additionally reduce the uncertainty in expected energy generation, which again can reduce the risk and the investment cost of the project [25]. Performance predictions are also important for forecasting of the energy generation from PV, which is expected to be necessary for energy systems with a large share of PV.

1.2 Knowledge gaps within PV performance analysis

In the beginning of this research project, the initial plan was to contribute to improvements in PV performance evaluation methodology by developing advanced models for PV systems that could enable automatic detection of performance losses based on output data from the monitoring system. This was not a new idea, considering the number of publications on this and similar topics. However, despite the amount of research on this topic, there is still a lack of standardized methods to accurately evaluate different PV performance aspects. Studies have shown that because of choices done in the analysis process different analysts can achieve different results in for example estimation of degradation rate [26] and in estimation of total performance loss rate (PLR) [27]. The same analyst dependency on the result is seen for PV performance predictions [28]. Before developing advanced methods for performance loss detection, we found that we therefore first should study the cause for this lack of standardized methods, and focus on the following question: Why is it so difficult to determine performance losses in a PV system?

A central part of the answer to this question is that there are many parameters impacting PV performance. Additionally, there are multiple factors that is not related to PV performance, for example related to data quality, that can impact the output data of a PV monitoring system and performance metrics calculations. When different analysts get varying results in for example a degradation rate estimation, it may be because 1) they consider different losses and parameters affecting the performance analysis, and 2) they choose to handle the losses and the effects caused by these parameters in different ways. To develop methods for accurate and reproducible performance evaluation, we therefore need improved understanding and documentation of which factors that impact both the *performance* and the *performance evaluation* in various operating conditions and system designs. We also need improved understanding of which solutions that can be used to handle the effect of the different factors efficiently.

Many of the factors affecting performance and performance evaluation are widely discussed in the literature. Typically, the discussed effects are related to losses that are common in regions where the installed PV capacity has been large for a long time, i.e. at lower latitudes and in warm climates. Losses caused by soiling [29] and component degradation [30] and which factors that impact the evaluation of these losses are for example frequently investigated. For regions with a more recent entry of PV installations, there is less knowledge about how the operating conditions specific to these regions influence performance and performance evaluation. For locations at high latitudes with cold climate, such as in Norway, it is expected that snow, cloudy weather, and irradiance with low intensity and high angle of incidence will affect both performance and performance evaluation. But *in what manner* and by *how much* are not necessarily well documented or known. This introduces uncertainties in the predictions of expected output. Uncertainties in expected output can results in PV systems not being installed, or increase the risk and consequently often the cost of the investment. Inaccurate predictions of PV performance can also have financial consequences if the expected energy generation is overestimated. In 2021, it was for example discovered that the PV installation at the University in Tromsø, Norway, generated significantly less energy than expected, because losses caused by snow were not considered [31]. Because of this underproduction, the system owner required compensations from the system installer.

1.3 Thesis scope and methodology

The main topic of this work is methodology for performance *evaluation.* Improved performance evaluation is also essential for improved performance *predictions.* The aim is to contribute to the development of standardized methodology for performance evaluation by investigating various factors impacting PV performance evaluations and assess solutions that can handle the effect of these factors. The work is concentrated on effects found in cold climates and at high latitude, and the presented analysis is based on monitoring data from multiple Norwegian PV installations. In addition to an investigation of which factors that affects performance evaluation in these conditions, particular attention is paid to the effect of snow, which can significantly impact PV performance in cold climates and challenge accurate identification of other losses. The overall methodology of this project is based on the following main principles: First the problem areas were identified through evaluation of existing solutions. Second, improved solutions are suggested and evaluated. Following this methodology, the following research questions were identified:

- What are the factors impacting performance evaluations of PV systems in high latitude, cold climate locations?

– What are possible solutions to handle these impact factors?

– What is the effect of snow in PV systems?

– Can losses caused by snow be identified and predicted?

The overall aim of discussing these questions, is to contribute to improved PV performance evaluations and predictions specifically in high latitude, cold climate locations. However, several of the different discussed aspects are transferable to other operating conditions, and I therefore hope the work also can be useful for reducing costs and improving PV performance in a more general context.

1.4 Thesis structure

This thesis summarizes my work as a Ph.D. student and consists of seven papers and six introductory chapters. In this first chapter, I describe the motivation for this work, the overall aim of the research, and summarize the papers. The purpose of chapter 2-5 is to contextualize the research described in the papers this thesis is based on. An overview of these four chapters and which of the papers that contribute to the main discussions in these chapters are given in Figure 2.

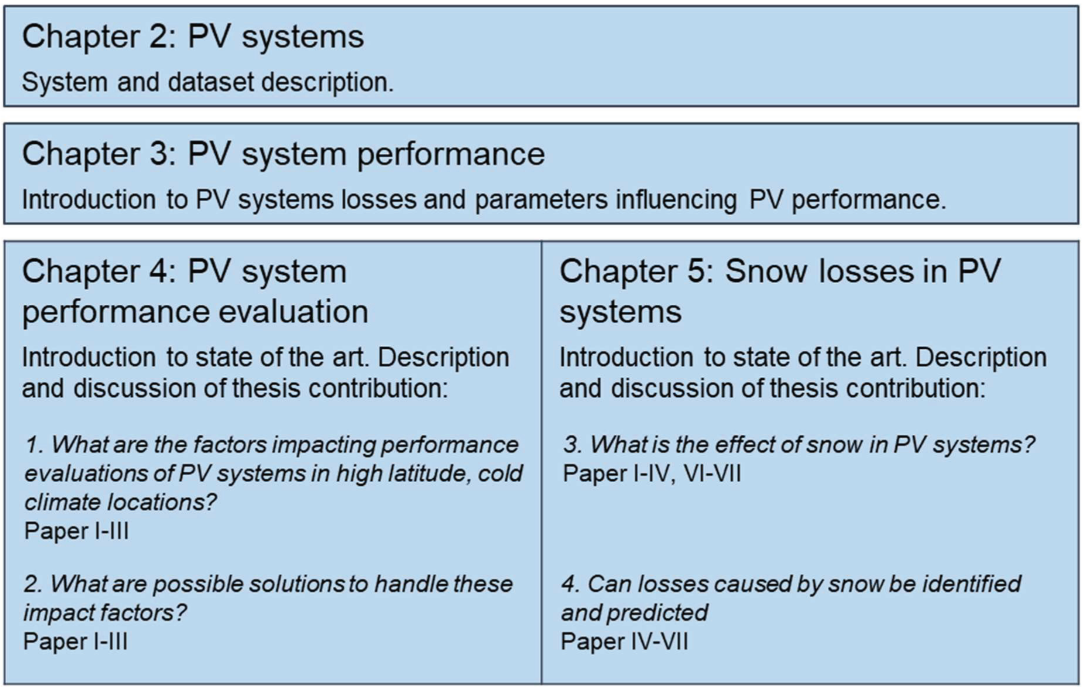


Figure 2: Overview of the four chapters where the research is contextualized, and how the published papers relate to the topics discussed in these chapters.

Chapter 2 provides a description of the PV systems evaluated in this work and an overview of the analyzed datasets. Chapter 3 describes why we need to know how a PV system perform and which parameters that impact PV system performance. Chapter 4 introduces the methodology for evaluation of PV system performance, and the various challenges related to accurate and reproducible performance evaluation, as well as possible solutions. The contributions of the thesis papers on identification and handling of factors impacting performance evaluation are also described. Chapter 5 gives an overview of how snow affects both PV performance and performance evaluation and prediction. Description as well as prediction of snow losses in PV systems are discussed, and our work on these topics is summarized.

1.5 Summary of papers

This section summarizes the papers this thesis is bases on, listed in *List of papers*. I was the main author of all the papers. I was responsible for the main idea, data processing and analysis, and writing the text.

- **Paper I** investigates the possibility for evaluating the quality of the sensor data from the PV monitoring system based on analyzing the measured output. Analyzing PV monitoring data is a central part of the methodology in the papers of this thesis, and the work presented in this paper is an important contribution to this analysis. The main contribution of this paper is validation of the use of clear sky modeling to evaluate irradiance data quality.
- **Paper II** investigates and identifies the root cause of detected performance deviations for a PV system in Norway. The paper also discusses how filtering can be used to improve condition monitoring and fault detection by removing expected losses/deviations in performance metrics. Having an overview of effects that need to be considered in analysis of PV monitoring data is an important first step to develop standardized methods for performance evaluation and fault detection. We also evaluate how typical filters suggested in the literature perform in the operating conditions typical for Norway.
- **Paper III** builds on Paper II and continues the evaluation of the root cause of the detected performance deviations and the assessment of filtering as a solution to process performance metrics for use in fault detection. In this paper, both the dataset and the set of performance metrics evaluated are extended. In addition to strengthening the conclusions from Paper II, we find that using machine learning to model expected output can contribute to improved fault detection.
- **Paper IV** investigates the effect of snow on PV systems, and both snow loss modeling and filtering are tested. The effect of snow is described by evaluation of signatures in monitoring data and simulations of IV curves. In the testing of existing snow loss models, we find that the model suggested by [32] yields best results, and we suggest additional improvements. Based on the snow signatures and the snow loss modeling results, new strategies for snow loss filtering are suggested.
- **Paper V** extends the work on snow loss modeling, and validates the findings from Paper IV. On an extended dataset with multiple different systems in different climatic conditions, the Marion snow loss model with the suggested improvements are tested.

Testing snow loss modeling on multiple systems in different operating conditions are essential to demonstrate the applicability of the snow loss model.

– **Paper VI** is an extended version of paper V. The validation datset is expanded, and it is also evaluated if the data signatures described in Paper IV are valid for the extended dataset. An additional contribution of this paper, is a review of the literature relevant to describe the accumulation and clearing of snow on PV modules.

– **Paper VII** aims to estimate snow loss for PV systems in Norway. Modeled snow loss is used to complement historical data, and the paper discusses how snow loss modeling could be employed to give improved understanding of typical and extreme snow loss values as well as inter-annual variation in monthly and annual loss for locations where long time series of historical data are not available. The paper also discusses how snow loss modeling should be implemented in yield predictions of future PV systems.

2 PV systems

This chapter provides a description of the PV systems evaluated in this work. The aim is not to describe how the different components (solar cells, bypass diodes, inverters, maximum power point trackers, etc.) work or how they are impacted by irradiance and temperature. This is explained in detail in for example [33,34]. The aim is additionally not to describe *all* the components of a PV system, but to focus on the main components and concepts relevant for the research. Section 2.1 gives a general description of the analyzed PV systems, while Section 2.2 presents the details for the analyzed systems, such as geographical location, installed capacity, available sensor data and length of the time series.

2.1 General system description

Energy generating systems based on solar cells exists in multiple forms. In addition to the variation in solar cell technology, PV systems can specifically be designed for different applications. PV applications/technologies expected to play a role in the future [15], include for example floating and building integrated PV, PV specially adapted to agriculture, PV systems with trackers, or systems with bifacial PV modules. It is not uncommon that systems adapted for a specific application or different PV technologies have specific losses or gains, and procedures for PV performance evaluations and predictions for these PV applications and technologies are also needed. In this work, however, the main focus is on performance evaluation of a very basic type of PV systems: grid-connected, fixed tilt systems with monofacial modules, installed on roofs (flat or tilted, as illustrated in Figure 3). The focus is on the output of the PV modules, and the analysis is therefore limited to the DC side of the system.

Figure 3: PV installations on tilted and flat roofs.

Additionally, only PV monitoring that relates to continuous data collection from permanent sensors is considered, and the discussed methodologies for performance evaluation builds on analysis of this type of data. Other types of monitoring techniques for use in PV performance evaluation campaigns do exist, such as imaging and IV-curve measurements. A comprehensive overview of the different techniques is given in [35].

While there are large variations in components and design of PV systems, the systems studied in this work, in similarity with most grid-connected systems, are based on the principles described in Figure 4. As illustrated in the figure, arrays of PV modules are connected to an inverter and a monitoring system.

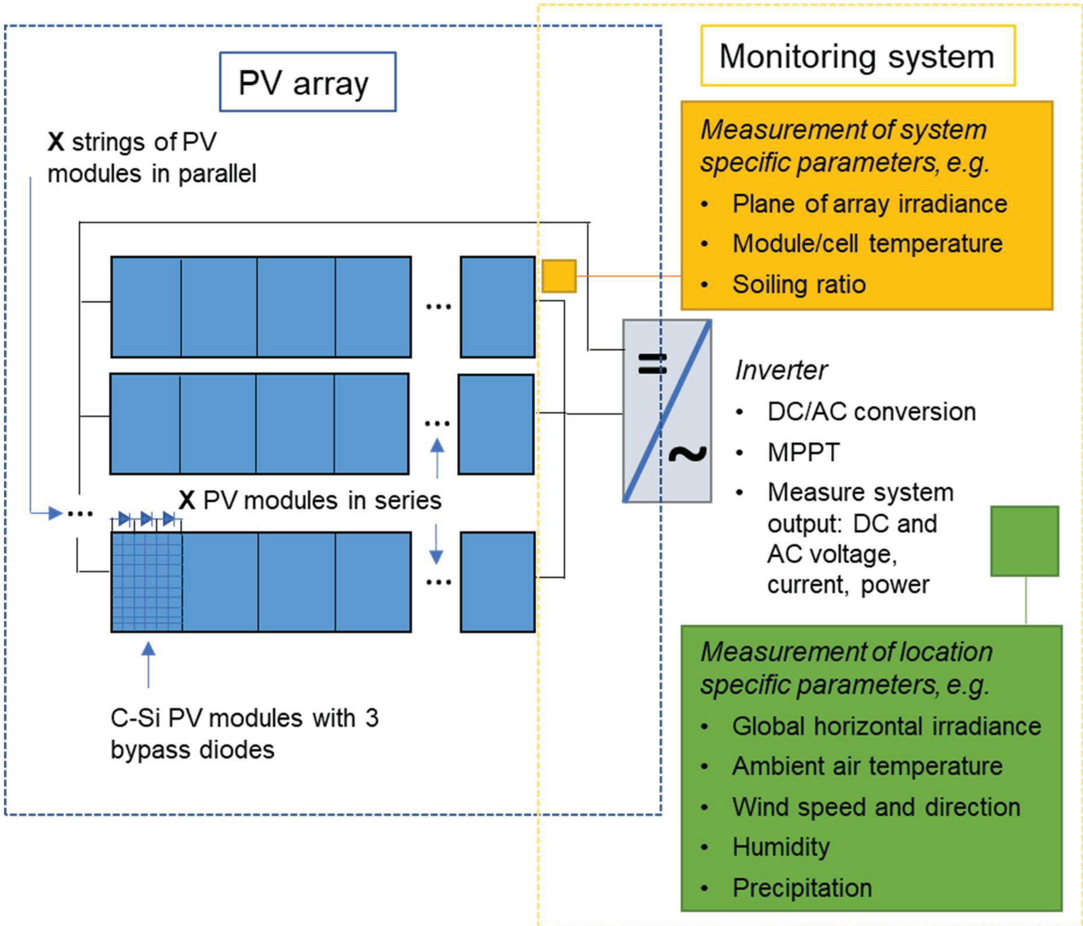


Figure 4: Overview of the PV system and central components in the PV array and in the monitoring system. The list of sensors in the monitoring system is based on [36].

2.1.1 PV array

Figure 4 gives an illustration of a typical PV array [34]. The modules have crystalline silicon (c-Si) solar cells and three bypass diodes. This is one of the most common types of module technologies in existing PV installations, and the type of module installed in the systems studied in this work. A PV array typically consists of series connected PV modules. The strings of series connected modules can additionally be connected in multiple parallels. The array of PV modules is then connected to an inverter. In addition to its main task – to convert the PV module output from DC to AC – the inverter typically also does the maximum power point tracking (MPPT) on the PV array and measure the system DC and AC output.

2.1.2 Monitoring system and sensors

Because the inverter measures the output of the system, it is often also a central part of the monitoring system. Whether the installation has additional sensors measuring influential parameters or not, is a question about cost efficiency. Sensors are added if they are expected to add more value than cost. How many sensors that are part of the monitoring system as well as the quality of the sensors will also depend on their benefit for the system. For a larger system, the cost of a comprehensive monitoring system can be a small share of the total cost, and the value of documenting system performance or detecting performance deviations can be large. For a smaller system, an advanced monitoring system is less likely to be cost efficient.

The irradiance and the temperature of the solar cell are, in addition to the electrical characteristics of the solar cell and module, the main parameters influencing the PV output. The main purpose of most sensors in a PV monitoring system is therefore to estimate these two parameters. Figure 4 presents a list of the sensors suggested in the IEC standard *Photovoltaic system performance – Part 1: Monitoring* (IEC 61724-1:2021). The list is here divided in two types of sensors, 1) sensors measuring parameters specific for the *system*, and 2) sensors measuring parameters specific for the *location*. The parameters specific for the system include plane of array irradiance, module/cell temperature and the effect of soiling on reducing solar cell irradiance. The plane of array irradiance (POA) for the system is typically measured with either a thermopile pyranometer or a reference device. In this work, the reference device installed at multiple of the analyzed systems is a reference cell, where the main component of the sensor is a small solar cell. While the pyranometer measures all the irradiance in the PV array plane, i.e. shortwave radiation (approximately 300 – 2800 nm) in a field of view of 180°, the reference cell aims to measure the *effective irradiance*, i.e. the irradiance the solar cells can utilize after spectral and reflection losses. The module or cell temperature is either measured with a sensor on the rear side of the PV modules, or by measuring the temperature of the reference cell. Soiling sensors are less common and mainly relevant for locations where large soiling losses are expected. The parameters specific for the location that influence PV energy generation listed in Figure 4 include global horizontal irradiance (GHI), ambient air temperature, wind, humidity and precipitation. For large PV plants or scientific systems, many or all these parameters can be measured on site. Nearby weather stations or estimates based on satellite data are also commonly used sources for these types of data.

2.2 Datasets evaluated in this thesis

Different types of datasets are used in the work presented in this thesis, but all are from PV systems located in Norway. The evaluated systems are quite representative of the PV installations we find in Norway, including large commercial systems on flat roofs with east/west oriented modules with 10˚ tilt, installations on tilted roofs (both residential systems and larger buildings), and smaller research systems. The studied installations are typically located in an area where the climate is, according to the Köppen-Geiger classification [37], *warm summer humid continental climate (Dfb).* However, many of the systems are located in areas bordering to *oceanic climate (Cfb)* or *subarctic climate* (*Dfc),* and a few are also located within these climate zones.

The installations studied in this thesis have variations in both array configuration and instrumentation. Table 1 presents an overview of the different system types in the dataset, with information about installed capacity, start of time series, instrumentation and system configuration. For every system type the table summarizes in which papers data from the given category is used. Figure 5 shows the position of the installations, illustrating the geographical distribution.

Table 1: Overview of the datasets studied in this thesis.

System type	Installed capacity	Start of time series	Instrumentation	Configuration	Paper
Commercial systems, flat roof	3.3 MW (6 systems)	Varies between: 2014-2017	Reference cells (POA + cell temperature).	10° tilt, ~east/west. Array: 3 strings in parallel per MPPT, a couple of MPPT per inverter.	One/multiple systems used in Paper I-VII
Residential systems, tilted roof	24 kW (5 systems)	2018 or 2019	No local instrumentation, only data from nearby weather stations used.	Tilt and orientation follow building. Array: one or two strings per MPPT/ inverter.	Multiple/all systems used in Paper V-VII
Commercial system, tilted roof	70 kW	2014	Pyranometer (POA), module temperature.	Tilt and orientation follow building. Array: unknown.	Paper I
Research system	4 kW	2016	Reference cells (POA, GHI), module and ambient temperature.	28° tilt, ~south, open rack. 1 module per optimizer/MPPT.	Paper I, IV

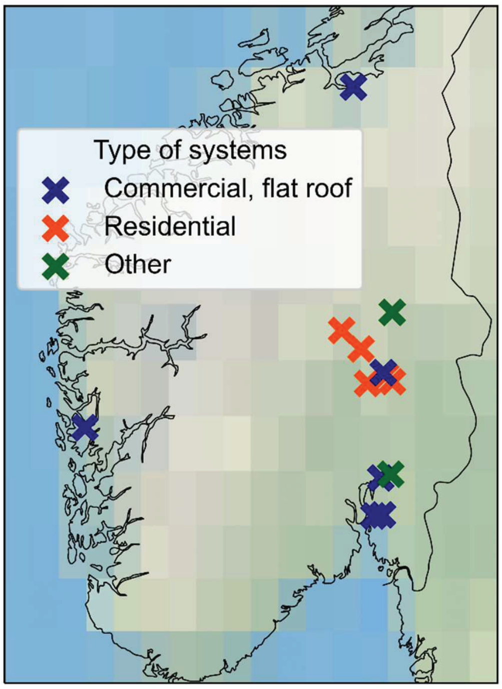


Figure 5: Overview of locations for the installations studied in this work.

3 PV system performance

This chapter describes the effects impacting PV performance and provides the foundation for further discussions on PV system performance *evaluations,* i.e. description and documentation of *how* the systems work and identification and quantification of system losses. In this work, the term *performance* is not used to describe how much energy a PV system will generate, as this is most of all dependent on the available solar resource. The aim of this chapter is to describe 1) the motivation for research on PV system performance, and 2) which parameters that impact PV system performance. A special focus is given to how the operating conditions typical for Norway are affecting PV system performance.

3.1 PV system performance research motivation

As described in Chapter 1, identification and documentation of the losses in PV systems are essential for assessing the potential for performance improvements and for predictions of PV energy generation, which both can contribute to reduction in costs and in increase in the performance of PV installations.

3.1.1 Performance improvement assessments

Assessment of the potential for PV system performance improvements, can both include identification of losses that can be corrected in the operation phase, and losses that can be avoided in the design or planning phase of future systems. This requires accurate quantification of both the losses and the cost of correcting the losses. Many of the losses in PV systems today, could in theory have been removed or significantly reduced in the design phase or through efficient *operation and maintenance* (O&M). To do this would in most cases be classified as over-engineering, and probably neither be cost nor energy efficient. For the sake of cost and energy efficiency, i.e. using our resources in the best possible way*,* we do *accept* a certain level of losses. Certain losses are also *unavoidable*. The only losses and faults we want to *correct* are the ones that are avoidable or recoverable in a cost-efficient manner. However, the line between *acceptable/unavoidable* losses and *correctable/avoidable* losses is moving with technology development, labor costs, energy prices, etc., and is dependent on operating conditions. For example, in a satellite, there are few other options for energy generation and challenging to repair broken components. The willingness to invest in efficiency and reliability is therefore high. In the case of a residential system, on the other hand, the house owner will not wish to pay more for the system than what she can save on the electricity bill. Competing with other energy sources in a sustainable energy system additionally requires consideration of energy payback time and the environmental impact of the potential improvement. To evaluate the potential for improvements can hence be a complex calculation, and accurate identification and quantification of losses would aid this process.

3.1.2 Predictions of energy generation

Identification of losses in PV systems are also necessary to improve predictions of PV energy generation. Estimations of losses in a system for various operating conditions are essential input in the prediction of PV system energy generation on both short and long time-horizons. Hence, estimations of the PV system performance are important input to the models of PV output used in for example PV forecasts and yield estimations. PV forecasting is predictions of PV energy generation in the near future and its importance is expected to increase with an increasing share of PV in energy systems and grids [38]. Yield estimations, which are predictions of how much energy a certain PV system will generate in its lifetime, are necessary for designing and dimensioning of the system. Yield estimations are also necessary for estimations of techno-economic KPIs for a system, such as levelized cost of electricity (LCOE) and return on investment (ROI), and in system life cycle analysis (LCA). Both improved accuracy of forecasting and yield estimations can reduce system costs. In multiple energy markets, errors in the forecasts of energy delivered to the grid can have a cost and reducing forecasting errors is consequently expected to reduce this cost [39]. Accurate yield estimations will ensure optimal system design with respect to energy generation and eventual matching with load profile. Additionally, accurate yield estimations reduce the uncertainty in calculated KPIs and can result in reduced risk and uncertainty in the investment phase, which can reduce the investment cost [25].

3.2 Parameters affecting PV system performance

A wide range of loss mechanisms impacts PV performance. The type and severity of losses in a system is additionally affected by *operating conditions* (including local weather and environment) and *system design* (including technology choice) – which are the parameters that impact the amount of lost energy, and *O&M* - which are the routines and practices implemented to reduce and correct losses for an existing system.

3.2.1 Loss mechanisms

Table 2 presents an overview of different PV system loss mechanisms often described in the literature. The loss mechanisms are divided in four categories: pre-cell losses, component losses, system losses, and degradation and failures. Pre-cell losses are defined as losses caused by effects giving reduced solar cell irradiance. Component losses are expected for all the components in the system that is used for energy conversion or transmission. The quality of the components can, however, impact the magnitude of the losses. The system loss category is here defined as losses that occur because different components are connected. For example, if modules with different capacity are series connected, the module with the lowest capacity will limit the output of the string, giving mismatch losses. In an energy system, PV output can be curtailed if the output exceeds the needs of the system. Losses due to component degradation and failure are expected and considered in the estimation of expected lifetime of the components. All components will experience wear and tear that can give degradation in their performance, and ultimately failure. However, not all cases of component failure and degradation are within the expectations. This can originate from a weakness within the component, or it can be caused by operating conditions that are harsher than the components are dimensioned for.

Table 2: Overview of commonly described loss mechanisms in PV systems [40–43]*, and range of reported annual losses estimated in performance evaluations or used in yield modeling.*

Loss mechanism	Range of expected/reported annual losses	
Pre-cell losses: irradiance attenuation		
Reflection	0–10%	[41,44]
Shading	0–4%	[40,41]
Snow	0–30%	[45]
Soiling	0–7%	[40–43,46]
Component losses		
PV module Conversion Thermal	 ~80% (~20% c-Si module efficiency) 0–15%	 [15] [41]
DC/AC wiring, connections	0–7%	[40,41]
Inverter Inverter efficiency MPPT efficiency Sizing losses, clipping	1–3%	[41]
Transformer	1–2%	[41]
System losses		
Mismatch	0–2%	[40–42]
Curtailment	0–13%	[47]
Degradation and failures		
Module faults and degradation: PID, LID, delamination, cracked cells, corrosion, discoloration of laminate, broken interconnects/solder bonds, frame breakage, junction box failures, backsheet failures, etc. [48,49]	Median annual overall degradation rate: 0.5 %	[30]
Inverter faults and downtime	Depends on repair time	[50]
Faults in other components: connections, cables, etc.	Depends on repair time	[50]

The table also shows a range of reported expected annual loss for the different mechanisms, based on results from performance evaluations and reports on loss values commonly used in yield modeling. The wide range of reported or expected annual loss can be related to 1) the fact that there can be large variation in the loss caused by different loss mechanisms for different systems and in different operating conditions, and 2) limited documentation of certain loss mechanisms can give variation in what different PV modelers expect in losses for a given loss mechanism.

3.2.2 The effect of system design, operating conditions and O&M on losses

System design and operating conditions are expected to influence the type and severity of the losses in a system. Table 3 gives examples of how operating conditions and system design can influence both the cause and the magnitude of the loss for the loss mechanisms listed in the pre-cell category in Table 2. Shading of the modules is caused by objects blocking the direct irradiance of the PV modules. The magnitude of shading losses will depend on both how the system is designed to handle partial shading (array configuration, bypass diodes), and the diffuse/direct shares of the irradiance. The accumulation of soiling on the module surfaces will primarily depend on available sources for dust and dirt in the environment [46]. Precipitation can give natural removal of soiling. Humidity, wind, ambient temperature, and module tilt and surface are also expected to impact the soiling process [51]. While accumulation of snow cover is primarily occurring after snowfall, the natural clearing of snow is impacted by multiple different parameters related to weather and snow conditions, and system design [45]. For cases with partial snow cover, the energy loss will depend on how well the system responds to partial shading. The reflection losses will depend on the glass and cell surface technologies and their anti-reflection properties, as well as the angle of incidence of the irradiance which relates to both solar position relative to the module plane and the diffuse/direct share.

The same complexity with many parameters impacting the loss is also observed for many of the other loss mechanisms. For example, both the ambient temperature and irradiance influence the temperature of a module and consequently thermal losses. But potential heat sources in the system and if the system design enables natural cooling of the modules are also influential [52]. The development of faults and degradations relates both to the robustness of the different components, and the stress they experience, which typically closely relates to the operating conditions [53].

Reduction and correction of losses

For several loss mechanisms, the energy losses can only be avoided or corrected in the design or installation phase of the system, by for example choosing efficient and reliable components, adapting the design to the shading scene, or choosing a mounting ensuring efficient ventilation. For certain loss mechanisms, however, it is possible to reduce the energy loss by efficient O&M. O&M are today considered important to keep the performance of a PV installation on a high level [54]. Loss reducing measures typically implemented in O&M include reducing the shading by soiling, snow or vegetation, and replacement of faulty components [55]. As discussed in Section 3.1.1, estimation of the need for corrective measures is a complex calculation. In addition to accurate identification and quantification of these losses, the cost and gain to correct them, as well as safety or reliability related issues connected to the relevant loss mechanism, also requires consideration.

Table 3: Examples of the effects of operating conditions and system design on the cause and magnitude of the energy loss for the different loss mechanisms in the pre-cell loss category.

Loss mechanism	**Operating conditions**		**System design**	
	Cause	Magnitude	Cause	Magnitude
Shading		Diffuse/direct share of irradiance	Shading objects	Array configuration, bypass diodes
Soiling	Soiling sources	Precipitation, wind, humidity, ambient temperature		Module tilt, surface coatings
Snow	Snow fall	Temperature, irradiance, type of snow		Tilt, array/module/system configuration, module technology
Reflection		Diffuse/direct share of irradiance, solar position	Glass/cell surface	Glass/cell surface technology

3.3 The effect of Norwegian conditions on PV performance

Because operating conditions can have a large impact on losses, the loss mechanisms that are most prominent in different geographical regions can vary. In this work the studied dataset consists of data from PV installations Norway, and the aim is to focus on how the operating conditions specific for Norway impact PV performance and performance evaluation.

3.3.1 Characterization (and prevalence) of the Norwegian conditions

Within Norway, there are significant variations in both climate and irradiance conditions. Certain qualities can, however, be generalized, and are additionally transferable to other geographical regions. The typical aspects of Norwegian conditions relevant for PV are: 1) high latitudes, which affect the irradiance conditions (solar position and irradiance level), and 2) a climate characterized by large seasonal variations, low temperatures and precipitation the whole year, including snow in the winter. The climate zones in the studied dataset (Dfb, Dfc and Cfb, according to the Köppen-Geiger classification [37]) cover well the areas in Norway with highest population density, and where most PV systems therefore are installed. Most of the studied systems are, however, located at latitudes below 62°, and the more extreme high latitude conditions are thus not represented.

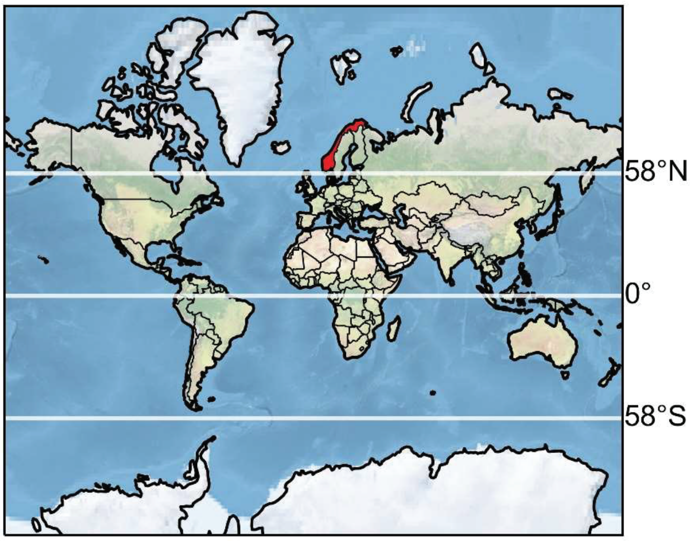


Figure 6: Map of the world illustrating which areas that are located above 58° (marked with white lines), the latitude of the southern point of Norway (marked in red).

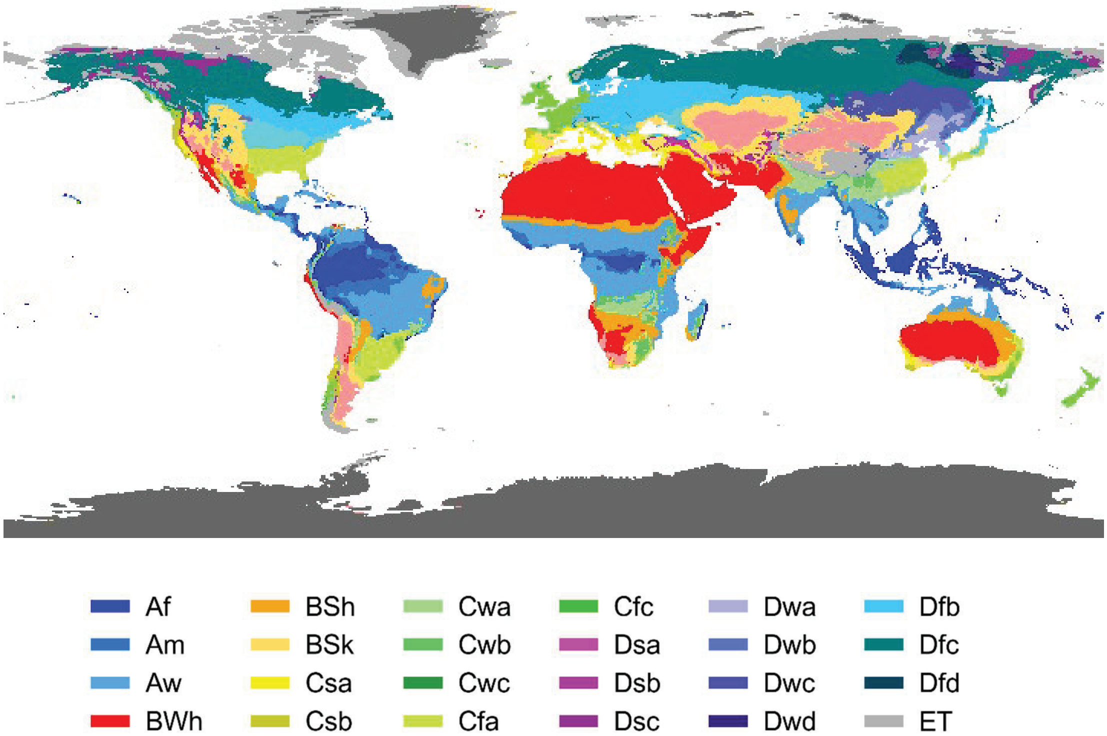


Figure 7: Köppen-Geiger climate classification based on historical (1980-2016) climate data. Figure from [37].

Figure 6 illustrates which areas that are located at the same latitude as Norway, and Figure 7 shows the geographical distribution of the different climate zones according to the Köppen-Geiger classification. From these figures, it can be observed that the combination of high latitude and climate we find in Norway is also representative for large areas in Sweden, Finland, Canada, Alaska and Russia. *Warm summer humid continental (Dfb)* and *oceanic (Cfb)* climates (barely visible in Norway on the map but represented in the coastal and near-coastal areas) are additionally represented in large areas at slightly lower latitudes in Europe and Northern America. The effects of high latitude on the irradiance conditions increases with increasing latitude but are expected to be noticeable in all the mentioned areas.

3.3.2 Impact of Norwegian conditions on PV

As described in [56], the type of irradiance and the climatic conditions in the northern countries impact PV energy generation and performance in various ways. Because of the tilt in the Earth's axis of rotation and its round shape, a high latitude position gives large variation in the day length through the year and low solar elevation. The interannual variations in day length introduces large interannual variation in the solar resource. It additionally has an impact on how the angle of incidence (AOI) of the PV module irradiance varies through the year. For fixed tilt systems, high AOI values for longer periods of the year/day are typical. Low solar elevation gives reduced irradiance and increased *air mass*, i.e. the optical length through the atmosphere of the Earth. The air mass affects both the intensity and the spectrum of the solar irradiance. The climate affects PV through low temperatures, overcast weather and precipitation. Overcast weather reduces the irradiance level and increases share of direct and diffuse irradiance [57]. Precipitation includes both rain and snow. Figure 8 shows how the daily mean solar elevation (for elevation >0), the daily mean TMY temperature and the weekly TMY irradiation of Oslo, Norway, compares to Hamburg, Germany and Milan, Italy. Oslo has lower temperature and solar elevation values than Hamburg and Milan, but the irradiation is similar to the irradiation of Hamburg.

The described operating conditions will affect PV energy generation both through affecting the available solar resource and the PV system performance. The efficiency of the PV modules drops at lower irradiance levels and the efficiency of the inverter drops when the generated power is much lower than the nominal capacity of the inverter. High angle of incidence of the irradiance leads to increased surface reflection and reduced solar cell irradiance. Lower temperatures will on the other hand give reduction in thermal losses, and frequent rain keeps the PV modules clean, reducing soiling losses. Snow blocking the irradiance can give very high losses, but if the modules are snow free, the increased POA caused by increased reflections from the ground can give performance gains. There are, however, still significant knowledge gaps with respect to both understanding and consequently also predicting the effect of snow. The effect of snow on PV performance is therefore given particular consideration in this work, and Chapter 5 is dedicated to this topic. Figure 9 shows how losses caused by reflection and temperature compare for Oslo, Hamburg and Milan. While the thermal losses are lower for Oslo, the reflection losses are higher,

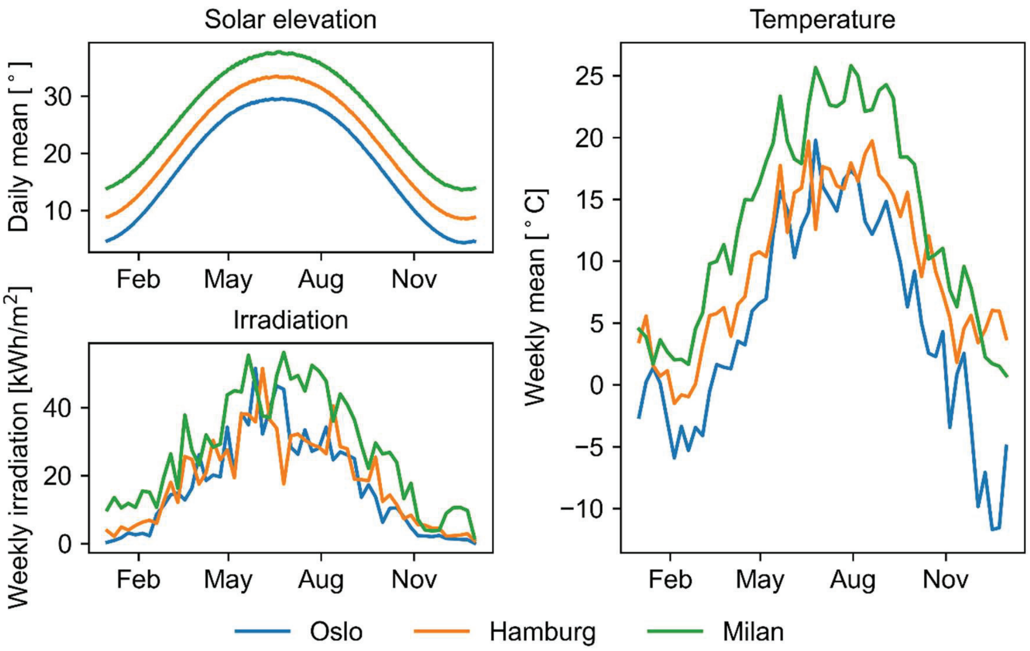


Figure 8: Daily mean solar elevation (for elevation >0), daily mean TMY temperature and weekly TMY irradiation of Oslo, Hamburg, and Milan. Solar elevation calculated with pvlib python [58]. *TMY data from* [59].

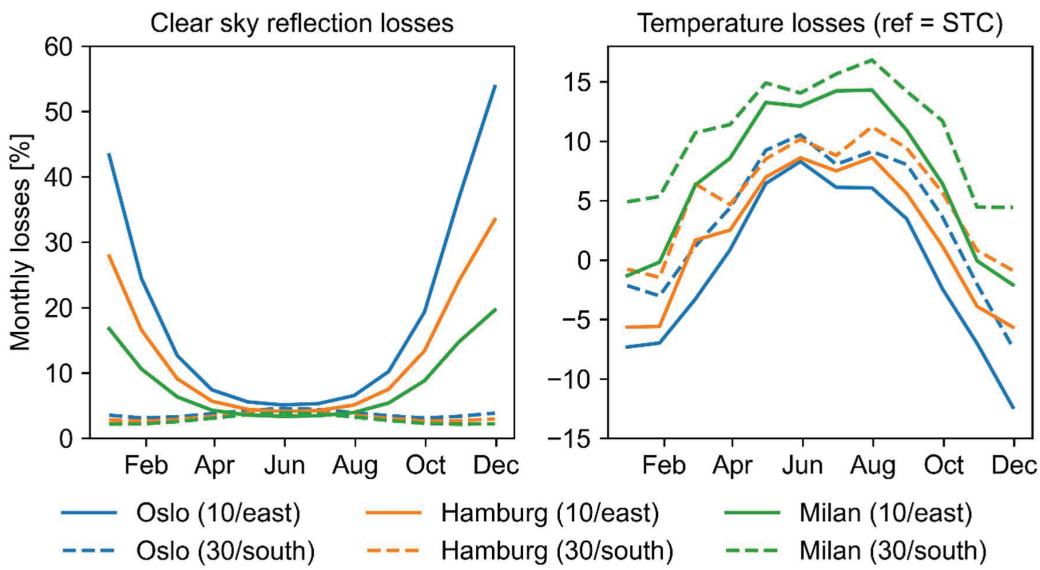


Figure 9: Daily reflection losses (at clear sky irradiance) and monthly temperature losses relative to STC for a system with a) 10° tilt, oriented east, and b) 30° tilt, oriented south, located in Oslo, Hamburg and Milan. The reflection losses are modeled with the pvlib [58] *implementation of the model described in* [60], *and the temperature losses are estimated from TMY data* [59] *with a temperature coeffcient of -0.004.*

especially when the modules are not oriented South. The reflection losses are estimated for clear sky conditions, and is thus expected to represent the maximum reflection loss.

In addition to the described losses, which closely correlates to variations in operating conditions, the climatic and irradiance conditions will also influence the wear and tear of the different components, i.e. how fast they *degrade* and ultimately *fail.* It is commonly known that climatic conditions play a major role in degradation [48,61]. High temperature, humidity and UV are for example expected to give increased degradation. Both analyses based on field data [62] and degradation models [53] shows higher degradation rates in hotter climates. Modeling of the degradation mechanisms related to temperature, irradiance and humidity, confirms that lower degradation rates caused by these stressors are expected at high latitude locations [53]. On the other hand, snow is an additional stressor that can impact degradation rates. Periods with increased mechanical load because of snow can contribute to increased degradation and failures, especially if the PV modules are installed with insufficient mechanical support. Is has been observed that snow loads have resulted in both cell and module fracture [63]. However, the effect of the cracks on PV performance is not necessarily large [63]. Additionally, it is complex to estimate what type of snow cover that gives a load that can lead to cracks. The mechanical load of a snow cover will depend on both the temperature and the homogeneity of the snow cover [64]. Published field data indicates low degradation rates in high latitude/snow climates, but more data is still required to robustly conclude on this [62].

While the theory on the effect of the operating conditions found in Norway on PV performance is well known, there is still a need for studies on existing PV systems providing detailed information on how significant these effects are. Compared to regions at lower latitudes which has had higher PV installation numbers for a longer time, the available data for such evaluations in Norway and similar locations have been limited. Several of the existing performance studies on PV systems in Norway and Sweden are focused on reporting the total system performance for different technologies. Studied technologies include building applied PV, both BIPV systems and systems on flat and tilted roofs [65–69], dual-axis tracker systems [70], and bifacial PV systems [56,71–73]. Shading, snow, and inverter outages are identified as influential loss mechanisms [65–67,72]. It has been found for two different Norwegian installation that the degradation rate is well below 0.5% per year [74,75], the median degradation rate reported by [62] for PV installations worldwide. Ref. [57] models how the measured irradiance and temperature in different locations in Norway impact the effective efficiency for different module technologies. Field data on comparing different module technologies for the given operating conditions is provided in [56,76], where CIGS and c-Si modules are compared.

4 Methodology for evaluating PV system performance

To evaluate PV performance based on the measured data in the monitoring system, a *performance metric* is typically calculated. How this metric is best defined and evaluated, depends on the application, i.e. if the aim is to determine the overall system performance or identification and quantifications of specific losses. Evaluating PV performance is, however, not straightforward. Various factors can impact both PV performance and the evaluation itself. These factors are not necessarily the same in all systems and operating conditions. A consequence of this is that methods to evaluate different performance aspects can be adapted to specific operating conditions or system designs, which gives a range of methods available for different tasks. To enable fast and easy implementation of performance analysis in automatic systems, we need *standardized* methods and procedures with broad applicability.

This chapters describes existing methodology for performance evaluation and discusses challenges and solutions. Section 4.1 describes the methodology for performance evaluations of PV systems based on monitoring data. Section 4.1.1 describes commonly used performance metrics and Section 4.1.2 discusses how the metrics are evaluated and used in different applications. In Section 4.1.3, standardized procedures for PV performance evaluations are discussed. In Section 4.2 factors impacting performance evaluation are classified, and potential methods to handle these impact factors are discussed in Section 4.3. The last two sections also describe how our research contributes to the classification and handling of impact factors.

4.1 Performance evaluation based on monitoring data

4.1.1 Performance metrics

Both PV system losses and PV energy generation are highly dependent on operating conditions and system design. Within the *performance metrics* used to quantify and evaluate PV performance, the output of the system is therefore typically compared to a *reference* or a *yield target* [77]. This section describes three different types of commonly used references and gives examples of how these reference types are used in the performance metrics defined in the IEC standard *Photovoltaic system performance – Part 1: Monitoring* (IEC 61724-1:2021). The performance metrics defined in this IEC standard are often implemented in PV system monitoring software. The three references are 1) *other systems or system units,* 2) *impactful parameters* and 3) *model of the system or system units/components*.

Other systems or system units/components

The electrical output (current, voltage, power, energy) of a PV module, PV array or inverter can be compared with other units of the same type within a system [77] or to the output of other systems [78]. If the units are not identical, this is done by normalizing the output to the rated capacity of the units. For the ratio between energy output (E_{out}) and the power rating (P_0) for a PV system, array or module, the terms *yield*, *energy yield*, or *specific yield* are commonly used. In the IEC standard [36], the final system yield (Y_{f}) is defined as:

$$Y_f = E_{out}/P_0. \qquad (1)$$

When comparison of output is used as a performance metric, the required similarity in design and operating conditions for the compared arrays or systems, depend on the application of the performance evaluation. For an evaluation of the overall system performance, it is sometimes useful to compare the yield of systems or arrays with different designs or operating conditions. To use comparisons of array specific yield to detect component faults within a system, the compared arrays should have the same configuration (tilt angle and orientation) and operating conditions (irradiance and temperature).

Impactful parameters (such as irradiance, temperature)

Solar irradiance is the parameter that most of all affects the energy output of a PV system. The irradiation (*H*) in the plane of the PV array (H_{POA}) is therefore a natural parameter to include when evaluating PV performance. For comparison of PV yield to plane of array irradiation, the term *performance ratio* (*PR*) [36] is commonly used [78,79]:

$$PR = (E_{out}/P_0)/(H_{POA}/G_{ref}). \qquad (2)$$

G_{ref} is the reference irradiance (*G)* at which P_0 is determined. For standard test conditions (STC), G_{ref} is 1000 W/m^2. For comparison of instantaneous values, energy can be replaced with power (P_{out}) and irradiation with irradiance (G_{POA}). As temperature has high impact on the PV system output, the *temperature corrected performance ratio* (PR'_T) has also been suggested for performance evaluations [80]. For temperature correction with respect to the STC temperature of 25 ℃ (T_{STC}), this is defined in [36] as:

$$PR'_{25°C} = (E/(P_0 * (1 + \gamma(T_{mod} - T_{STC}))))/(G_{POA}/G_{ref}). \qquad (3)$$

The material dependent module power temperature coefficient γ, determines the loss or gain in power caused by temperature changes. Comparison with an impactful parameter could also be done for other system output parameters: temperature could for example be used as a reference for output voltage, and irradiance could be used as a reference for output current. The performance ratio is commonly used to quantify the system performance and the total losses in the system [78], and comparing for example voltage and current to other influential parameters have been suggested to identify different loss mechanisms [79].

Model of the system or system units/components

A model of the output of the system or system components is another potential reference for the measured output. The term *performance index* (*PI*) is defined in [36] as:

$$PI = Measured\ output/Expected\ output. \qquad (4)$$

A wide range of models for the expected output of PV systems exists, from models of the different components to full system models. Both physical and empirical/machine learning models are suggested [27,77,81,82]. The losses that should be included in the model will depend on the application/purpose of the performance evaluation. For example, to evaluate if the system works as planned, all the expected losses accounted for in the dimensioning of the system should be included. Quantification of the energy loss caused by one specific loss mechanism requires accurate estimations of all other losses.

4.1.2 Evaluation of performance metrics

While the calculation of the described performance metrics is straightforward, the *evaluation* of the performance metrics to determine potential for improvements and identify specific loss mechanisms can be done in various ways. This section describes how the references described in the Section 4.1.1 could be utilized if the aim is 1) *estimation of system performance*, or 2) *identification of loss mechanisms*.

All the suggested references can be used to *quantify* the overall *system performance*. The choice of reference will, however, impact the type of information achieved from the performance evaluation. The system yield will for example describe how a specific system with the given design and given operating conditions compares to other systems. With the performance ratio, the effect of varying solar resource is removed, and all the losses in the system are quantified. If other expected losses are considered the loss estimation will be more specific. For example, the temperature losses could be corrected for. Including an increasing number of weather effects impacting PV performance, results in an improved description of how the systems work independent of weather conditions. A detailed model of the system with all expected losses enables comparisons between expected performance and actual performance.

To *identify* different *loss mechanisms*, for example the mechanisms previously described in Table 2, and detect potential for improvements in system performance, comparing the output to one reference alone is not always enough. Often multiple variables must be considered to identify different loss mechanisms. This can be relevant for both *detection* and *diagnosis* of performance loss. With losses of a few percent, for example caused by module faults and degradation, detecting the performance deviation in itself can be challenging. A typical parameter that is often used to detect the losses caused by module fault and degradation, is the development of the performance metric *with time*. For example, to both identify and quantify the total performance loss because of faults and degradation, the development of the performance metrics over multiple years can be evaluated [27,83]. We show in [84] that calculation of the cumulative energy losses over time can give fast detection of faults that give small relative losses. If the performance loss caused by a component fault is not gradually increasing, but the onset is more abrupt, the component fault can be identified by detecting steps in the performance metric time series [85].

Describing *signatures* or *fingerprints* in the data is an approach that is commonly used to identify specific loss types [79,86]. Often, multiple parameters can be evaluated to identify or detect one specific loss mechanism, and there are multiple pathways that can be followed to answer the relevant questions. The suitability of the different signatures will depend on data availability (and the quality of the available data), and the system design and operating conditions. The influence of system design and operating conditions follows from the fact that both can impact the energy losses related to different loss mechanisms, as previously illustrated in Table 3.

To show an example of how signatures in the time series data can be used to identify different loss mechanisms, Figure 10 illustrates the difference in the response in the monitoring data timeseries for shading caused by snow and shading caused by a fixed object. The plane of array irradiance is here used as a reference for the expected current output, as these quantities are expected to be directly correlated. The snow shading loss is here gradually recovered with time, related to how fast the snow melts. The shading loss caused by a fixed object occurs at a specific time of the day and is larger at clear sky conditions. As described in Section 3.2.2., the energy loss caused by shading from a fixed object will depend on solar position, and it will increase with increasing direct irradiance. This means that shading can be identified by relating the losses detected in for example PR'_T or PI with solar position and clear sky index. Shading can also be identified by comparing the output power or current with a reference, such as an irradiance measurement or another PV module/array and evaluate how the differences change with time. Snow will give large losses with similarity to both full and partial shading, but how snow losses develop with time is following the development of snow cover, which is related to various weather parameters. Identification of snow losses will be further discussed in Chapter 5.

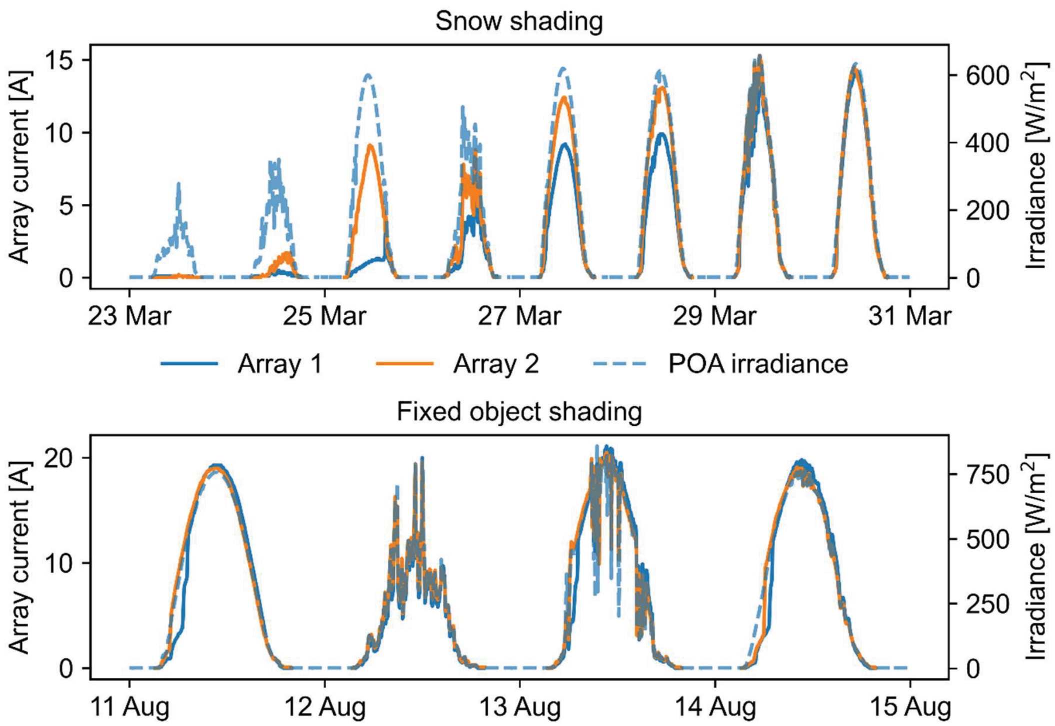


Figure 10: Illustration of the difference response in the monitoring data timeseries for shading caused by snow and shading caused by fixed objects.

4.1.3 Standardized performance evaluation methodology

A consequence of the multiple possible pathways in performance evaluation, in combination with variations in available output data from monitoring systems, is that it is challenging to develop *standard analysis procedures* to either estimate overall system performance, or to quantify specific losses. The IEC technical specification *IEC TS 61724-3:2016 Photovoltaic system performance – Part 3: Energy evaluation method* suggests a procedure for comparing the actual energy generation of a system with the expected energy generation for the given weather and irradiance conditions. However, in this procedure there are many choices that must be made and many effects to consider that are dependent on the system and the purpose of the evaluation. Due to the need for evaluating the available input data and the methodology, an accurate performance evaluation of a PV system will require an experienced analyst. Relying on a human analyst does, however, open for variation in interpretation of the data. In a study where a group of different analysts were given the task of estimating the PV degradation rate in the same dataset, the obtained results had significant variations [26]. In this study, the authors evaluate many of the choices that could be made in a degradation rate estimation. Specifically, decisions related to data filtering can have large implications on the results. It has been shown that the results of an estimation of total PLR [27,87] and a module degradation analysis [88] can vary with use of filtering, performance metrics, and models of the expected output. It has also been shown that the output of yield predictions can depend on both the modeling tool and the modeler [28]. More standardized procedures for evaluation, and also prediction, of PV performance are clearly needed. Standardized procedures would give analyses with improved repeatability and enable less experienced analysts to do the evaluation as well as automation and implementation of the analysis in monitoring platforms. This could give cheaper and faster evaluation of PV system performance.

There are examples in the literature of work contributing to standardized methodologies for PV system performance evaluation. These standardized methods are typically targeted for specific applications, such as fault detection or degradation rate estimation. Standardized procedures for fault detection or degradation estimation are for example suggested in [84,88–90] where steps for processing the data, calculation and evaluation of performance metrics are suggested. RdTools is a framework developed to enable robust [83] and reproducible [26] degradation rate estimation, where data filtering and a year-on-year analysis procedure are central elements. Ref. [91] suggests estimating degradation rate using an unsupervised machine learning approach based on estimating, based on the measured power, what the clear sky power output would have been and then calculate the change in this estimated signal from year to year. This methodology utilizes the systematic trend in the clear sky irradiance through the day and the year. Together with the suggested automatic cleaning procedure [92], the method is fully automatic and requires only power output data.

Ref. [27] tests several procedures for performance loss estimation on multiple different datasets with the aim of standardizing performance loss rate estimation. The procedures typically include data cleaning and filtering, choice of performance metrics and statistical evaluation of the metric. The authors conclude that a perfect procedure that can be used for all systems is probably not existing, and they recommend that the choice of filters, performance metrics, etc., should be adapted to the dataset characteristics. To adapt procedures to different datasets, the different effects in a dataset that can impact the performance analysis should be well known, as well as the methods to handle these effects in different operating conditions and with different system designs. The authors of ref. [27] also suggest that the performance loss rate could be estimated through an ensemble approach, where multiple different procedures are used, and the average output value is used as the estimate. A potential issue with this approach is that if there are losses or other factors impacting the performance loss rate calculation that are not targeted by most of the methods, this could bias the result. To use this approach, it is therefore also important that the effects that potentially could impact the performance evaluation in different datasets is well known, as well as how the different procedures handle these effects and losses. To identify impact factors, and implement methods that efficiently handle these impact factors, can aid the development of standardized performance evaluation procedures with broad applicability.

4.2 Classification of factors impacting performance evaluation

The aim of this section is to classify the factors that have an impact on PV performance evaluation. Performance evaluation and identification of loss mechanisms are not only challenged by the large number of parameters influencing PV performance, as described in Chapter 3, but also that there are multiple parameters affecting the calculated performance metric that are not related to PV performance. In the papers included in this thesis, factors that affect fault detection in the studied Norwegian installations are analyzed and classified. Based on these results, summarized in Section 4.2.1, a broader classification is discussed in Section 4.2.2. The discussion in 4.2.2 is added to generalize the classification, both with respect to operating conditions and aim for performance evaluation, and to put the results into a broader context and compare with the existing literature.

4.2.1 Thesis contribution

In Paper I-III, we have done an evaluation of which factors that impact performance evaluation of systems in the operating conditions typical for Norway, described in Section 3.1.1. The performance evaluation case discussed in the papers is mainly fault detection. In Paper I, we discuss data quality challenges in PV monitoring systems, and how this impacts performance evaluation in general. The data quality effects discussed are drifts and shifts in the irradiance measurements, misalignment of irradiance sensor, and detachment of module temperature sensor. In Paper II and III we identify different factors that impact the calculated performance metric time series by introducing noise, offset or systematic trends. These effects include expected losses not considered in the performance metrics or factors that give similar signals as performance losses or gains. In Paper III we categorize these factors in the following three categories: 1) invalid data, 2) data quality and availability, and 3) unstable periods. Table 4 presents the identified factors in the different categories for the case of fault detection. The invalid data category includes mechanisms introducing large losses or factors that give a signal similar to large losses. These types of signals are a significant challenge for fault detection because the related impact on the performance metric typically is much larger than the impact of other loss mechanisms. The data quality and availability category summarizes factors that can give systematic trends or offsets in the performance metrics because of the data quality, or because the availability of input data are not sufficient to include these effect. The unstable periods category include factors introducing noise in the performance metric time series.

Table 4: Summary of factors impacting fault detection and/or performance evaluation identified/described in Paper I-III.

Invalid data	Data quality and availability	Unstable periods
Sensor/monitoring system downtime Snow Clipping Curtailment	Lacking quantification of expected system losses because of insufficient system data avilability Systematic differences in irradiance in PV array and/or between PV array and irradiance sensor (shading, variations in tilt angles) Drift/shifts in irradiance measurements Detachment of module temperature sensor	Irradiance with low intensity and AOI, low solar elevation Rapid, large changes in irradiance

These types of impact factors are also previously described in the literature for different types of performance evaluations (fault detection, degradation and performance loss rate estimation, evaluation of overall system performance, etc.) [27,79,83,89,93]. For example, it is not surprising that periods with high angle of incidence (~ 60°) and low intensity of the irradiance (~ 200 W/m^2) would be challenging. These conditions result in losses that are difficult to accurately quantify, as well as increased measurement uncertainties [94]. In analyses where the aim is to identify losses caused by for example faults or degradation, periods with irradiance with low irradiance and high angle of incidence are typically filtered out [27]. This filtering is done because the related noise gives a stronger signal in the calculated performance metric than the faults or the degradation.

The main challenge in the studied dataset is, however, the prevalence of many of these factors. As discussed in Section 3.3.2, periods with large share of irradiance with low intensity and high angle of incidence can last for longer periods and will be more severe at the studied locations than at locations at lower latitudes. Snow can also give large losses for long periods. Clipping and curtailment are highly dependent on system design and grid operation but could be expected to be prevalent in locations like Norway. With large variations in irradiance through the year and low electricity prices in the summer, it can be beneficial not to optimize the system for the few summer months with highest irradiance, and systems can therefore be designed with expected clipping losses.

The effect on the performance evaluation of systematic differences in irradiance within a PV array and between PV array and irradiance sensor because of small variations in tilt angles are not widely discussed in the literature. For PV modules that are installed on racks on the ground or on flat roofs, all the modules will not always have the exact same tilt because the surface they are installed on not always will be completely flat. To measure the effective plane of array irradiance of an array with small variations in tilt will be difficult. Additionally, even if the modules have the exact same tilt, the irradiance sensor can be inaccurately installed with respect to tilt and orientation. Ref. [36] recommends a sensor alignment accuracy of 1° in a high accuracy monitoring system. Figure 11 shows the relative difference between the daily clear sky irradiation of planes with respectively 1- and 2-degrees difference for Oslo, Hamburg and Milan, and how this difference varies through the year. In the figure we observe a seasonal variation in the difference, and that this variation is largest for Oslo and low tilt angles. The difference will be largest at clear sky when the share of direct irradiation is at maximum, and the figure consequently illustrates the maximum difference in irradiation for the given tilt angles. The seasonal variation in the irradiation difference between two planes with small deviations in tilt, is expected to give a seasonal effect on a performance metric when the utilized irradiance has a tilt that deviates from the PV arrays. Figure 12 shows how the *PPI* of a system can vary through the day when the irradiance sensor has a tilt that is 1 and 2 degrees steeper than the tilt of the PV array. The systems are modeled with no losses, and the *PPI*s are thus expected to be equal to 1. Despite the small absolute differences in irradiation between the different planes, the effect when comparing to the PV output is large. Because of the deviations in tilt in the irradiance, the performance metric indicates performance gains or losses in the morning and in the afternoon. We see that this trend is stronger for the day with clear sky conditions.

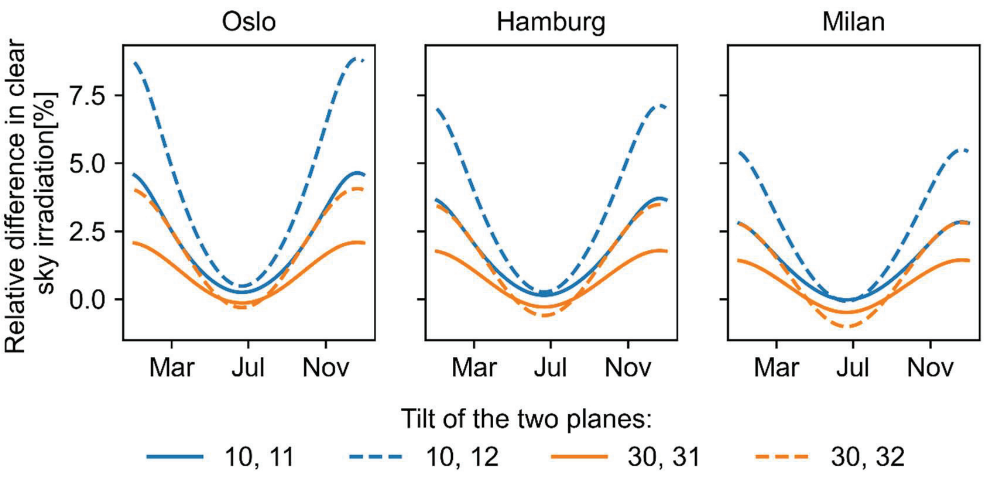


Figure 11: The relative difference between the daily clear sky irradiation of two south faced planes with respectively 1 or 2 degrees difference.

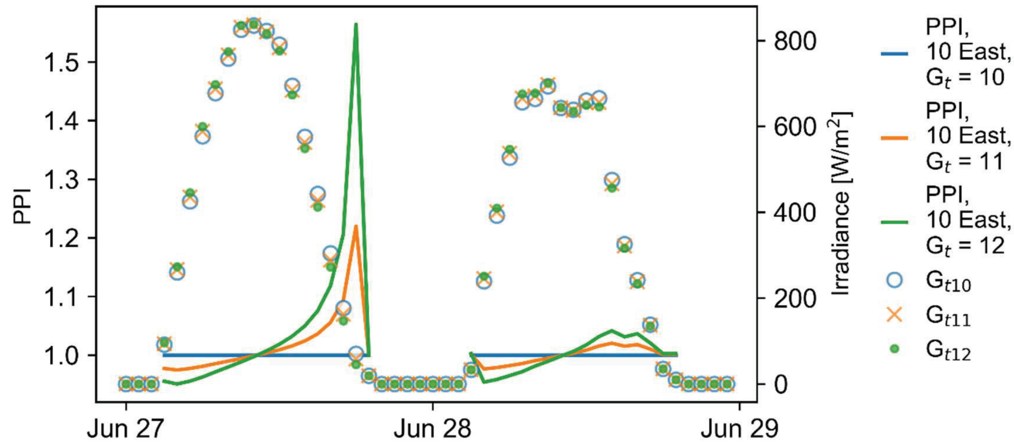


Figure 12: The calculated power performance index of a system located in Oslo with moduels with 10° tilt oriented East, calculated with the irradiance (G) of planes with a tilt (t) of 10, 11 and 12 degrees. No losses are included, and the PPI is thus expected to be equal to 1.

4.2.2 General classification

The categorization of impact factors presented in the previous section is specifically targeted for fault detection. However, these impact factors could be classified more broadly to also enable implementation in other types of performance evaluations. Generally, the discussed factors were either related to:

– Data quality (drift/shifts in irradiance measurements, detachment of module temperature sensor).

– Non representative references used in the performance metrics (systematic irradiance differences in the system caused by small variations in tilt, rapid changes in irradiance).

– Expected losses not (accurately) considered in the performance metric (snow, curtailment, clipping, shading, irradiance with low intensity and high angle of incidence).

This section elaborates further on these three categories. This is done to generalize the classification of factors impacting performance evaluation, with respect to both operating conditions and aim for performance evaluation, and to put the results into context and compare with the existing literature. While classification of factors impacting performance evaluation is not widely discussed in the literature, many of the different factors are previously evaluated. Table 5 summarizes the impact factors found in our work supplemented with factors described in the literature and gives examples of how the factors can impact the calculated performance metric. The categorization is not considered to be absolute, and the impact factors can in several cases fit into more than one category.

Table 5: Overview of the different factors impacting performance evaluation, and common consequences of these factors in the calculated performance metric and the data analysis.

Cause	Possible effects on performance metric
Data quality	
Missing data	Holes in time series
Erroneous data, e.g. stale or unphysical values, duplicated data, incorrect time stamp	Erroneous results because logged data not representative for measured physical quantity or time stamp
Non calibrated sensors	Missing information on uncertainty in performance metric
Temperature sensor with poor thermal contact	Measured module temperature systematically too low, giving systematic trends in performance metric
Drift or shifts in irradiance measurements	Drift or shifts in performance metric
Soiling/shading on irradiance sensor	Indication of performance gain, systematically and/or random (depending on soiling types)
Sensor uncertainties	Noise, offset in performance metric time series
Non-ideal references	
Choice of sensor Placement of sensor Inadequate reference model	Noise, offset, systematic trends in performance metric time series if measured or modeled value is not representative for the quantity it represents
Inaccurate quantification of expected losses	Noise, offset, systematic trends in performance metric time series

The typical effect these factors can have on the performance evaluation, is to induce noise, shifts, offsets or systematic trends in the calculated performance metrics. Parts of these signals in the performance metrics will be related to actual losses, but other parts of the signals can be related to data quality issues or the use of non-representative references used in the performance metric. Figure 13 shows daily PR'_{T} for three years, and 5-minute PR'_{T} values for six days for one of the studied systems. For both time resolution, we observe *noise* and *systematic trends.* On high time resolution there is a *daily* systematic trend, and on lower time resolution we observe a *seasonal* systematic trend. The seasonal trend is a commonly observed systematic trend in calculated performance metrics [88].

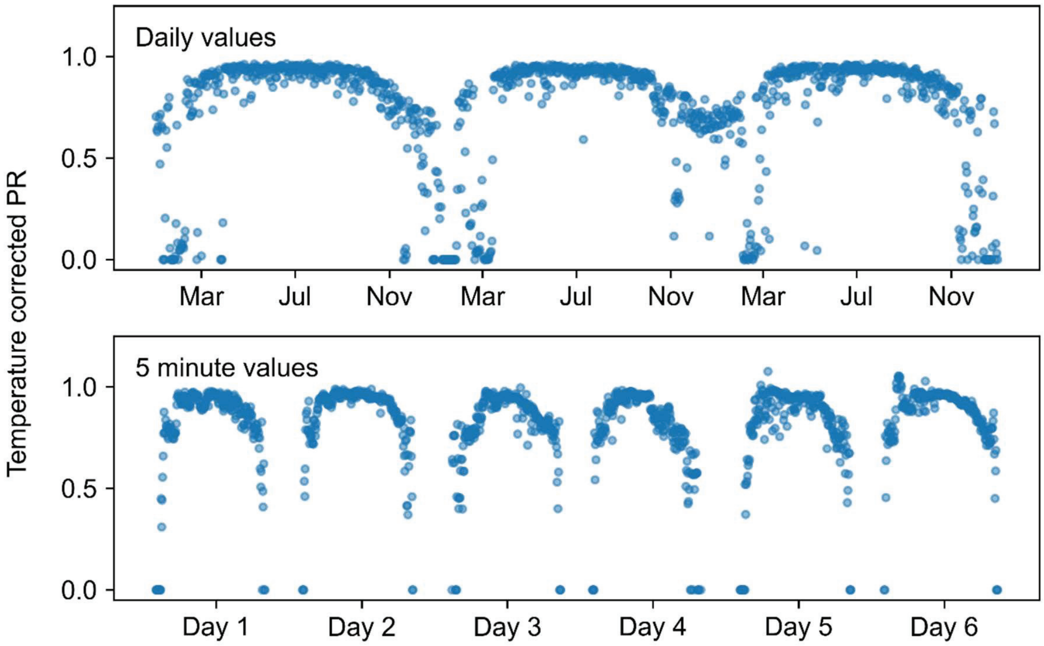


Figure 13: PR'$_T$ with daily and 5 minute time resolution for one of the studied systems, illustrating possible variation in the performance metric through the year and through the day.

Data quality

Poor data quality is a potential weak link in any data analysis. To avoid this, well-performing monitoring system should be implemented from the start. This necessitates both the sensors, installation, and maintenance to be of a certain quality. The list of requirements to achieve a high-quality monitoring system enabling thorough PV performance analysis is long (including various high-quality sensors and continuous maintenance), as for example described in the IEC standard for PV monitoring [36]. Because of the additional costs and that the exact cost of high-quality data is not known, the quality of data in existing PV installation is varying. If the data quality is not ensured from the beginning and additionally properly maintained, it must be considered in the analysis. Consequently, methods for recognizing and handling quality issues in the data is typically necessary.

Common data quality issues in PV monitoring data, is described in for example [36,89,95,96]. Potential issues in PV monitoring data include lack of sensor maintenance, non-functioning logging, communication or sensors, or conditions giving biases or noise in the sensor data. Quality issues connected to lack of maintenance of sensors are for example lacking sensor calibration and irradiance sensors that are not cleaned. Non-functional sensors or logging can result in missing data, unphysical values, duplicated data or stale values. Errors or missing information related to time stamps is another potential issue. This includes lacking information on time zone, daylight saving time or method for time sampling (for example description of what the time stamp represents: the beginning, end or the middle of the logging interval). Inaccurate time synchronization is another potential time-related issue. Irradiance sensors have expected uncertainties linked to for example time lag, drift, and angular response, as well as temperature related zero offsets that can give positive readings at nighttime [97,98]. Additionally, they can be installed inadequately, for example with inaccurate tilt angle or in a location with shading [96]. For module temperature sensors attached to the rear side of the module, it has been observed that they can have poor thermal contact and detach from the module temperature surface [79].

If data quality issues are not detected and handled, they can have large impact on the result of the system performance quantification and loss identification. As discussed in [99], data quality issues giving errors in the irradiance and module temperature measurement can introduce severe errors in the performance evaluation. For estimations of the overall system performance the irradiance measurement is often essential, and errors in the irradiance value would consequently give erroneous results. In analysis of how the performance evolve with time, shifts or drifts in the irradiance measurements can be misinterpreted as faults or degradation in the PV system, or mask degradation in the system. Detached temperature sensors will lead to underestimation of the module temperature which again may lead to underestimation of temperature losses and give an over-optimistic estimation of expected output. Periods of missing data can reduce the robustness of the analysis, as the basis for evaluation is reduced.

Non-ideal references

Another issue that can challenge accurate performance evaluations, is that the *reference* used in the performance metric is not *representative* for the evaluated system. This can be an issue related to *reference measurement* or to the *reference model.* The suitability of the reference measurement is to a certain degree related to data quality. The sensors in the monitoring system must be designed to correctly give a representative estimate of the conditions of the module, both with respect to sensor suitability and sensor placement/installation. This is relevant for sensors measuring both system and location specific parameters, but as system specific parameters are the most common measurements used in our research, these are the parameters discussed in this section. Another potential challenge with comparing the output to a reference, is how well the reference follows the target output under *fast changes* in operating conditions, for example under conditions with fast moving clouds.

Different sensors will give different information, impacting how *suitable* a sensor is for a specific task. As presented in Section 2.2, the plane of array irradiance in the systems evaluated in this work is either measured by a pyranometer or a reference cell. While the pyranometer is measuring the total irradiance in the plane of array, the reference cell is designed to measure the *effective* irradiance, i.e. the irradiance the solar cells can utilize [97,100]. The aim of the reference cell is thus to measure the irradiance after reflection and spectral losses, and with the same time response as a solar cell. A pyranometer, on the other hand, has nearly uniform sensitivity for all the wave lengths within its measurement range and the angular response is almost constant. Also, its measurable response to irradiance is a few seconds slower than a solar cell. Irradiance measured by a reference cell could thus be a better reference device for cases where we want to consider reflection and spectral losses accurately in the performance evaluation. A prerequisite here is, however, that the reference cell is matched with the relevant PV technology. While the spectral responses of crystalline silicon cells are expected to be quite similar, there could potentially be variations in the quality of the antireflective properties of the cell, glass or both. Multiple methods for developing anti-reflecting coatings for PV modules exists, as described in the literature [101]. With differences in reference cell technology and the installed PV modules, unknown deviations between target and reference can occur (affecting the suitability of the reference cell as a reference measurement).

Cell/module temperature measurements can also be performed in different ways. The cell temperature for the systems evaluated in this work is either measured by sensors on the rear side of a module or integrated in a reference cell. The rear side sensor is not measuring the actual cell temperature, but it is measuring the actual operating conditions of the module. To get an accurate value of the expected cell temperature, the expected difference between these two temperatures should be modeled [100]. In the reference cell the cell temperature is directly measured, but the cell has significant deviations in design compared to the PV modules, and the operating temperature could therefore be different.

Issues related to the placement or installation of the sensor are relevant for both in-plane irradiance sensors and module/cell temperature sensors. Section 4.2.1 presented the issue of deviations in tilt between the irradiance sensor and the PV array, and variations in exact tilt within the PV array. A similar issue is relevant for the module/cell temperature measurement. Temperature variations within the system are expected, and the temperature sensor should thus be placed in a location that gives a good indication of what the system overall is experiencing.

If the non-representative sensor measurements are used as input to a model, this will consequently also impact how representative the model will be for the PV system output. Models can also be less suitable as a reference if the effect of influential parameters and expected losses are not correctly considered, as further discussed in the end of this section. Another potential issue with a reference model or reference measurement in performance evaluations, is the difference in temporal response between the reference and the target. This can be crucial for situations with rapid changes in irradiance under cloudy conditions. Moving clouds are a challenge due to potential inhomogeneous shading effects, the time lag of the pyranometer, and the efficiency with which the MPPTs responds to the varying conditions.

Sensor measurements that are not adequate references of the irradiance or temperature of the PV modules, can lead to both offset, systematic trends and noise in the performance metric time series. Fast changes in irradiance, are for example expected to give noise in the performance metric. One severe effect of using non-representative references in performance metrics is that the performance metrics indicate losses or gains in the system that are not real, as previously illustrated in Figure 12. Slight differences in tilt between the array and the irradiance sensor or the reference array, are expected to be one of the causes of seasonal variations in performance metrics, illustrated by the relative irradiation differences between different planes in Figure 11. In general, with a reference model that is not accurately considering all expected losses, the basis to evaluate if the system is under or over performing is thin.

Unconsidered or inaccurately quantified losses

The co-existence of several different loss mechanism is a major challenge when seeking to identify and quantify specific loss mechanisms. To estimate the effect of one loss mechanism, accurate quantitative knowledge of other losses is required. For example, to evaluate the effect of snow, it is necessary to consider other common wintertime losses such as losses caused by irradiance with low intensity and high incident angles or shading. Accurate consideration of other loss mechanisms is particularly relevant for identification of small losses, for example caused by soiling, degradation or module faults. The authors of [102,103] find that improved estimation of soiling and degradation losses can be achieved if the effect of both is considered in the analysis. Typical challenges are loss mechanisms that are difficult to accurately quantify, such as losses related to low intensity irradiance, high angle of incidence, cloudy conditions and temperature. Accurate estimation of reflection loss can for example be challenging when the angle of incidence is above 60°, as the resulting loss is increasingly sensitive to changes in AOI above this limit [60]. Accurate calculations of these types of losses require more detailed input information than what is typically available. This can include information about exact tilt angles, overview of the non-uniformity in irradiance and temperature, and exact information on how all the components respond to the different conditions. Inaccurate estimation of these losses yield uncertainty that can lead to noise in the data and hide the signals related to recoverable losses such as soiling and module faults. Methods for accurate modeling of the expected losses are not only necessary for performance evaluation of historical data. Estimations of these types of losses are also required to predict performance of future systems.

4.3 Potential solutions for improved performance evaluation

When a comprehensive understanding of how various effects impact performance evaluation is established, effective solutions to handle these effects can be discussed. This section presents methods for improving performance evaluation through handling the impact factors discussed in Section 4.2.

4.3.1 Previous work

Various strategies to handle different effects in the data are suggested in the literature. Filtering is perhaps the most common, and also the simplest, solution suggested. Another solution, typically used to handle systematic trends in the performance metrics, is more detailed modeling and trend correction. A third approach is to utilize data signatures or performance metrics to evaluate the data that are less impacted by the discussed impact factors.

Filtering

The principle of filtering is to remove the periods where the performance metric fails to represent the quantity it is determined to measure. This could for example be periods with erroneous input data, noise in the performance metrics, or temporary losses that are not relevant for the effect under study. Filtering could thus be used to handle factors related data quality, non-ideal references, and losses that are challenging to properly quantify. Data cleaning with filtering is typically an important step in suggested procedures for degradation and performance loss analysis [27,88]. The first step is often removal of data quality issues such as non-physical values, invalid readings and periods with missing or duplicated data [36,89,96]. For analyses where the aim is to identify small losses caused by for example soiling, degradation and faults, it is also often suggested to remove periods with-irradiance with low intensity or high angle of incidence, or cloudy conditions [27,83,99,104,105]. In these conditions the output of the PV array can be difficult to predict, because the related expected losses can be difficult to quantify accurately. Which other effects that are filtered out, typically depend on the aim of the analysis. In degradation analysis for example, shading, clipping, curtailment, downtime, and outliers in the calculated performance metrics are often also removed [83,99,105].

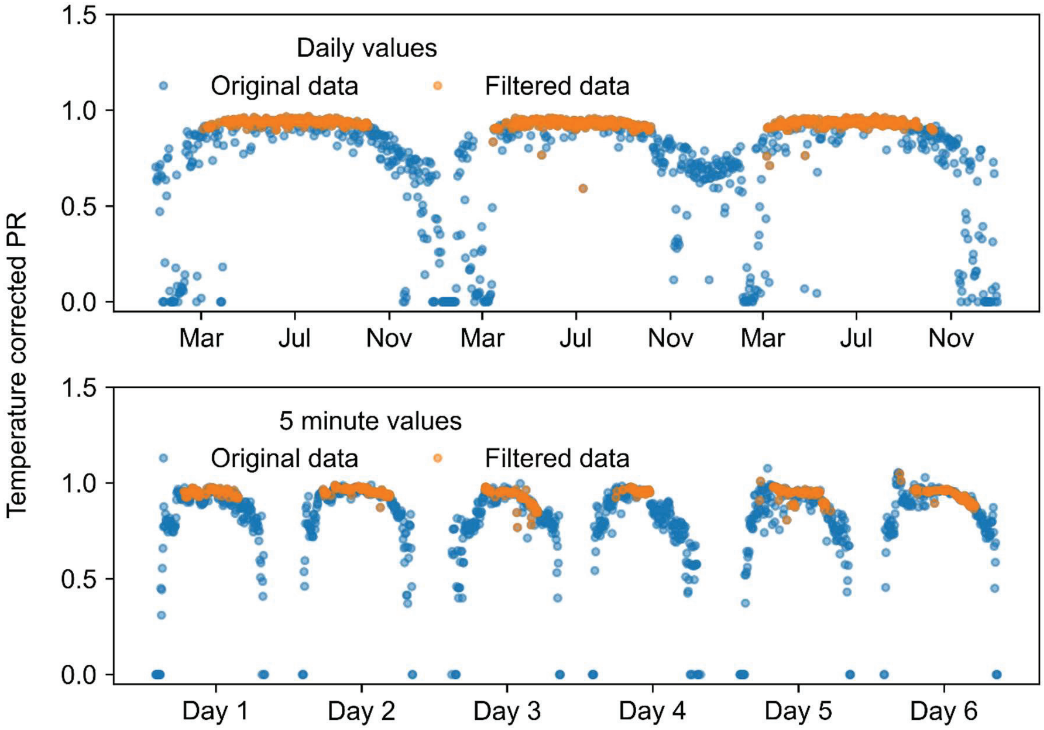


Figure 14: Illustration of the effect on the output of the performance metric calculation of filtering out periods with low irradiance and periods with large changes in irradiance caused by partly cloudy weather.

Figure 14 shows the effect of filtering out periods with low irradiance and large variation in irradiance caused by partly cloudy weather on the output of the performance ratio calculation for one of the studies systems. We observe less variation in the output value, which would make it easier to identify changes caused by component degradation or faults. However, for this example based on data from a Norwegian PV system, we now lack data for large parts of the year.

However, as discussed in [27] the implementation of filtering can be influenced by the preferences of the analyst, and the role and the effect of filtering are not always discussed when implemented in performance evaluation. It is not uncommon that one filter is used to remove multiple effects. For example, a low irradiance filter can be used to remove the effect of low irradiance, high angle of incidence and shading. If not carefully considered, filtering can easily remove more data than necessary. If the aim is to evaluate the overall performance of the system, it is critical that it is only issues related to data quality and non-ideal references that are removed, and not periods with central performance losses. A common challenge introduced with filtering, is when long time periods are removed, giving holes in the time series. Imputation of data has been considered to solve this problem [96]. While imputation of modeled sensor data is considered as a valid substitution, imputation of modeled power data will mostly be relevant for cases where the estimation of how much energy that is generated is central. For evaluation of PV performance, the measured PV array output will be essential, and imputation of modeled power data should therefore be carefully considered.

To improve the filtering process, efforts to identify and filter specific effects are needed. For clipping, for example, a recent contribution in this regard was made by [106]. More automated methods to find the optimal filtering thresholds to reduce the uncertainty in the performance estimation, such as suggested by [88] for degradation rate estimation, is another pathway for improving the filtering process.

Modeling and correction of losses and trends

Another method to handle losses that are not properly quantified in the performance metric is to develop improved models to consider the relevant loss mechanisms. This could be either through physical models, or through empirical approaches, such as machine learning. The use of statistical methods and machine learning to achieve improved PV modeling is gaining more attention. Statistical methods and machine learning can also be used to correct for systematic trends in data, irrelevant of if they are caused by losses that are not properly quantified or non-ideal references giving for example seasonal trends in the data.

Ref. [105] reduces the uncertainty in an estimation of PLR by including physical modeling of the effects caused by variations in the solar spectrum, a mechanism that rarely is included in performance analysis, in the PV output model. In our related work presented in [76], the aim is to compare the performance of two different PV technologies in Norwegian operating conditions. Filtering out low light conditions is thus not an option. To enable comparison at given conditions, the performance metrics is corrected for the losses caused by irradiance with low intensity and high angle of incidence. The analysis shows, however, that while the correction for irrelevant losses improves the basis of comparison, it is challenging to obtain sufficient accuracy.

Using statistical methods, such as seasonal trend decomposition [107], to quantify and correct for systematic trends is often suggested to handle the seasonal trend often observed in performance metric time series. This has for example been suggested for development of fault detection algorithms [84], degradation rate [108] and PLR estimation analyses [27,109]. Figure 15 shows the effect on the output of the performance metric calculation of correcting for the seasonal trend. The seasonal trend is quantified using seasonal trend decomposition [107]. We observe that the systematic trend is reduced, giving a nearly constant performance metric through most of the year. The large losses in the wintertime that occurs every year, but in a more random manner (assumed to be caused by snow), are not successfully corrected.

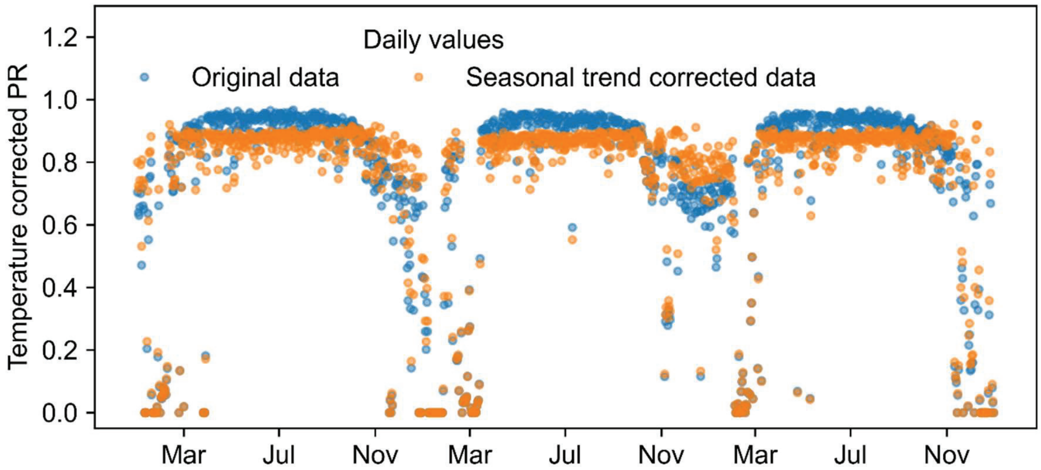


Figure 15: Illustration of the effect on the output of the performance metric calculation of correction for the seasonal trend. The seasonal trend is quantified using seasonal trend decomposition.

Tailored use of performance metrics and data signatures

Another way to handle the discussed impact factors, is to utilize performance metrics or data signatures that avoid the effect of these factors. A method suggested to handle systematic seasonal trends in data, not sorting under neither filtering nor modeling/correction is the year-on year method used to estimate degradation trends [83]. This method is implemented in the open-source library RdTools that is commonly used in degradation analysis. In this method, calculated performance metric values from the same day of the year are compared to each other, reducing the effect of potential systematic seasonal trends. In the suggested approach, extensive use of filtering is additionally applied. This gives a more stable comparison because it compares days under the same condition, but a weakness is that the filtering can reduce the number of comparable days from year to year. As shown in our related work [75] where the aim is degradation rate estimation in Norwegian conditions, it is quite common that whole days are filtered out, challenging comparison of one day of the year to the same day the next year. To improve on this, improved versions of the year-on-year method are tested and developed in this paper. The improved methods enable estimation of the rate of change not only from one year to the next, but by considering the rate of change in the whole analysis period. In [83], it is also suggested to use modeled clear sky irradiance as a reference in the performance metrics instead of measured irradiance, and only evaluate clear sky periods. This removes the challenge of sensor data quality but requires long periods with clear sky weather. Another way to avoid the impact of poor irradiance sensor data quality, is to use yield comparison of similar/identical PV arrays, as we do in [84,90].

A commonly used method to reduce noise in the calculated performance metric, is to aggregate the data and reduce the temporal resolution of the performance metric [27,84]. If the performance metric is only impacted by random noise or if the noise is caused by effects giving large relative differences between the output and the reference in the performance metrics but small absolute differences, this can be an efficient solution. However, if there are systematic trends in the performance metric time series, aggregation could conceal these effects.

4.3.2 Thesis contribution

In the work presented in this thesis, methods for evaluating sensor data quality and the use of filtering to handle different performance evaluation impact factors are tested. Additionally, it is evaluated how different performance metrics are impacted by the discussed factors. In Paper I, we show that comparison with clear sky irradiance modeling can be used to identify drift and shift in the irradiance sensor measurements, in addition to misalignment of the sensor. This discussion is partly continued in Paper II and III, where we find that the statistical clear sky fitting algorithm proposed by [92] can be used to find deviations between the tilt angle of the irradiance sensor and the effective tilt angle of the PV arrays. In Paper I we also assess if the relationship between irradiance, ambient temperature and module temperature can be used to identify detachment of the module temperature sensor.

Paper II and III evaluate the use of filtering to improve the sensitivity in fault detection. Paper II shows that filtering with low irradiance thresholds commonly used in the literature and clear sky filtering do not solve the challenges we have with noise in the performance metric time series for the studied high latitude, cold climate datasets. In paper II and III it is found that filtering thresholds directly aimed at the origin of the noise and systematic trends in the performance metric time series give more stable performance metrics that are more suitable to use in fault detection. This shows that the optimal filtering thresholds depend on operating conditions. This further supports the need for development of filtering threshold optimization methods and other methods to select filters tailored to the specific effects and the purpose of the analysis.

In Paper III we also evaluate how different commonly used performance metrics perform in a fault detection analysis and handle the impact factors discussed in Section 4.2. The tested performance metrics are array Y_f comparison, $PR'_{25°C}$ and PPI based on both physical and machine learning modeling. The results show that choosing the right performance metric based on the quality and condition of the available input data can also be a strategy to handle the impact factors, as suggested in the end of Section 4.3.1. For example, for situations where the irradiance sensor is not a representative reference or has quality issues, yield comparisons between the individual units of the system can be efficient. Machine learning can improve modeling of the expected output of the system, especially for cases where all the parameters needed to physically model all losses in the system are not known. In Paper III, we see that with machine learning the estimation of the expected output in situations with shading and with deviations in tilt between the irradiance sensor and the PV array is improved compared to using a physical model.

4.3.3 Discussion

The main responses in the calculated performance metrics caused by the factors discussed in Section 4.2 are erroneous data, noise and systematic trends or offsets. The two preceding sections discuss how these effects can be handled to improve and develop standardized methodology for performance evaluation. In general, filtering appears to be an efficient method to handle errors and noise in the data, i.e. effects that is difficult to model. Correction based on physical models or statistical methods could be used to handle systematic trends or offsets. As also discussed, filtering and correction can be bypassed if performance metrics or data signatures that are not impacted by the relevant impact factors are utilized. In addition to handling the factors impact performance metric calculation, improved identification of different loss mechanisms and detailed description of the signatures in the data for the specific loss mechanism would also be necessary for improved performance evaluation methodology.

There are two main pathways to develop strategies to handle the factors impacting the performance evaluation. Either the methods could be target at directly handling the different factors, or the aim of the methods could be to handle the resulting signals in the data. In Paper II and III we suggest handling the impact factors directly by using filtering that is specifically targeted for the different impact factors. For example, the conditions where irradiance with low intensity and high AOI give noise in the data are specifically identified and removed. Development of analysis procedures based on this strategy requires methods for accurate identification of the various impact factors. For the other pathway, where the goal to a larger degree is to handle the signal in the performance metric, the methods could for example be targeted at reducing noise and systematic trends in general. One example is optimization methods for finding the filtering thresholds for different parameters (for example intensity and angel of incidence of the irradiance, or clear sky index) that gives minimal noise in the performance metrics. Seasonal trend decomposition to quantify and correct for the effects giving seasonal trends in data or machine learning modeling (as discussed in Paper III) are other potential methods for handling the more systematic signals in the performance metrics. With seasonal decomposition all the effects introducing a seasonal trend in the performance metric is corrected for simultaneously, the same way as all the losses in the system are estimated as one effect with a machine learning model.

For both pathways, in-depth knowledge of the impact factors in the dataset, methods accurately targeting these factors, as well as a clear aim for the analysis are required. For example, we do not want to correct for or remove losses we want to quantify. When using methods for selecting filters and filtering thresholds based on optimization methods, it will be critical that the optimization is done on the parameters that actually introduce noise in the data. Machine learning modeling requires proper data cleaning. The automatic methods can easily fail if there are issues giving stronger signals in the data than the ones we try to correct for. For example, to use seasonal decomposition if there are stronger signals in the data than the seasonal signal, can give a correction that introduce new deviations and biases in the data. Solving all the issues with targeted correction and filtering for different effects, on the other hand, can give a very complex and time-consuming analysis.

However, we find that focusing on how to specifically handle the impact factors in the data is in general a promising strategy for improved performance evaluation. In our related work presented in [84] a data processing procedure for fault detection in larger PV plants is suggested. This procedure is based on using array yield comparison as a performance metric to avoid irradiance sensor quality issues. Effects giving noise and erroneous data are filtered out, and seasonal correction of the performance metric time series is utilized. In [85] it is found that with this performance metric calculation and processing, loss as small as the power loss caused by activation of bypass diodes in PV modules can be detected from the output data of a larger PV array.

5 Snow losses in PV systems

The assessment of factors impacting PV performance evaluation in high latitude cold climate conditions presented in Chapter 4, shows that snow is an important factor to consider because of its potential large effects on PV performance. This chapter presents the challenges snow introduce to in PV performance evaluations, and potential solutions. In addition to this, the chapter discusses how snow should be considered in PV performance *predictions*. Methods for *identification* and *prediction* of snow losses are essential in order to handle snow in performance evaluations and prediction. Section 5.1 describes the impact of snow on performance evaluation and predictions, while Sections 5.2. and 5.3. summarize previous work and the contributions of this thesis on snow loss identification and prediction, respectively.

5.1 The impact of snow on PV performance and performance analysis

5.1.1 The prevalence of snow losses

Snow can cause severe and long-lasting shading of PV modules that result in large power losses. Daily and monthly energy losses due to snow of up to 100 % and annual losses of more than 30 % have been reported [45]. There are, however, large variations in the reported snow loss values. The energy loss correlates with the lost irradiance, which again depends on how often and for how long the PV modules are covered by snow. Excluding areas with permanent snow cover (such as Antarctica and Greenland) and mountainous areas, snow is typically found in the northern hemisphere [110]. As illustrated in Figure 16, the lower latitude limit where snowfalls are expected will vary with continent. With increasing latitude, the number of days with ground snow covers generally increases. From the figure, it is observed that snowfalls are expected in significant PV markets [15] such as China, Japan, EU and USA. It is, however, anticipated that global warming will cause less snow in the future. In Norway, shorter snow seasons, reduction in snow depth and increase in snowline elevation are expected [111]. Figure 17 shows the historical monthly mean snow water equivalent (SWE) for a location in Eastern Norway compared to modeled SWE for the period 2030-2050 based on two different emission scenarios. SWE is a measure of the amount of liquid water stored in a snowpack. The reductions in SWE for the modeled values illustrate the expected reduction in snow depth and length of snow season.

5.1.2 Applications of snow loss estimation and identification

Because of the large potential losses, snow is an essential parameter in performance evaluation and prediction. Often the loss is *accepted*, but as active and passive snow clearing are possible, the loss can also be *correctable* or *avoidable*. Snow loss could for example be reduced through system design or O&M. Snow loss *estimations* are therefore useful in both predictions of system output, optimization of system design, and for efficiently implementing corrective measures in O&M. Additionally, *identification* of snow losses is essential in performance analysis where separation between different loss mechanisms is necessary.

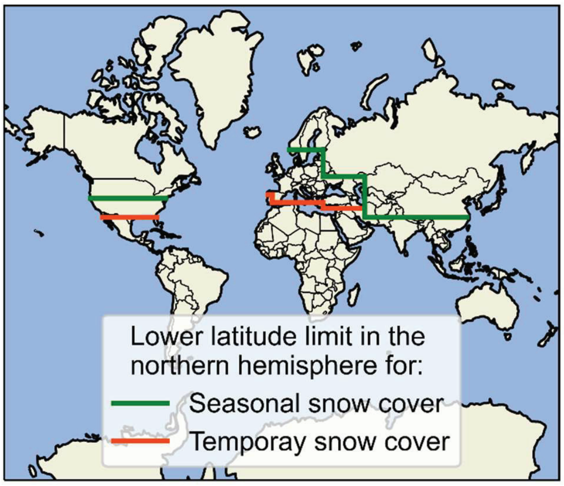


Figure 16: The lower latitude limit where temporary snow covers (lasting a few days) and seasonal snow covers (lasting for months) is found, according to [110].

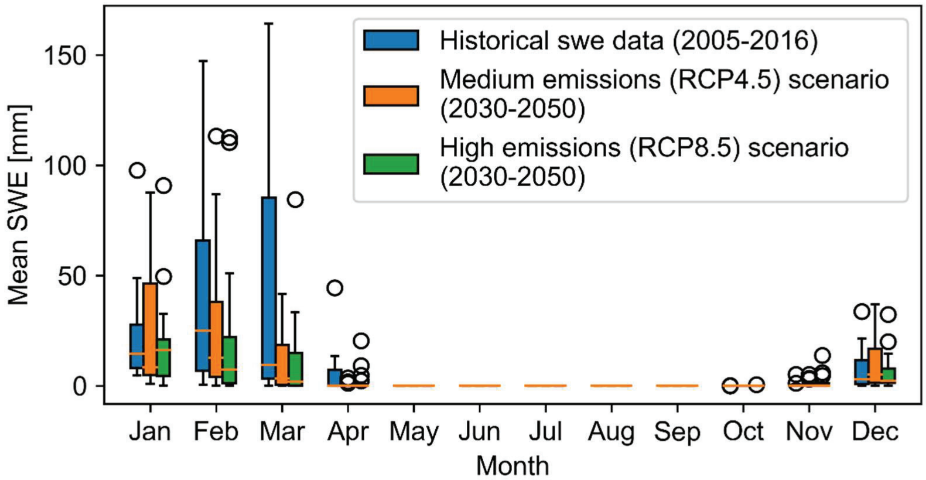


Figure 17: Monthly mean snow water equivalent for a location in Eastern Norway for 1) historical data, 2) modeled data for 2030-2050 given the RCP4.5 emission scenario, and 3) modeled data for 2030-2050 given the RCP8.5 emission scenario [112,113]. Data from [114].

Snow loss estimation

As discussed in Section 3.1.2, estimations of expected PV system losses are necessary for predictions of PV energy generation, which again is necessary in multiple different applications. With its potential large impact, snow losses are no exception to this. Because energy losses in PV systems related to snow are expected to vary with both snow and weather conditions, as well as system design [45], historical data from one system will not necessarily be transferable to other systems. A snow loss model is therefore preferred. Additionally, a model enables inclusion of the potential effect of reduction in snow caused by global warming on PV snow losses.

Estimations of snow loss is also important when evaluating the potential for reducing the losses caused by snow. Snow losses can be reduced by designing the system to induce increased passive clearing, for example by increased tilt, use of frameless [115] or bifacial modules [116], or ice phobic surface coating [117]. Alternatively, or additionally, snow losses can be reduced by active snow clearing as a part of O&M of the system, for example by manual shoveling or heating [118]. The estimation of potential snow losses is important input to the evaluation of the cost of *not* clearing and the *gain* of actively clearing the snow. Estimation of natural snow cover development is particularly important. When would the snow have been cleared off the modules if no active measures were implemented? Is it snowing soon again, making the active snow clearing useless? If the snow is not cleared now, will it freeze to the modules and persist for a long time? The cost of not clearing the snow is, however, not only related to lost energy generation, but also the potential damage of the system and/or the roof caused by the mechanical load.

Identification of snow losses

Identification of snow losses is essential in analysis of loss mechanisms. Because of the large and non-systematic losses, snow could for example have a large impact on the output of a degradation analysis or in a monitoring system designed for fault detection. If snow is not separated from failures in a monitoring system, alarms indicating component failure can be issued when the actual issue is temporary snow. This could give reduced trust in the monitoring system from the operators. A consequence of this could be that serious system issues are not tended to because they are assumed to be caused by snow. Identification of snow losses in PV monitoring data can also provide more knowledge with respect to snow cover and resulting losses based on historical data. Such data could be important input to development and validation of snow loss models.

Method development challenges

From the discussion in the previous sections, it can be concluded that it is essential with reliable methods for 1) identification of snow in PV monitoring data, and 2) predictions of snow losses for a given system under given weather conditions. However, neither identifying snow in PV monitoring data nor predicting the presence of snow cover and resulting loss are trivial. As we discuss in Paper IV, because of the potential varying coverage and transmittance of snow, snow can shade PV modules in multiple different ways, which consequently can give multiple different snow shading responses in the measured output. Which response a given snow cover leads to in measured parameters in a PV system can also depend on the module orientation/bypass diodes and the array configuration, i.e. how many modules that are series and parallel connected to the same MPPT. Figure 18 shows how a partial snow cover on the lower part of the module, a typical situation if the snow slides down the modules, will shade relative to the module substrings and the bypass diodes. The challenge of estimating the resulting loss in PV parameters due to a given snow cover will complicate both the identification and prediction of snow losses. Additionally, as we describe in Paper VI, the process of natural accumulation and clearing of snow on PV systems depend on a large range of weather, snow, and installation parameters, and are consequently also difficult to predict. The impact of multiple different parameters on the snow clearing can give complex snow shedding patterns, as shown in Figure 19. An additional challenge for development of methods based on monitoring data, is that snow also can impact the sensor values. For example, the irradiance sensor can also be covered by snow.

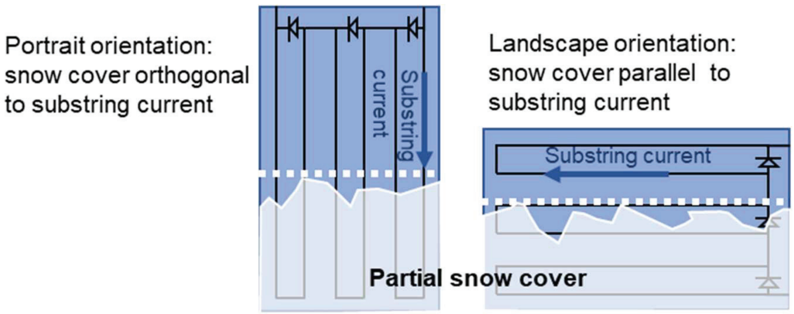


Figure 18: Illustration of how a partial snow cover on the lower part of the module will shade relative to the module substrings and the bypass diodes.

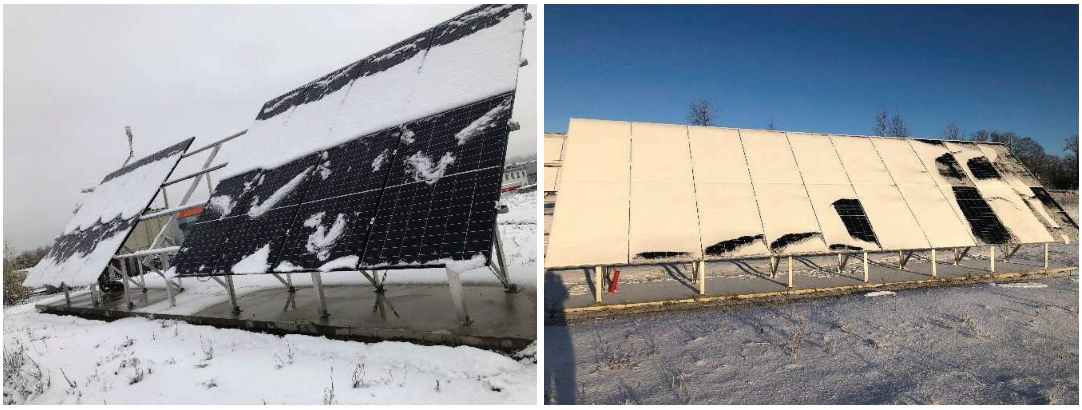


Figure 19: Partial snow shedding at the PV test facilities at IFE, Kjeller, Norway.

5.2 Identification of snow losses

5.2.1 Previous work

Because of the potential large and long-lasting losses, how to identify and handle snow has been discussed in various performance analysis studies. Filtering out outliers [27,88] or removing periods with low performance ratios are potential methods to handle snow in performance evaluation. This will remove most snow events, but also other types of large losses, and will not be accurate enough for detailed snow identification. This type of filter can be used in for example a degradation rate analysis, but will fall short in for example fault detection, as it could not separate between snow losses and severe component failures, one of the most important fault a monitoring system should detect. To improve the snow identification, other parameters have been included, giving more detailed description of the signature in the data caused by snow. Ref. [119] suggests identifying snow losses by detection of power losses in periods with temperatures close to or below freezing. In the snow detection algorithm suggested in [120] for larger power plants, it is added that the power loss should be identified in most of the arrays, and the irradiance and power should in general be low. The development in plant and array values of *PR'* is further tracked to confirm that it follows the characteristic for snow loss recovery (snow loss recovery is expected to be *gradual*). The snow event is closed when the *PR'* returns to normal for most of the arrays, and ambient temperatures above 20°C. The snow detection suggested by [120], do enable more detailed snow identification than just considering power loss and/or ambient temperature. The method does, however, to a certain degree appear to be adapted to the system configurations and weather type of the studied systems. For example, it requires comparable arrays, and the temperature threshold giving most accurate closing of snow events and the characteristics of snow loss recovery could be expected to vary with the weather and snow conditions. To efficiently implement this, a thorough description of the characteristic of snow loss recovery could additionally be useful.

In this work, the focus is on methods to identify snow cover using parameters typically measured in a PV plant. However, methods using dedicated measurements to identify snow also exist. Imaging [121] is a promising method, as this can give a good overview of the total snow coverage, including its non-uniformity. This method requires automatic image processing and analysis, and a model for translating the estimated snow cover to resulting power loss. A potential challenge for this method is semitransparent snow covers. Other methods have also been introduced, such as weight sensors [118]. We have also found that the Kipp & Zonen Dust IQ soiling sensor can be used in snow identification [122]. A challenge with point measurements is, however, the potentially large non-uniformity of snow. A point measurement is not necessarily representative for the snow coverage, especially in the melting period, but it can be a useful measurement to include in snow cover detection.

5.2.2 Thesis contribution

Improved understanding on how snow influences PV monitoring data for both different snow conditions and system design, enable improved separation of snow loss from other types of losses and development of general identification methods with broad applicability. A more comprehensive understanding would give a better basis to decide if results from one system, with its specific system design and snow conditions, can easily be transferred to another system, or when adjustments are needed. The main contribution of the work in this thesis on this topic, is analysis of signatures in PV monitoring data caused by snow. In Paper II and III, we describe how snow can give large and varying losses in PV systems and consequently have a significant impact on fault detection. In Paper VII we quantify the magnitude and the interannual variation in the snow losses for many of the systems in the studied dataset. The effect of snow on the output parameters in the monitoring system, i.e. the *snow signatures* in

PV monitoring data, is described in more detail in Paper I, IV and VI. In addition, we discuss how the snow signature can be used to improve performance analysis.

In Paper I we evaluate how well the relation between measured module temperature and irradiance work as a snow cover identifier. The correlation between irradiance and module temperature is well known, but when the modules are covered by snow, it is expected that the impact of the irradiance on the module temperature will be reduced. When the irradiance is high and the module temperature sensor is measuring the temperature of snow-covered cells, identification of snow based on this relation works quite well, as the module temperature is much lower than what would have been expected at the given irradiance. For situations with lower irradiance, evaluating this relation is not equally efficient, as the absolute difference between the measured temperature and the expected temperature at the given irradiance is small. We also conclude that partial snow shading can be challenging. Because the module temperature sensor gives a point measurement, it can potentially measure the temperature of the part of the module that is not covered.

In Paper IV the assessment of snow signatures in PV monitoring data is extended. We analyze the effect of full and partial snow cover on time-series of module temperature, DC current, voltage and power for a small system with portrait-oriented modules and a large-scale system with landscape-oriented modules. The analysis is supported by simulations of IV curves for snow covered PV modules in both landscape and portrait orientation (see Figure 18 for description on how partial snow cover is expected to be different for the two orientations), where both the transmittance and coverage of the snow cover are varied. The identified possible responses in output data for different types of snow covers are summarized in Table 6. In Paper VI, the most typical snow signatures in an expanded dataset are identified, marked in bold in the table. When the losses in all the parameters are not close to 100 %, the response in the electrical parameters is mostly characterized by large loss in current and typically much lower losses in voltage, both for systems with modules in landscape and portrait orientation. From the large current losses, it seems that the typical snow cover on the systems is impacting most module substrings in the array, and/or that there is a large degree of semitransparent snow covers. Semitransparent and partial snow covers explain the lower voltage losses. It is concluded that the non-uniformity in snow transmittance and coverage, as well as the array configuration, impacts the output of the system in the situations where the snow cover is not full and opaque.

Based on the identified signatures in PV monitoring data caused by snow, improved snow filtering in fault detection is discussed. In the fault detection study presented in Paper III, the noise caused by snow is removed by filtering out all periods with snow on the ground. This removes almost all the data impacted by snow, but it also removes large parts of the data where there are no snow losses, reducing the period where fault detection is possible. With the aim of developing more accurate filters, Paper IV discusses the inclusion of snow signatures, as well as using snow loss modeling to indicate the probability of snow cover instead of measured snow on the ground. The signature included in the snow detection evaluation is voltage loss, as the modules are installed in landscape orientation. This will not separate snow from all types of losses (for example other loss mechanisms that leads to voltage loss), but it will reduce the amount of data removed compared to take away all the data when there is snow on the ground. Together with the snow loss modeling, 97 % of the data with significant snow losses are removed, but there is still potential to reduce the number of false positives. We conclude that using the snow signatures is useful, but adding more signatures is required to develop automated snow detection that is both accurate and enables separation of snow from faults.

Table 6: Overview of the expected response in PV monitoring output parameters for different types of snow cover, as described in Paper IV. The typical response found in the validation study (evaluating electrical parameters) presented in Paper VI is marked with bold font.

	Full, opaque	**Full, semi-transparent**	**Partial, opaque**		**Partial semitransparent**	
			Portrait	Landscape	Portrait	Landscape
Module temperature	<< Normal operating temperature		Typically < Normal operating temperature			
DC current loss	**100 %**	Large[1]	100 %	0/Large[2]	Large	0-**large[3]**
DC voltage loss	**100 %**	0	Large	Large-medium[2]	Negative loss	**0-large[3]**
Power loss	**100 %**	Large[1]	100 %	Large-medium[2]	Large	**Medium-large[3]**
PV plant	No output	All inverters have large loss	Large losses in many/all inverters. There may be large variations in power, current, voltage and module temperatures.			

[1]*Depending on snow transmittance.*
[2]*Depending on snow coverage/number of covered module substrings.*
[3]*Depending on snow transmittance and coverage.*

While the identified snow signatures are not implemented in automated snow detection, we have made use of the to improve the loss identification in the performance analysis of a bifacial system [73] and in estimation of soiling losses on a farm [122]. In both cases the snow signatures are useful to identify and quantify the snow losses for the two system types, and to separate the snow losses from losses caused by shading or irradiance with low intensity and high angle of incidence. Because the bifacial solar cells also can utilize the rear side irradiance, and typically will not experience snow shading on the rear side, the snow signatures of bifacial systems are expected to be different than for monofacial systems. Still, the identified data signatures from the studied monofacial systems appear to be useful in identification of snow losses.

In the work presented in the preceding sections, the aim has been to describe what snow looks like in PV monitoring data, and how this knowledge can be used. However, the studied datasets may not include all potential variations, and the description is not expected to be exhaustive with respect to the dependency on system design and weather/snow conditions. Other measurements and data signatures, for example how the different parameters develop over time, can also be relevant. More detailed current and voltage data, for example IV curves, is expected to provide more information on the impact of snow covers on PV modules/arrays. More detailed data on the snow coverage (for example snow cover images and transmittance measurements/estimates) and data spanning a greater set of different snow conditions could also provide additional information.

5.3 Prediction of snow losses

5.3.1 Previous work

Multiple models for predicting snow cover and/or snow loss based on weather and system parameters have been described in the literature [45]. Most of the suggested models, according to the overview of PV snow loss models given in [45], are based on making empirical correlations directly between influential parameters and snow loss. In the model suggested by Marion et al. [32], the modeling is to a larger extent based on physical processes. In the Marion model empirical correlations and physical considerations are used to 1) estimate when snow accumulates on the PV modules, and 2) when and 3) how fast the snow clears of the modules. This is used to estimate the snow coverage, and from the number of shaded module substrings, the power loss is calculated.

A typical limitation of the suggested models is that they are developed using data from one system, and often the time series used are short – about one or two winters. Additionally, few validations of the models are published [123]. Exceptions do however exist, the Marion model [32] is for example developed based on data from multiple systems and is also evaluated in other studies [123,124]. In general, the lack of validation on a larger dataset can mean that the models are biased because they will be strongly connected to how snow accumulates/clears off the studied system and the specific types of snow conditions represented in the dataset. This means that the models will not necessarily be transferable to other systems and other snow conditions. While developing a model that can handle *all* system configurations and *all* snow conditions indeed is very ambitious, a model should be developed for broad applicability. Building general models using a fully empirical approach, requires large amounts of data and good measurements/estimates of all the potential influencing factors, covering the different potential combination of system design and snow conditions. The model development studies do, however, often give valuable information on which processes and parameters that can be important for snow accumulation and clearing.

5.3.2 Thesis contribution

To contribute to improved and generalized snow loss modeling, we have tested existing snow loss models, identified which types of models that seem to be most promising for generalization, identified room for improvements and used this to suggest improved models. We have also studied how snow loss modeling should be used in yield estimations for future systems.

In Paper IV we test four of the snow loss models [32,125–127] proposed in the literature. The evaluated models are chosen because they are of varying complexity and utilize commonly available input parameters. We find the Marion model [32] to be the most promising. With just some adjustments to the empirical coefficients used in the model, we achieve a satisfactory fit with the snow power losses estimated for the studied system. A potential explanation for why this model may be more suitable for generalization is that it aims to predict the different processes in snow accumulation and snow clearing separately, as described in the previous section. Additionally, the separation in the model of different processes makes modifications easier. Models based directly on empirical correlations are to a larger extent describing the conditions in the test dataset, and if these conditions are not general, the model will also not be general.

In Paper IV we find that the Marion snow loss model could be improved by using empirical snow clearing rate coefficients estimated for the specific system design, to consider how the system design can impact snow accumulation and clearing. For the system evaluated in the paper we also find that different snow clearing rates should be used for thin and thick snow covers, to better include how snow clearing and accumulation can vary for different snow conditions. The results are validated for multiple roof mounted systems (smaller residential systems and large-scale commercial systems) in Paper V and VI. Figure 20 shows for one snow season how the snow coverage modeled with the suggested improvements to the Marion model follows the loss estimated from the monitoring data for the system used to test the model in Paper IV. The snow coverage is not expected to be directly correlated to the energy snow loss, as the activation of bypass diodes also will have an impact and the snow coverage is not expected to be exactly the same for all modules. The snow coverage is still included here instead of the modeled loss (which is based on how many bypass diodes that are expected to be covered, and therefore is either 0, 33 %, 66 % or 100 %) to give more detailed description of the prediction. Figure 21 shows the performance of the model on the extended dataset in Paper VI for monthly and annual losses. There are two types of system configurations in the dataset: residential systems on tilted roofs with portrait-oriented modules, and commercial systems on flat roofs on with landscape-oriented modules. Within the same type, the same set of snow clearing coefficients are used to the test the model applicability. It is found that separate snow clearing coefficients for thin and thick snow covers better include the effect of varying snow condition, giving good results for the same type of systems independent of climate zone. The large modeling errors for some of the systems in Figure 21, is either related to small absolute losses, or that the systems had additional parameters impacting snow clearing, such as heat leakage from the building or shading. This complicates snow loss modeling. The PV installations had the same design, but the overall systems were not similar in all the parameters that impact snow clearing.

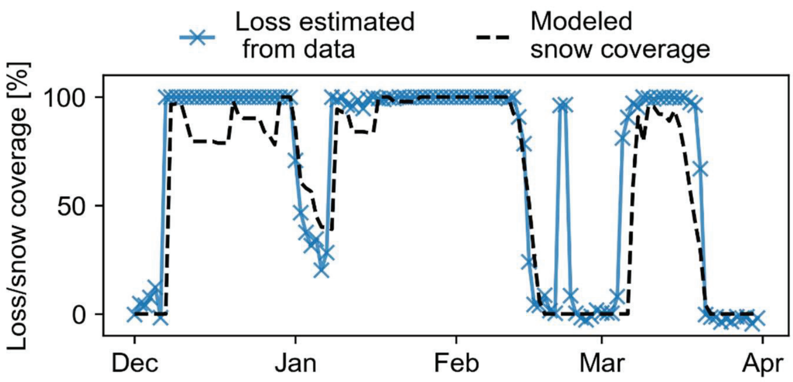


Figure 20: Daily values of energy loss estimated from data and modeled snow coverage for one snow season. The snow coverage is modeled with the suggested improvements to the Marion model.

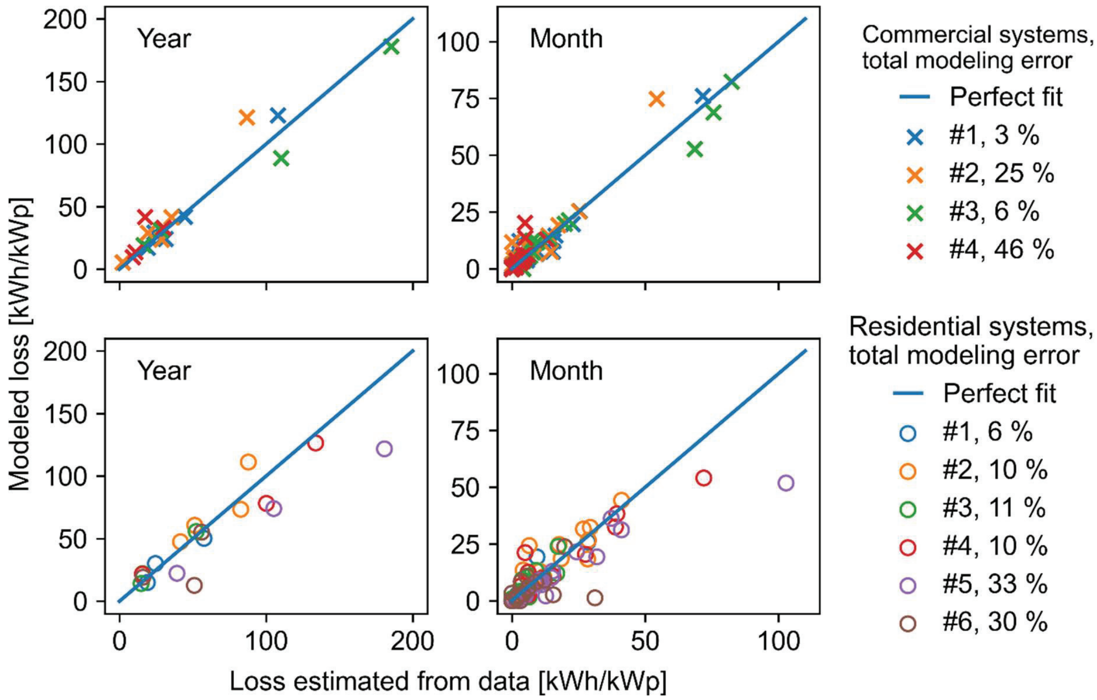


Figure 21: Annual and monthly snow loss modeled with the proposed improvements in the Marion model compared to losses estimated from the output data.

In Paper VII we further discuss how snow loss modeling could be applied in yield modeling of future systems, as there are many choices here that need evaluation. Which input data that should be used is a central concern, as a typical meteorological year (TMY) that is commonly used in PV yield estimation is not necessarily a typical snow year nor give a good estimate with respect to year-on-year variations. We describe a procedure using long time series of weather data and snow loss modeling, which gives both a typical monthly/annual loss value, and a range of typical variation. An example of how this could be implemented is presented in [128], where we use the monthly typical values as input to a PVsyst simulation as part of an energy system dimensioning analysis.

Compared to describing and identifying the effect of snow in PV monitoring data, there are probably even more unanswered questions related to PV snow loss modeling. For example, how accurately can snow loss modeling be done with the weather data that typically is available? Which parameters (including system, snow, and weather parameters) have significant influence on snow cover and resulting losses, and should be specifically included in the model? Which system parameters should be similar to use snow clearing rate coefficients extracted from one system to another system? The work in this thesis only focuses on data from roof mounted systems. But what about ground mounted systems, where the rear side is open, and the module temperature might be influenced by for example wind? There are indications that ground mounted systems shed snow faster than roof mounted systems [32]. There is additionally limited research on how snow will impact the yield of systems with trackers or bifacial modules, both technologies that are increasingly common. It is assumed that systems with tracking (especially if the tracking algorithm is adapted with the aim of active snow clearing) [123,126,129] or bifacial [116] will have less snow losses, but more data confirming to which extent this is expected is needed. It is clear that a range of parameters significantly impacts snow loss, and more data is necessary to identify the most critical parameters for different types of system designs and to extend the validations.

6 Conclusions and further work

6.1 Conclusions

In order to obtain accurate predictions of PV energy yield or to identify the optimal system design, *predictions* of the system performance under given operating conditions are essential. Performance *evaluations* of existing systems can give valuable information on which losses that should be included in the performance predictions, and additionally determine the potential for improved performance. Accurate methods to predict and evaluate performance could thus contribute to cost reductions and performance increases for PV installations. The main topic of the work presented in this thesis is methodology for PV performance evaluations. The study includes assessment of methodology for performance evaluation of PV systems for a dataset consisting of installations located in Norway, and classification of factors impacting the evaluation. The identified main categories of the factors impacting the evaluation is: 1) data quality, 2) the use of non-representative references in the performance metric calculation, and 3) inaccurate quantification of expected losses. For example, small variations in tilt between the PV array and irradiance sensor are found to give a seasonal trend in the calculated performance metrics, illustrating the potential effect of non-representative references utilized in performance metrics. For the evaluated dataset, especially snow and the irradiance conditions specific for the Norwegian conditions can result in losses that are difficult to accurately quantify and include in performance evaluation.

Additionally, this thesis describes an assessment of methodologies for improving and standardizing performance analyses by handling these identified impact factors. In the published papers, we test filtering of the effects caused by the impact factors, and how the impact factors affect different performance metrics. We find that the sensitivity in a fault detection analysis is improved when specifically targeting the different effects with filtering. We also find that the choice of performance metric can be used to avoid certain impact factors. For example, yield comparison of identical units can be used for fault detection in a larger PV plant if there are irradiance data quality issues. Machine learning can be used to achieve improved modeling of expected PV array output in systems with shading or variations in array/sensor tilt caused by for example topography variations.

From the assessment of performance evaluation methodologies, we find that improved methods to identify and predict snow losses in PV systems are necessary in both performance evaluations and predictions. To contribute to improved snow identification, we describe the effect of snow on various parameters measured in PV monitoring systems. It is found that snow give specific signatures in DC current and voltage time series. Combined with weather data it is therefore possible to separate snow losses from other types of losses. However, the potential non-uniformity in snow coverage and transmittance, combined with how various system designs can impact the shading response of the system, result in a wide range of potential snow signatures, complicating automatic snow identification. We additionally evaluate existing snow loss models, and suggest an improvement to the commonly used Marion snow loss model. The improved model results in a reduction in modeling error of 23 percentage points for the studied dataset compared to the default implementation of the Marion model.

1. What are the factors impacting performance evaluations of PV systems in high latitude, cold climate locations?

Paper I: Evaluation of sensor data quality issues.

Paper II, III: Identification of root cause factors for deviations in calculated performance metrics.

2. What are possible solutions to handle these impact factors?

Paper I: Suggestion of methods for evaluation of sensor data quality based on analysis of monitoring data.

Paper II, III: Testing og targeted filtering of the effects caused by factors impacting performance evaluation. Evaluation of method for quantification of the effect of tilt variations on PV output/measured irradiance.

Paper III: Evaluation of performance metric choice and machine learning based modeling for improved fault detection.

3. What is the effect of snow in PV systems?

Paper I: Description of the relationship between snow cover and PV module temperature measurements.

Paper II, III: Description of the effect of snow on a fault detection analysis.

Paper IV: Description of how snow cover impacts measured PV monitoring data parameters.

Paper VI: Validation of results in Paper IV.

Paper VII: Estimation of snow losses for a set of PV systems in Norway.

4. Can losses caused by snow be identified and predicted?

Paper IV: Evaluation of existing snow loss models. Suggestion of improvements to the Marion snow loss model.

Paper V, VI: Evaluation of the improved Marion model on extended dataset.

Figure 22: Overview of how the published papers contributes to answering the research questions outlined in the first chapter.

Figure 22 summarizes how the published papers contribute to answering the research questions outlined in the first chapter. The studied dataset used in the evaluation of impact factors is limited to one type of system design and with the same type of monitoring system. Consequently, the classification of factors impacting performance evaluation is not expected to be exhaustive, although the identified impact factors are anticipated to be central also for other types of systems in similar operating conditions. The evaluation of methods to handle impact factors could also be extended. As indicated in Section 4.3, there are more methods that could be evaluated. Thus, the presented work does not aim to present a final solution for improved performance evaluation, but rather to contribute to improved understanding on how improved performance evaluation could be achieved. The work presented in this thesis contributes to improved understanding of the effect of snow in PV systems and proves that snow losses can be identified and predicted. However, more data and in-depth analysis is required to evaluate the broadness of applicability and enable automatic snow identification and accurate snow loss modeling. While this work identifies and validates improvements to a well-established snow loss model, multiple aspects that appears to be influential for snow losses and could further improve snow loss modeling are also identified.

6.2 Further work

The results from this work could be generalized and given wider applicability if a variety of additional datasets are analyzed in the same way. A larger number of datasets could give a more general evaluation of performance analysis impact factors and snow signatures, and a more solid validation of the suggested methodology for predicting snow losses. Additionally, various methods to handle different performance evaluation impact factors should be further assessed and developed. To further improve snow loss predictions, more work should also be done on determining parameters that influence snow cover development and evaluating the importance of the different parameters. Central points of interest for further work are:

– *Extension of classification of performance evaluation impact factors and suggestions and validation of solutions to handle the impact factors.* A wide evaluation of impact factors and solutions to handle them, may help to identify the factors that have the most impact on the analysis for different system designs and operating conditions. Additionally, an extended evaluation may help identify which factors that are typical and which factors that only occur in very specific situations. A broad validation of methods for handling the impact factors, would identify for which system designs and operating conditions the methods are applicable, aiding development of standardized and automated performance analysis.

– *Loss identification in existing systems.* Improved methodology to identify and quantify losses in PV systems enable a broad analysis of system losses and performance loss rate in existing systems, which could be used to improve the prediction of performance for future systems.

– *Use more detailed data to describe the effect of snow on PV.* Detailed data on snow coverage/transmittance and PV module (IV-curves) would give improved and more detailed understanding of the effect of snow on PV modules. This could quantify the variation in the output caused by different types of snow shading and improve both identification and prediction of snow loss.

– *Validate snow signatures and develop snow identification algorithms.* Broad identification of snow signatures in PV monitoring data on multiple different datasets with different monitoring output would enable separation between general, typical snow signatures and typical cases. This could facilitate development of general snow identification algorithms with broad applicability.

– *Evaluation of influential parameters to improve snow loss modeling.* To improve snow loss modeling, an assessment of the different parameters (system design, snow and weather conditions) that impact on snow coverage and the resulting loss is useful. Based on this, the most central parameters could be identified, and directly implemented in the snow loss modeling. Additionally, it is necessary with broad validation of the models for different system designs and different snow conditions.

7 References

[1] Matt McGrath, Climate change: IPCC report is “code red for humanity,” (2021). https://www.bbc.com/news/science-environment-58130705 (accessed March 18, 2022).

[2] IPCC, AR6 WGI Summary for policymakers headline statement, 2021.

[3] United Nations, Secretary-General Calls Latest IPCC Climate Report ‘Code Red for Humanity’, Stressing ‘Irrefutable’ Evidence of Human Influence, (2021). https://www.un.org/press/en/2021/sgsm20847.doc.htm (accessed March 18, 2022).

[4] D.I. Stern, The role of energy in economic growth, Ann N Y Acad Sci. 1219 (2011) 26–51. https://doi.org/10.1111/j.1749-6632.2010.05921.x.

[5] H. Schandl, S. Hatfield-Dodds, T. Wiedmann, A. Geschke, Y. Cai, J. West, D. Newth, T. Baynes, M. Lenzen, A. Owen, Decoupling global environmental pressure and economic growth: scenarios for energy use, materials use and carbon emissions, Journal of Cleaner Production. 132 (2016) 45–56. https://doi.org/10.1016/j.jclepro.2015.06.100.

[6] S. Chu, A. Majumdar, Opportunities and challenges for a sustainable energy future, Nature. 488 (2012) 294–303. https://doi.org/10.1038/nature11475.

[7] R. Golombek, A. Lind, H.K. Ringkjøb, P. Seljom, The role of transmission and energy storage in European decarbonization towards 2050, Energy. 239 (2022). https://doi.org/10.1016/j.energy.2021.122159.

[8] S. Nonhebel, Renewable energy and food supply: Will there be enough land?, Renewable and Sustainable Energy Reviews. 9 (2005) 191–201. https://doi.org/10.1016/j.rser.2004.02.003.

[9] A. Gasparatos, C.N.H. Doll, M. Esteban, A. Ahmed, T.A. Olang, Renewable energy and biodiversity: Implications for transitioning to a Green Economy, Renewable and Sustainable Energy Reviews. 70 (2017) 161–184. https://doi.org/10.1016/j.rser.2016.08.030.

[10] I. Hanski, Habitat loss, the dynamics of biodiversity, and a perspective on conservation, AMBIO. 40 (2011) 248–255.

[11] E. Brondizio, S. Diaz, J. Settele, H.T. Ngo, Global assessment report on biodiversity and ecosystem services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services, IPBES. (2019). https://doi.org/https://doi.org/10.5281/zenodo.3831673.

[12] L. Temper, S. Avila, D. del Bene, J. Gobby, N. Kosoy, P. le Billon, J. Martinez-Alier, P. Perkins, B. Roy, A. Scheidel, M. Walter, Movements shaping climate futures: A systematic mapping of protests against fossil fuel and low-carbon energy projects, Environmental Research Letters. 15 (2020). https://doi.org/10.1088/1748-9326/abc197.

[13] A. Jäger-Waldau, Snapshot of photovoltaics - March 2021, EPJ Photovoltaics. 12 (2021). https://doi.org/10.1051/epjpv/2021002.

[14] Fraunhofer ISE, Photvoltaics report, updated 24 February 2022, 2022. https://www.ise.fraunhofer.de/en/publications/studies/photovoltaics-report.html (accessed April 6, 2022).
[15] G. Masson, I. Kaizuka, Trends in photovoltaic applications 2021, Report IEA PVPS T1-41:2021. (2021).
[16] IRENA, IRENA statistics, 2021. https://irena.org/Statistics (accessed December 1, 2021).
[17] D.J. van de Ven, I. Capellan-Peréz, I. Arto, I. Cazcarro, C. de Castro, P. Patel, M. Gonzalez-Eguino, The potential land requirements and related land use change emissions of solar energy, Scientific Reports. 11 (2021). https://doi.org/10.1038/s41598-021-82042-5.
[18] A. Sahu, N. Yadav, K. Sudhakar, Floating photovoltaic power plant: A review, Renewable and Sustainable Energy Reviews. 66 (2016) 815–824. https://doi.org/10.1016/j.rser.2016.08.051.
[19] K. Komoto, T. Ehara, H. Xu, F. Lv, S. Wang, P. Sinha, E. Cunow, A. Wade, D. Faiman, K. Araki, et al., Energy from the desert: Very large scale PV power plants for shifting to renewable energy future, Report IEA-PVPS T8-01:2015. (2015).
[20] M. Popkin, A. Krishnan, The future of landfills is bright, RMI Report. (2021).
[21] M. Littwin, F.P. Baumgartner, M. Green, W. van Sark, Performance of New Photovoltaic System Designs, Report IEA-PVPS T13-15:2021. (2021).
[22] A.K. Shukla, K. Sudhakar, P. Baredar, Recent advancement in BIPV product technologies: A review, Energy and Buildings. 140 (2017) 188–195. https://doi.org/10.1016/j.enbuild.2017.02.015.
[23] D. Jordan, T. Barnes, N. Haegel, I. Repins, Build solar-energy systems to last-save billions, Nature. 600 (2021) 215–217.
[24] M.J. Brunisholz, IEA PVPS annual report 2020, (2021).
[25] E. Vartiainen, G. Masson, C. Breyer, D. Moser, E. Román Medina, Impact of weighted average cost of capital, capital expenditure, and other parameters on future utility-scale PV levelised cost of electricity, Progress in Photovoltaics: Research and Applications. 28 (2020) 439–453. https://doi.org/10.1002/pip.3189.
[26] D.C. Jordan, C. Deline, M.G. Deceglie, A. Nag, G.M. Kimball, A.B. Shinn, J.J. John, A.A. Alnuaimi, A.B.A. Elnosh, W. Luo, A. Jain, M.U. Saleh, H. von Korff, Y. Hu, J.-N. Jaubert, F. Mavromatakis, Reducing interanalyst variability in photovoltaic degradation rate assessments, IEEE Journal of Photovoltaics. 10 (2019) 206–212. https://doi.org/10.1109/jphotov.2019.2945191.
[27] S. Lindig, D. Moser, A.J. Curran, K. Rath, A. Khalilnejad, R.H. French, M. Herz, B. Müller, G. Makrides, G. Georghiou, A. Livera, M. Richter, J. Ascencio-Vásquez, M. van Iseghem, M. Meftah, D. Jordan, C. Deline, W. van Sark, J.S. Stein, M. Theristis, B. Meyers, F. Baumgartner, W. Luo, International collaboration framework for the calculation of performance loss rates: Data quality, benchmarks, and trends (towards a uniform methodology), Progress in Photovoltaics: Research and Applications. 29 (2021). https://doi.org/10.1002/pip.3397.

[28] J.S. Stein, B. Farnung, PV performance modeling methods and practices, Report IEA-PVPS T13-06:2017. (2017).

[29] M.R. Maghami, H. Hizam, C. Gomes, M.A. Radzi, M.I. Rezadad, S. Hajighorbani, Power loss due to soiling on solar panel: A review, Renewable and Sustainable Energy Reviews. 59 (2016) 1307–1316. https://doi.org/10.1016/j.rser.2016.01.044.

[30] D.C. Jordan, S.R. Kurtz, Photovoltaic degradation rates - An Analytical Review, Progress in Photovoltaics: Research and Applications. 21 (2013) 12–29. https://doi.org/10.1002/pip.1182.

[31] A.Ø. Helland, Forsto ikke hvorfor solcelleanlegget produserte så lite strøm – svaret førte til reklamasjon, Nordlys. (2022). https://www.nordlys.no/forsto-ikke-hvorfor-solcelleanlegget-produserte-sa-lite-strom-svaret-forte-til-reklamasjon/s/5-34-1582057?key=2022-03-13T07:46:59.000Z/retriever/b850237a701fe281c25a795db2de69cef93e844b (accessed March 18, 2022).

[32] B. Marion, R. Schaefer, H. Caine, G. Sanchez, Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations, Solar Energy. 97 (2013) 112–121. https://doi.org/10.1016/j.solener.2013.07.029.

[33] C. Honsberg, S. Bowden, Photovoltaics education website, (2019). www.pveducation.org (accessed February 2, 2022).

[34] A. Smets, K. Jager, O. Isabellea, R. van Swaaij, M. Zeman, Solar energy: The physics and engineering of photovoltaic conversion technologies and systems, UIT Cambridge LTD, 2016.

[35] W. Hermann, G.C.; Eder, B. Farnung, G. Friesen, M. Köntges, B. Kubicek, O. Kunz, H. Liu, D. Parlevliet, I. Tsnakas, J. Vedde, Qualification of photovoltaic (PV) power plants using mobile test equipment, Report IEA-PVPS T13-24:2021. (2021).

[36] IEC, Photovoltaic system performance - Part 1: Monitoring, IEC 61724-1:2021. (2021).

[37] H.E. Beck, N.E. Zimmermann, T.R. McVicar, N. Vergopolan, A. Berg, E.F. Wood, Present and future köppen-geiger climate classification maps at 1-km resolution, Scientific Data. 5 (2018). https://doi.org/10.1038/sdata.2018.214.

[38] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F.J. Martinez-de-Pison, F. Antonanzas-Torres, Review of photovoltaic power forecasting, Solar Energy. 136 (2016) 78–111. https://doi.org/10.1016/j.solener.2016.06.069.

[39] Y. Wang, D. Millstein, A.D. Mills, S. Jeong, A. Ancell, The cost of day-ahead solar forecasting errors in the United States, Solar Energy. 231 (2022) 846–856. https://doi.org/10.1016/j.solener.2021.12.012.

[40] A.P. Dobos, PVWatts version 5 manual, Report NREL/TP-6A20-62641. (2014). https://doi.org/10.2172/1158421.

[41] M. Richter, C. Tjengdrawira, J. Vedde, M. Green, L. Frearson, B. Herteleer, U. Jahn, M. Herz, M. Köntges, Technical assumptions used in PV financial models - review of current practices and recommendations, Report IEA-PVPS T13-08:2017. (2017).

[42] PVsyst, Array losses, general considerations, (2021). https://www.pvsyst.com/help/array_losses_general.htm (accessed April 11, 2022).
[43] M.M. Fouad, L.A. Shihata, E.S.I. Morgan, An integrated review of factors influencing the performance of photovoltaic panels, Renewable and Sustainable Energy Reviews. 80 (2017) 1499–1511. https://doi.org/10.1016/j.rser.2017.05.141.
[44] N. Martín, J.M. Ruiz, Annual angular reflection losses in PV modules, Progress in Photovoltaics: Research and Applications. 13 (2005) 75–84. https://doi.org/10.1002/pip.585.
[45] R.E. Pawluk, Y. Chen, Y. She, Photovoltaic electricity generation loss due to snow – A literature review on influence factors, estimation, and mitigation, Renewable and Sustainable Energy Reviews. 107 (2019) 171–182. https://doi.org/10.1016/j.rser.2018.12.031.
[46] L. Micheli, M. Muller, An investigation of the key parameters for predicting PV soiling losses, Progress in Photovoltaics: Research and Applications. 25 (2017) 291–307. https://doi.org/10.1002/pip.2860.
[47] E. O'Shaughnessy, J.R. Cruce, K. Xu, Too much of a good thing? Global trends in the curtailment of solar PV, Solar Energy. 208 (2020) 1068–1077. https://doi.org/10.1016/j.solener.2020.08.075.
[48] M. Köntges, S. Kurtz, C. Packard, U. Jahn, K.A. Berger, K. Kato, T. Friesen, H. Liu, M. van Iseghem, Review of failures of photovoltaic modules, Report IEA-PVPS T13-01:2014. (2014).
[49] M. Köntges, G. Oreski, U. Jahn, M. Herz, P. Hacke, K.-A. Weiss, et al., Assessment of photovoltaic module failures in the field, Report IEA-PVPS T13-09:2017. (2017).
[50] S. Baschel, E. Koubli, J. Roy, R. Gottschalg, Impact of component reliability on large scale photovoltaic systems' performance, Energies (Basel). 11 (2018). https://doi.org/10.3390/en11061579.
[51] Z. Song, J. Liu, H. Yang, Air pollution and soiling implications for solar photovoltaic power generation: A comprehensive review, Applied Energy. 298 (2021). https://doi.org/10.1016/j.apenergy.2021.117247.
[52] P. Jia, Y. Chua, S. En, R. Tay, Thermal model for heat transfer in photovoltaic modules - what meaning does validation hold?, in: 2021 IEEE 48th Photovoltaic Specialists Conference (PVSC), 2021.
[53] J. Ascencio-Vásquez, I. Kaaya, K. Brecl, K.A. Weiss, M. Topič, Global climate data processing and mapping of degradation mechanisms and degradation rates of PV modules, Energies (Basel). 12 (2019). https://doi.org/10.3390/en12244749.
[54] G.T. Klise, J.R. Balfour, T.J. Keating, Solar PV O&M standards and best practices - existing gaps and improvement efforts, Report SAND2014-19432. (2014).
[55] C. Whaley, Best practices in photovoltaic system operations and maintenance, Report NREL/TP-7A40-67553. (2016).
[56] M. Lindh, M. Svedjeholm, A. Granlund, J. Petersson, A.M. Petersson, Handbok för nordlig solel, RISE Rapportnummer: 2020:61. (2020).

[57] A.G. Imenes, J. Selj, Irradiance and temperature distributions at high latitudes: Design implications for photovoltaic systems, in: 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC), 2017: pp. 619–625. https://doi.org/10.1109/PVSC.2017.8366376.

[58] W.F. Holmgren, C.W. Hansen, M.A. Mikofski, pvlib python: A python package for modeling solar energy systems, Journal of Open Source Software. 3 (2018).

[59] T. Huld, R. Müller, A. Gambardella, A new solar radiation database for estimating PV performance in Europe and Africa, Solar Energy. 86 (2012) 1803–1815. https://doi.org/10.1016/j.solener.2012.03.006.

[60] W. de Soto, S.A. Klein, W.A. Beckman, Improvement and validation of a model for photovoltaic array performance, Solar Energy. 80 (2006) 78–88. https://doi.org/10.1016/j.solener.2005.06.010.

[61] K.-A. Weiss, L.S. Bruckman, R.H. French, G. Oreski, T. Tanahashi, Service life estimation for photovoltaic modules, Report IEA-PVPS T13-16:2021. (2021).

[62] D.C. Jordan, S.R. Kurtz, K. VanSant, J. Newmiller, Compendium of photovoltaic degradation rates, Progress in Photovoltaics: Research and Applications. 24 (2016) 978–989. https://doi.org/10.1002/pip.2744.

[63] C. Buerhop, S. Wirsching, A. Bemm, T. Pickel, P. Hohmann, M. Nieß, C. Vodermayer, A. Huber, B. Glück, J. Mergheim, C. Camus, J. Hauch, C.J. Brabec, Evolution of cell cracks in PV-modules under field and laboratory conditions, Progress in Photovoltaics: Research and Applications. 26 (2018) 261–272. https://doi.org/10.1002/pip.2975.

[64] P. Romer, A. Beinert, A.J. Beinert, Effects of inhomogeneous snow load on the mechanics of a PV module, in: Proceedings of the 38th European PV Solar Energy Conference and Exhibition, 2021. https://doi.org/10.4229/EUPVSEC20212021-4BO.4.5.

[65] A.G. Imenes, Performance of BIPV and BAPV installations in Norway, in: 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC), 2016: pp. 3147–3152. https://doi.org/10.1109/PVSC.2016.7750246.

[66] A.G. Imenes, H.G. Beyer, K. Boysen, J.O. Odden, R.E. Grundt, Performance of grid-connected PV system in Southern Norway, in: 2015 IEEE 42nd Photovoltaic Specialist Conference (PVSC), 2015. https://doi.org/10.1109/PVSC.2015.7355823.

[67] T. Haumann, A brief look at the performance of PV in Norway, Master thesis, UiT The Arctic University of Norway, 2016.

[68] A.G. Imenes, Performance of zero energy homes in smart village Skarpnes, 2016 IEEE 43rd Photovoltaic Specialists Conference (PVSC). (2016) 3153–3158. https://doi.org/10.1109/PVSC.2016.7750247.

[69] E. Schelin, Photovoltaic system yield evaluation in Sweden, Master thesis, Mälardalens Högskola, 2019.

[70] Ø. Kleven, H. Persson, C. Good, W. Sulkowski, T. Boström, Solar cells above the arctic circle - A comparison between a two-axis tracking system and simulations, in: Proceedings of the 24th European PV Solar Energy Conference, 2009: pp. 4090–4093.

[71] E. Molin, B. Stridh, A. Molin, E. Wackelgard, Experimental yield study of bifacial PV modules in Nordic conditions, IEEE Journal of Photovoltaics. 8 (2018). https://doi.org/10.1109/JPHOTOV.2018.2865168.

[72] J.S. Stein, C. Reise, J.B. Castro, G. Friesen, G. Maugeri, E. Urrejola, S. Ranta, Bifacial photovoltaic modules and systems: Experience and results from international research and pilot applications, Report IEA-PVPS T13-14:2021. (2021).

[73] H.N. Riise, M. Øgaard, J. Zhu, C.C. You, F. Andersson, J. Young, S.E. Foss, Performance analysis of a BAPV bifacial system in Norway, in: 2021 IEEE 48th Photovoltaic Specialists Conference (PVSC), 2021: pp. 1304–1308.

[74] B.R. Paudyal, A.G. Imenes, Performance assessment of field deployed multi-crystalline PV modules in Nordic conditions, in: 2019 46th IEEE Photovoltaic Specialists Conference (PVSC), 2019.

[75] E.B. Sveen, M.B. Øgaard, J.H. Selj, G. Otnes, PV system degradation rates in the Nordics, in: Proceedings of the 37th European PV Solar Energy Conference and Exhibition, 2020.

[76] G. Otnes, M.B. Øgaard, L.T. Milde, S.E. Foss, J.H. Selj, Detailed loss analysis for wall mounted photovoltaic systems at high latitude; a comparison of multicrystalline Si- to CIGS-modules, in: Proceedings of the 36th European PV Solar Energy Conference and Exhibition, 2019: pp. 1621–1626.

[77] S. Daliento, A. Chouder, P. Guerriero, A.M. Pavan, A. Mellit, R. Moeini, P. Tricoli, Monitoring, diagnosis, and power forecasting for photovoltaic fields: A review, International Journal of Photoenergy. 2017 (2017). https://doi.org/10.1155/2017/1356851.

[78] B. Marion, J. Adelstein, K. Boyle, H. Hayden, B. Hammond, T. Fletcher, D. Narang, A. Kimber, L. Mitchell, S. Richter, Performance parameters for grid-connected PV systems, in: Conference Record of the Thirty-First IEEE Photovoltaic Specialists Conference, 2005., 2005: pp. 1601–1606. https://doi.org/10.1109/PVSC.2005.1488451.

[79] A. Woyte, M. Richter, D. Moser, M. Green, S. Mau, H.G. Beyer, Analytical monitoring of grid-connected photovoltaic systems, Report IEA-PVPS T13-03. (2014). https://doi.org/10.13140/2.1.1133.6481.

[80] T. Dierauf, A. Growitz, S. Kurtz, C. Hansen, Weather-corrected performance ratio, Report NREL/TP-5200-57991. (2013).

[81] A. Triki-Lahiani, A.B.-B. Abdelghani, I. Slama-Belkhodja, Fault detection and monitoring systems for photovoltaic installations: A review, Renewable and Sustainable Energy Reviews. 82 (2018) 2680–2692. https://doi.org/10.1016/j.rser.2017.09.101.

[82] S. Rodrigues, H.G. Ramos, F. Morgado-Dias, Machine learning in PV fault detection, diagnostics and prognostics: A review, in: 2017 IEEE 44th Photovoltaic Specialist Conference (PVSC), 2017: pp. 3178–3183. https://doi.org/10.1109/PVSC.2017.8366581.

[83] D.C. Jordan, C. Deline, S.R. Kurtz, G.M. Kimball, M. Anderson, Robust PV degradation methodology and application, IEEE Journal of Photovoltaics. 8 (2017) 525–531. https://doi.org/10.1109/JPHOTOV.2017.2779779.

[84] Å.F. Skomedal, M.B. Øgaard, H. Haug, E.S. Marstein, Robust and fast detection of small power losses in large-scale PV systems, IEEE Journal of Photovoltaics. 11 (2021) 819–826. https://doi.org/10.1109/JPHOTOV.2021.3060732.

[85] B.L. Aarseth, Å.F. Skomedal, M.B. Øgaard, E.S. Marstein, Detecting permanently activated bypass diodes in utility scale PV plant monitoring data, Manuscript Submitted for Publication. (2022).

[86] S. Stettler, P. Toggweiler, E. Wiemken, W. Heydenreich, A. de Keizer, W. van Sark, S. Feige, et al., Failure detection routine for grid-connected PV systems as part of the PVSAT-2 project, in: Proceedings of the 20th European PV Solar Energy Conference, 2005: pp. 2490–2493.

[87] A.J. Curran, C.B. Jones, S. Lindig, J. Stein, D. Moser, R.H. French, Performance loss rate consistency and uncertainty across multiple methods and filtering criteria, in: 2019 IEEE 46th Photovoltaic Specialists Conference, 2019: pp. 1328–1334.

[88] D.C. Jordan, S.R. Kurtz, The dark horse of evaluating long-term field performance - Data filtering, IEEE Journal of Photovoltaics. 4 (2014) 317–323. https://doi.org/10.1109/JPHOTOV.2013.2282741.

[89] A. Livera, G. Paphitis, M. Theristis, J. Lopez-Lorente, G. Makrides, G.E. Georghiou, Photovoltaic system health-state architecture for data-driven failure detection, Solar. 2 (2022) 81–98. https://doi.org/10.3390/solar2010006.

[90] Å. Skomedal, M.B. Øgaard, J.H. Selj, H. Haug, E.S. Marstein, General, robust and scalable methods for string level monitoring in utility scale PV systems, in: Proceedings of the 36th European PV Solar Energy Conference and Exhibition, 2019: pp. 1283–1287.

[91] B. Meyers, M. Deceglie, C. Deline, D. Jordan, Signal processing on PV time-series data: Robust degradation analysis without physical models, IEEE Journal of Photovoltaics. 10 (2019) 546–553. https://doi.org/10.1109/JPHOTOV.2019.2957646.

[92] B. Meyers, E. Apostolaki-Iosifidou, L.T. Schelhas, Solar data tools: Automatic solar data processing pipline, in: 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), 2020: pp. 655–656.

[93] IEC, Photovoltaic system performance - Part 3: Energy evaluation method, IEC TS 61724-3:2016. (2016).

[94] M.B. Strobel, T.R. Betts, G. Friesen, H.G. Beyer, R. Gottschalg, Uncertainty in photovoltaic performance parameters - dependence on location and material, Solar Energy Materials and Solar Cells. 93 (2009) 1124–1128. https://doi.org/10.1016/j.solmat.2009.02.003.

[95] Åsmund Skomedal, Data-based approaches to efficient operation and maintenance of PV systems, PhD thesis, University of Oslo, 2021.

[96] S. Lindig, A. Louwen, D. Moser, M. Topic, Outdoor PV system monitoring—Input data quality, data imputation and filtering approaches, Energies (Basel). 13 (2020) 1–18. https://doi.org/10.3390/en13195099.

[97] J. Meydbray, K. Emery, S. Kurtz, Pyranometers and reference cells, what's the difference?, NREL/JA-5200-54498. (2012).

[98] M.G. Kratzenberg, H.G. Beyer, S. Colle, A. Albertazzi, Uncertainty calculations in pyranometer measurements and application, ASME 2006 International Solar Energy Conference. (2006) 689–698. https://doi.org/10.1115/ISEC2006-99168.

[99] K. Kiefer, D. Dirnberger, B. Müller, W. Heydenreich, A. Kröger-Vodde, A degradation analysis of PV power plants, in: Proceedings of the 25th European PV Solar Energy Conference, 2010: pp. 5032–5037. https://doi.org/10.4229/25thEUPVSEC2010-5BV.4.26.

[100] D.L. King, J.A. Kratochvil, W.E. Boyson, Photovoltaic array performance model, Report SAND2004-3535. (2004). https://doi.org/10.2172/919131.

[101] N. Shanmugam, R. Pugazhendhi, R.M. Elavarasan, P. Kasiviswanathan, N. Das, Anti-reflective coating materials: A holistic review from PV perspective, Energies (Basel). 13 (2020). https://doi.org/10.3390/en13102631.

[102] M.G. Decegli, M. Muller, D.C. Jordan, C. Deline, Numerical validation of an algorithm for combined soiling and degradation analysis of photovoltaic systems, in: 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), 2019: pp. 3111–3114.

[103] Å. Skomedal, M.G. Deceglie, Combined estimation of degradation and soiling losses in photovoltaic systems, IEEE Journal of Photovoltaics. 10 (2020) 1788–1796. https://doi.org/10.1109/JPHOTOV.2020.3018219.

[104] S. Silvestre, L. Mora-López, S. Kichou, F. Sánchez-Pacheco, M. Dominguez-Pumar, Remote supervision and fault detection on OPC monitored PV systems, Solar Energy. 137 (2016) 424–433. https://doi.org/10.1016/j.solener.2016.08.030.

[105] G. Belluardo, P. Ingenhoven, W. Sparber, J. Wagner, P. Weihs, D. Moser, Novel method for the improvement in the evaluation of outdoor performance loss rate in different PV technologies and comparison with two other methods, Solar Energy. 117 (2015) 139–152. https://doi.org/10.1016/j.solener.2015.04.030.

[106] K. Perry, M. Muller, K. Anderson, Performance comparison of clipping detection techniques in AC power time series, in: 2021 IEEE 48th Photovoltaic Specialists Conference (PVSC), 2021: pp. 1638–1643. https://doi.org/10.1109/PVSC43889.2021.9518733.

[107] R.B. Cleveland, W.S. Cleveland, J.E. McRae, I. Terpenning, STL: A seasonal-trend decomposition procedure based on loess, Journal of Official Statistics. 6 (1990).

[108] A. Phinikarides, N. Kindyni, G. Makrides, G.E. Georghiou, Review of photovoltaic degradation rate methodologies, Renewable and Sustainable Energy Reviews. 40 (2014) 143–152. https://doi.org/10.1016/j.rser.2014.07.155.

[109] K. Kunaifi, A. Reinders, S. Lindig, M. Jaeger, D. Moser, Operational performance and degradation of PV systems consisting of six technologies in three climates, Applied Sciences. 10 (2020). https://doi.org/10.3390/APP10165412.

[110] W.G. Rees, Remote sensing of snow and ice, CRC Press, 2005.
[111] T. Saloranta, J. Andersen, Simulations of snow depth in Norway in a projected future climate (2071-2100), NVE Report 12. (2018).
[112] W.K. Wong, I. Haddeland, D. Lawrence, S. Beldring, Gridded 1 x 1 km climate and hydrological projections for Norway, NVE Report 59. (2016).
[113] D. Jacob, J. Petersen, B. Eggert, A. Alias, O.B. Christensen, L.M. Bouwer, A. Braun, et al., EURO-CORDEX: New high-resolution climate change projections for European impact research, Regional Environmental Change. 14 (2014). https://doi.org/10.1007/s10113-013-0499-2.
[114] Norwegian Water Resources and Energy Directorate (NVE), Norsk klimaservicesenter - nedlasting av griddata, (2017). https://nedlasting.nve.no/klimadata/kss.
[115] D. Riley, L. Burnham, B. Walker, J.M. Pearce, Differences in snow shedding in photovoltaic systems with framed and frameless modules, in: 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), 2019: pp. 558–561. https://doi.org/10.1109/PVSC40753.2019.8981389.
[116] L. Burnham, D. Riley, B. Walker, J.M. Pearce, Performance of bifacial photovoltaic modules on a dual-axis tracker in a high-latitude, high-albedo environment, in: 2019 IEEE 46th Photovoltaic Specialists Conference (PVSC), 2019: pp. 1320–1327. https://doi.org/10.1109/PVSC40753.2019.8980964.
[117] A. Dhyani, C. Pike, J.L. Braid, E. Whitney, L. Burnham, A. Tuteja, Facilitating large-scale snow shedding from in-field solar arrays using icephobic surfaces with low-interfacial toughness, Advanced Materials Technologies. (2021). https://doi.org/10.1002/admt.202101032.
[118] B.B. Aarseth, M.B. Øgaard, J. Zhu, T. Strömberg, J.A. Tsanakas, J.H. Selj, E.S. Marstein, Mitigating snow on rooftop PV systems for higher energy yield and safer roofs, in: Proceedings of the 35th European PV Solar Energy Conference and Exhibition, 2018: pp. 1630–1635. https://doi.org/10.4229/35thEUPVSEC2018-6CO.3.5.
[119] A.C. de Keizer, W. van Sark, S. Stettler, P. Toggweiler, E. Lorenz, A. Drews, et al., PVSAT-2: Results of field test of the satellite-based PV system performance check, in: Proceedings of the 21st European Photovoltaic Solar Energy Conference, 2006: pp. 2681–2685.
[120] Z. DeFreitas, G. Binnard, T. Delsart, Using on-site ambient temperature and performance ratio to identify days when snow cover affects PV plant production, in: 2021 48th IEEE Photovoltaic Specialists Conference (PVSC), 2021: pp. 1860–1864. https://doi.org/10.1109/pvsc43889.2021.9518551.
[121] J.L. Braid, D. Riley, J.M. Pearce, L. Burnham, Image analysis method for quantifying snow losses on PV systems, in: 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), 2020: pp. 1510–1516. https://doi.org/10.1109/PVSC45281.2020.9300373.

[122] H.N. Riise, M.B. Øgaard, T.U. Nærland, Soiling and snow impact on a PV plant at a farm in Norway, in: Proceedings of the 38th European PV Solar Energy Conference, 2021: pp. 1241–1244.

[123] D. Ryberg, J. Freeman, Integration, validation, and application of a PV snow coverage model in SAM, Report NREL/TP-6A20-68705. (2017).

[124] M. van Noord, T. Landelius, S. Andersson, Snow-induced PV loss modeling using production-data inferred PV system models, Energies (Basel). 14 (2021). https://doi.org/10.3390/en14061574.

[125] R.W. Andrews, J.M. Pearce, Prediction of energy effects on photovoltaic systems due to snowfall events, in: 2012 38th IEEE Photovoltaic Specialists Conference, 2012: pp. 3386–3391. https://doi.org/10.1109/PVSC.2012.6318297.

[126] T. Townsend, L. Powers, Photovoltaics and snow: An update from two winters of measurements in the Sierra, in: 2011 37th IEEE Photovolt. Spec. Conf., 2011: pp. 3231–3236. https://doi.org/10.1109/PVSC.2011.6186627.

[127] L. Powers, J. Newmiller, T. Townsend, Measuring and modeling the effect of snow on photovoltaic system performance, in: 2010 35th IEEE Photovoltaic Specialists Conference, 2010: pp. 973–978. https://doi.org/10.1109/PVSC.2010.5614572.

[128] J. Fagerström, L. Kvalbein, J. Danebergs, T. Uberg Nærland, M. Øgaard, K. Aamodt Espegren, Hybrid PV systems and colocalization of charging and filling stations for electrification of road transport sector, Solar RRL. (2022). https://doi.org/10.1002/solr.202100461.

[129] L. Burnham, D. Riley, J. Braid, Design considerations for photovoltaic systems deployed in snowy climates, in: Proceedings of the 37th European PV Solar Energy Conference and Exhibition, 2020: pp. 1626–1631.

Papers

I

Paper I

Methods for quality control of monitoring data from commercial PV systems

M.B. Øgaard, H. Haug, and J.H. Selj

In: *Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition* (2018), pp. 2083-2088.
DOI: 10.4229/35thEUPVSEC20182018-6DV.1.53

METHODS FOR QUALITY CONTROL OF MONITORING DATA FROM COMMERCIAL PV SYSTEMS

Mari B. Øgaard[1,2*], Halvard Haug[1,2] and Josefine Selj[1,2]
[1]Department of Technology Systems, University of Oslo
Gunnar Randers vei 19, 2007 Kjeller, Norway
[2]Department of Renewable Energy Systems, Institute for Energy Technology
Instituttveien 18, 2007 Kjeller, Norway
*Corresponding author. E-mail: mari.ogaard@ife.no, Tel.: +47 976 356 08

The aim of this work is to develop and test new methods for quality control of data from commercial monitoring systems for small and medium sized PV installations. Such installations often have limited or non-existent maintenance of their monitoring systems. Quality issues in e.g. irradiance and temperature measurements will cause errors in the analysis of the PV system performance and might lead to non-optimal maintenance of the system. To determine the condition of the sensors and the monitoring system based on the measured data itself is therefore essential to improve performance analysis algorithms and to understand historical data from these types of PV systems, and consequently this is of significant economical and practical value. In this work, we use data from both commercial and research systems in Norway to assess the robustness of the methods in a real-world scenario. We demonstrate that drift and deviations in the sensitivity of irradiance sensors, in addition to misalignment of the sensors, can be accurately quantified and detected based on comparison with clear sky irradiance modeling. Furthermore, we show that analysis of temperature data potentially can be used to detect snow cover of modules, in addition to identification of detached temperature sensors.

Keywords: monitoring, PV system, data quality

1 INTRODUCTION

Regulations and irradiation conditions greatly influence the size and type of PV systems installed in a given market. In Norway, as in other northern climates, moderate solar irradiance and incentives for self-consumption has resulted in a market dominated by relatively small PV systems with installed capacity less than 1 MWp. Moderately sized commercial systems constitute a significant fraction. Typically, these systems have a simple monitoring system, measuring both electrical and environmental data. Most commonly, the electrical data is collected from the inverter, the plane of array (POA) irradiance is measured by a reference cell, and there are sensors measuring ambient temperature and the module temperature. In addition to the typically low accuracy of commercial monitoring solutions [1,2], often very little maintenance is performed on these systems, as the cost of maintenance of distributed monitoring systems may exceed the expected benefit if local personnel are not available. In most cases, necessary maintenance like cleaning of the irradiance sensors and visual inspection of the system and sensors is not performed, and the sensors are not regularly recalibrated. Drift in sensors, dirty irradiance sensors and detached module temperature sensors can lead to significant misinterpretation of the PV system performance, and consequently also suboptimal or unnecessary maintenance of the PV system. The distribution of the systems, combined with little local competence and small economic incentives, makes increased manual supervision an unrealistic solution. A more realistic path ahead is quality control of the sensor data based on data analysis. In addition to saving money by reducing the need for maintenance, this approach also enables validation of the quality of historical data for any system, independent of previous monitoring system maintenance routines.

Data-based evaluation of the condition of the sensors and the monitoring system is not widely discussed in the literature. The quality control work done today is mostly limited to detecting abnormal points (i.e. outliers and data exceeding physical possible limits) [3–5], as opposed to detecting permanent changes in the measurement over longer periods. In this work we suggest a new data-based method for detecting changes in the quality or accuracy of the irradiance measurements. We also discuss a method for quality control of module temperature measurements and how this method can be used to identify snow cover on the modules. Analysis is performed using data from commercial PV systems in addition to scientific test sites to assess the robustness of the methods in a real-world scenario.

The challenge of missing maintenance and supervision of irradiance sensors is also discussed by Jordan et al. [6], who proposed to calculate PV system degradation based on clear sky irradiance simulations instead of measured irradiance. However, ground measurements provide valuable additional information on the performance of the PV system under cloudy conditions and enable a more comprehensive analysis of PV systems. This is particularly important for locations with few clear sky days.

In this work we use clear sky irradiance simulations to assess the reliability and quality of the measured irradiance. This has previously been suggested by Reno et al. [7], but is to our knowledge not tested before. The method is assessed for both high and low-quality irradiance sensors.

A relatively common problem with measurements of module temperature, is detachment of temperature sensors, which are normally attached to the back sheet of the module. This could be detected by monitoring the difference between the module and the ambient temperature. To our knowledge, this method has not been tested beyond the observations presented by Woyte et al. [8]. In this work we test this method for two different types of PV installation, and assess it for a new application; snow detection. The possibility of detecting snow coverage using measurement equipment that already exists on-site, could improve both the monitoring algorithms for PV system fault detection and the energy generation forecasts by separating snow events from other failures. Additionally, it will also simplify the analysis of historical data, with respect to estimation of performance and validation of snow loss models. In previous studies, the probability of snow cover on PV modules is estimated by detecting when the PV production is low relative to irradiance or estimated irradiance in combination with evaluation of other parameters, including ambient temperature [9], predictions of snow depth and temperature [10], and satellite observations [11]. Using ambient temperature alone will not separate snow events from e.g. total black outs or other serious failures [12]. The aim of this work is to increase the accuracy of snow detection at a specific location by also using the module temperature sensor data.

2 METHODS

2.1 Test of irradiance measurements based on clear sky irradiance simulations

Global horizontal clear sky irradiance is simulated using the Ineichen and Perez clear sky model based on zenith angle, air mass, elevation and Linke Turbidity [13]. The model error has been shown to have low dependency of time of the day and day of the year compared to other clear sky models [7]. The simulated clear sky irradiance is then transposed to a tilted plane by calculating beam, reflected and diffuse irradiance in the plane. The sky diffuse irradiance is calculated using the isotropic sky model [14], where the sky diffuse irradiance in the plane of the PV array is found using the diffuse horizontal irradiance, the tilt angle of the plane of array, and the assumption that the sky is a uniform source of irradiance. Periods with measured irradiance equivalent to clear sky conditions, i.e. a smooth irradiance curve, was detected using the algorithm proposed by Reno and Hansen, which compares GHI time series statistics to the Ineichen clear sky model [15]. The length of the clear sky time periods selected are at least two hours, to optimize between amount of data and correct selection of clear sky periods. All models are implemented in compliance with the methods in the Matlab version of the PV_LIB Toolbox [16], using the default Link turbidity values provided by SoDa in the clear sky modeling.

The modeled clear sky irradiance is in this case used as a reference, and the measured data is compared to the modeled results in the periods with irradiance equivalent to clear sky conditions. The relative difference (ΔI) between measured (I_{meas}) and modeled irradiance (I_{CS}) is given by:

$$\Delta I = (I_{meas} - I_{CS})/I_{meas} \qquad (1)$$

To evaluate how the measurements change relative to the model from year to year, a scaling factor (α) is estimated for every year by minimizing RMSE between measured and clear sky irradiance, given by [15]:

$$RMSE(\alpha) = \sqrt{\frac{\sum_{i=1}^{n}(\alpha \times I_{meas} - I_{CS})^2}{n}} \qquad (2)$$

2.2 Method for detection of sensor detachment and snow

The difference in module and ambient temperature is strongly correlated with the irradiance (Eq. 3) [17]. Hence, by monitoring this correlation, signatures of temperature sensor detachment can be identified [8]. We have tested this method for two different installation configurations, by investigating the changes in linear regression fits of hourly values of temperature difference and irradiance for different weeks. Additionally, we have tested if the same approach can be used for snow cover detection.

$$T_{mod} - T_{amb} = I \times e^{a + b \times Wind\ speed} \qquad (3)$$

Where a and b are coefficients compensating for site specific configurations.

2.3 Measurement sites

The proposed method for quality control of irradiance sensors was tested for the uncorrected raw data, given as 10-minute averages, measured by an old pyranometer installed at the supervised weather station at the Norwegian University of Life Sciences. The method was also tested on a commercial flat roof PV system, where the POA irradiance was measured by tilted reference cell and given in 5-minute averages. The tilt of the system and the reference cell is 10°, and the orientation of the modules is south-east and north-west.

The method for detection of sensor detachment is tested for two different types of PV installations: South-orientated modules installed in an open-rack configuration with a tilt of 28°, and modules installed on a roof with a tilt of 35°. The module temperatures are measured by a resistant thermometer attached to the back sheet of the modules, and the POA irradiance is measured by reference cells. For the rack installation, the ambient temperature sensor is installed under the modules, and for the tilted roof installation the ambient temperature is measured by a weather station at the same location.

For both systems, the same approach is tested for snow cover identification. Additionally, snow cover identification is tested for the flat roof system used in the irradiance quality control tests. For this system, the cell temperature of a reference cell is used as an estimation of module temperature, the ambient temperature is measured by a PT-1000 element, and irradiance is measured by a ventilated pyranometer. Normally, the reference cell is covered by snow in the same period as the modules because of the low tilt angles, and the pyranometer is less affected, most likely because of the ventilation and a more elevated installation position. Both roof installations are commercial systems, while the open rack system is a scientific test site. Hourly averages are used in the analysis.

3 RESULTS AND DISCUSSION

3.1 Quality control of irradiance measurements

The irradiance measurement control method was tested on raw data from the pyranometer at the Norwegian University of Life Sciences. As presented in Figure 1, the relative difference between measured and modeled irradiance is scattered and high, especially in periods with low irradiance. In periods with low irradiance and high angle of incidence, the relative errors of both the measurements and the modeled results is probably higher [7,18].

To improve the analysis, the periods with low irradiance were filtered out. The difference between measured irradiance and modeled clear sky irradiance when the measured irradiance was more than 500 W/m^2 is given as a function of time in Figure 2. The sensor was recalibrated and adjusted in 2008, after this there were no significant changes in sensitivity for two years, followed by a decrease in sensitivity by 1% per year in a four-year period (2010-2014). The pyranometer was replaced in 2014, which lead to an increase in the sensitivity of the irradiance measurements at the site of 6%. All these alterations were detectable through comparison with clear sky modeling. The shift in measured irradiance relative to the modeled irradiance at the time of the sensor adjustment and the pyranometer replacement is clearly shown in Figure 2. After the sensor replacement, the scaling factor α also increased by 6 percentage points. The analysis indicates a degradation of the sensor of 1% per year in the period 2010-2014, the same as the independent instrument calibration, based on a reduction of α of 1 percentage point per year.

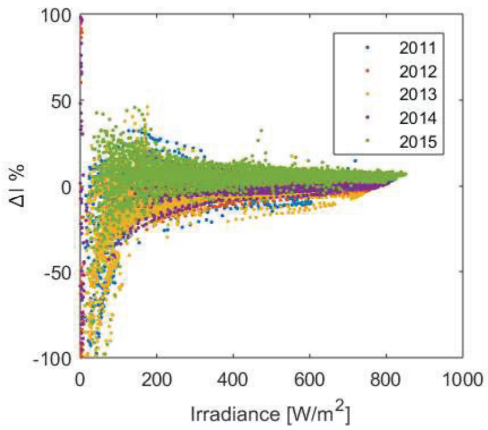


Figure 1: Relative difference between GHI measurements and clear sky irradiance modeling at different irradiance levels.

We observe that the modeled irradiance is slightly biased and most of the time is lower than the measured irradiance, as expected for higher latitudes when using the Ineichen and Perez model [7]. Additionally, the variations between single points is high, indicating that this method cannot be applied for quality control of individual measurements. However, the ability of the method to indicate drift or other permanent changes in the irradiance measurements over time is remarkable.

It is well known that pyranometers are subject to a thermal offset during clear sky periods [19]. This introduce a bias in the comparison between the clear sky model and the measurements. However, as it is the change and not the absolute value that is of importance, this effect is assumed negligible in this context. A more relevant challenge is changes in atmospheric conditions with respect to transmission and scattering. Changes in air pollution are one of the major contributions to what is referred to as global brightening and dimming, which may have an effect of a magnitude that might influence long time PV performance analysis [18]. More practical challenges related to this method for commercial systems, is the low time resolution data these types of systems often have, and potentially also the number of clear sky hours through the year at the specific location.

3.2 Quality control of tilted reference cell measurements

The quality control method for irradiance measurements was also tested for tilted reference cells measuring the POA irradiance of a PV system in a north-west, south-east configuration with a tilt of 10°. Reference cells are, as mentioned earlier, common for medium sized PV systems with monitoring. The uncertainties are however higher than when pyranometers are used. The relative difference between the measurements and the clear sky modeling for irradiance above 500 W/m^2 is presented in Figure 3. The scaling factor between the measured and modeled irradiance is approximately constant from year to year for the irradiance measured in the north-west direction, with a reduction of less than 0.1 percentage points. This is lower than expected degradation for a c-Si reference cell [8].

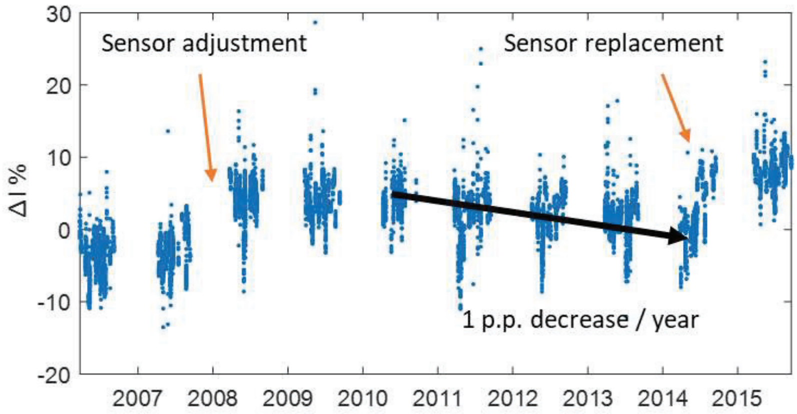


Figure 2: Relative difference between GHI measurements and clear sky irradiance modeling at measured irradiance above 500 W/m^2. Time of adjustment of the sensor, pyranometer replacement, and the period where 1 % decrease per year was measured is marked.

It should be noted that the north-west orientation results in lower irradiance and fewer data points with irradiance values above 500 W/m^2. This could potentially affect the robustness of this result. For the irradiance measured in the south-east direction, the difference between measured and modeled irradiance appeared to be season dependent. It was found that this sensor was misaligned, which could be an explanation for this behavior.

This misalignment also gave a shift between the position of the measured and modeled irradiance curves (as well as the PV power curve), as shown in Figure 4. A consequence of this was unlikely high performance ratios for the PV system in the morning, as the irradiance sensor measured less irradiance than what the PV modules received, and opposite in the evening.

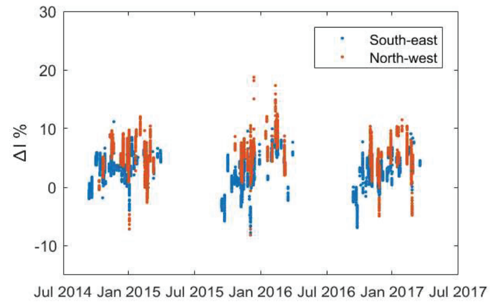


Figure 3: Relative difference between POA measurements and clear sky irradiance modeling at measured irradiance above 500 W/m^2.

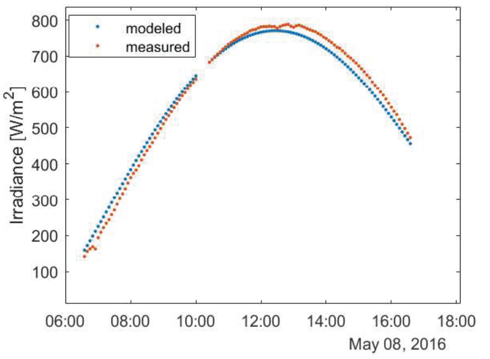


Figure 4: Measured and modeled irradiance with temporal shift in curves because of misalignment of sensor.

3.3 Control of module temperature sensor detachment

The linear relationship between module and ambient temperature difference and irradiance is presented in Figure 5-6 for two different types of installations. Data was selected from four different weeks and different months and years to highlight some important aspects: The correlation between the temperature difference and the irradiance is almost linear and it typically has relatively small variations from year to year and month to month for the ground mounted open rack system, and to some degree also for the roof mounted system. This indicates that it is possible to detect detachment of the module temperature sensor by monitoring changes in the slope. If detached, or partially detached, the module temperature sensor measure a value closer to ambient temperature, and the slope would be less steep or zero. The scattering of the data points is greater for the close roof mounted system and this system also experience the highest temperature difference. One natural explanation for this is the cooling effect of the wind, which is not taken into account in the analysis. Lack of rear side ventilation or heat leakage from the building itself may also be influential. The slope of the regression lines of the close roof mounted system is twice as steep as for the rack mounted system, in agreement with the experimental results of King et al. [17].

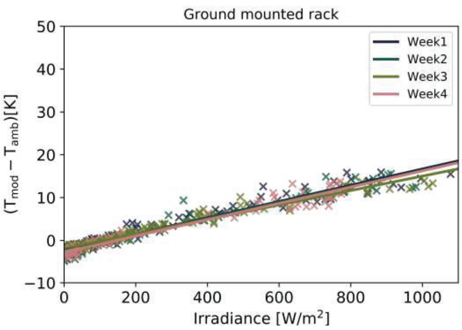


Figure 5: Absolute difference between module and ambient temperature as a function of measured irradiance. Data from four different weeks is shown.

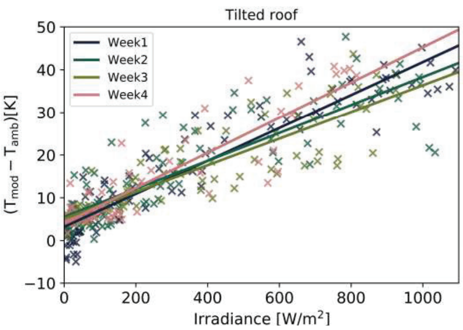


Figure 6: Absolute difference between module and ambient temperature as a function of measured irradiance. Data from four different weeks is shown.

3.4 Use of temperature sensors for snow detection

Snow cover on the PV modules will, like the sensor detachment, have an effect on the slope of the temperature/irradiance regression line. As presented in Figure 8-10, when snow is covering the modules, the correlation between the temperature difference and the irradiance will change substantially, as the module surface and back sheet will be significantly less heated. For the rack mounted system the measured module temperature is lower than the measured ambient temperature. The difference is increasing with increasing irradiance and ambient temperature. For the roof systems, the module temperature is higher than the ambient, most likely because the snow has an isolating effect.

Generally, it might not be possible to separate situations with partly snow-covered modules from situations with low irradiance, or situations when snow is covering the irradiance sensor. The figure showing the snow-covered modules at the tilted roof system, illustrates how it can be challenging to separate periods with low irradiance from periods with snow cover.

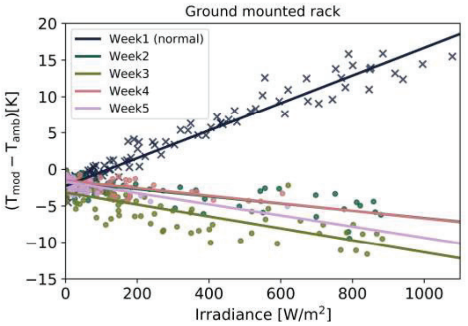


Figure 7: Absolute difference in measured module temperature and ambient temperature as a function of irradiance for weeks where snow is covering the modules compared to a week with normal production.

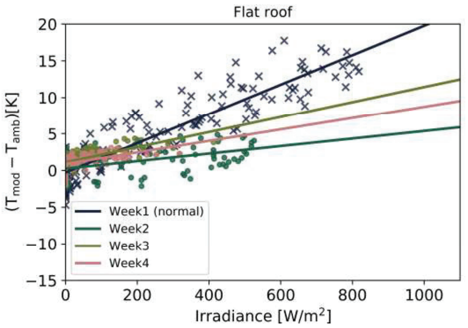


Figure 8: Absolute difference in measured module temperature and ambient temperature as a function of irradiance for weeks where snow is covering the modules compared to a week with normal production.

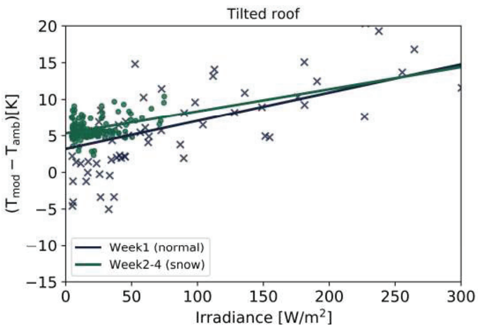


Figure 9: Absolute difference in measured module temperature and ambient temperature as a function of irradiance for weeks where snow is covering the modules compared to a week with normal production.

The main limitation with this method is its dependency of undisturbed irradiance measurements. Low irradiance situations might also be challenging, as the modules do not heat up as much as under normal operation. To give more robust snow detection algorithms, the difference in module and ambient temperature should be combined with PV efficiency data, and potentially also estimated irradiance to both indicate if the sensor is covered by snow and replace the irradiance measurement data.

The placement of the temperature sensors is of importance, as this will define the ability to detect partly covered modules, and give different temperature difference characteristics (i.e. positive, negative or zero). The temperature difference characteristics can also be studied through night time values.

4 CONCLUSION

In this work, a method for detection of permanent changes in irradiance measurements has been tested, as well as a method for detection of module temperature sensor detachment and how the latter can be used to improve detection of snow cover on the PV modules.

Quality control of irradiance measurements based on data anlaysis has been performed for different PV systems with high and low-accuracy irradiance sensors. We show that by using this approach it is possible to accurately detect both abrupt changes, as well as slow gradual changes such as the yearly drift of 1%.

A method for detection of module temperature sensors detachment and snow cover was tested for typical commercial installation configurations in Nordic climates, including open rack mounted, close roof mounted and tilted flat roof mounted systems. The linear relationship between temperature differences and irradiance proved to be stable enough to indicate sensor detachment and snow cover. The robustness of this method might depend on reliable irradiance measurements, sensor placement and installation configuration. Periods with low irradiance and partly snow-covered modules appeared to be challenging. However, in combination with other methods, the presented results show that the use of module temperature sensor increase the confidence and accuracy of forecasting and fault diagnostics with respect to snow cover events.

REFERENCES

[1] M. Fanni, L., Giussani, M., Marzoli, M., & Nikolaeva-Dimitrova, How accurate is a commercial monitoring system for photovoltaic plant?, Prog. Photovoltaics Res. Appl. 22 (2014) 910–922.

[2] M.B. Strobel, T.R. Betts, G. Friesen, H.G. Beyer, R. Gottschalg, Uncertainty in Photovoltaic performance parameters - dependence on location and material, Sol. Energy Mater. Sol. Cells. 93 (2009) 1124–1128.

[3] C. Ventura, G.M. Tina, Utility scale photovoltaic plant indices and models for on-line monitoring and fault detection purposes, Electr. Power Syst. Res. 136 (2016) 43–56.

[4] I. Moradi, Quality control of global solar radiation using sunshine duration hours, Energy. 34 (2009) 1–6.

[5] IEC 61724:1 Photovoltaic system Performance - Monitoring, 2017.

[6] D.C. Jordan, C. Deline, S.R. Kurtz, G.M. Kimball, M. Anderson, Robust PV Degradation Methodology and Application, IEEE J. Photovoltaics. (2017) 1–7.

[7] M.J. Reno, C.W. Hansen, J.S. Stein, Global horizontal irradiance clear sky models : implementation and analysis., (2012).

[8] A. Woyte, M. Richter, D. Moser, M. Green, S. Mau, H.G. Beyer, Analytical Monitoring of Grid-connected Photovoltaic Systems, 2014.

[9] S. Stettler, P. Toggweiler, E. Wiemken, W. Heydenreich, a de Keizer, W. van Sark, S. Feige, M. Scheider, G. Heilscher, E. Lorenz, Failure detection routine for grid-connected PV systems as part of the PVSAT-2 project, 20th Eur. Photovolt. Sol. Energy Conf. (2005) 2490–2493.

[10] E. Lorenz, D. Heinemann, C. Kurz, Local and regional photovoltaic power prediction for large scale grid integration: Assessment of a new algorithm for snow detection, Prog. Photovoltaics Res. Appl. 20 (2011) 760–769.

[11] G. Wirth, M. Schroedter-Homscheidt, M. Zehner, G. Becker, Satellite-based snow identification and its impact on monitoring photovoltaic systems, Sol. Energy. 84 (2010) 215–226.

[12] E. Lorenz, J. Betcke, A. Drews, A.C. de Keizer, S. Stettler, M. Scheider, S. Bofinger, H.G. Beyer, W. Heydenreich, E. Wiemken, W. van Sark, P. Toggweiler, G. Heilscher, D. Heinemann, Intelligent performance check of pv system operation based on satellite data, 2007.

[13] P. Ineichen, R. Perez, A new airmass independent formulation for the linke turbidity coefficient, Sol. Energy. 73 (2002) 151–157.

[14] P.G. Loutzenhiser, H. Manz, C. Felsmann, P.A. Strachan, T. Frank, G.M. Maxwell, Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation, Sol. Energy. 81 (2007) 254–267.

[15] M.J. Reno, C.W. Hansen, Identification of periods of clear sky irradiance in time series of GHI measurements, Renew. Energy. 90 (2016) 520–531.

[16] J.S. Stein, W.F. Holmgren, J. Forbess, C.W. Hansen, PVLIB: Open source photovoltaic performance modeling functions for Matlab and Python, Proc. 43rd IEEE Photovolt. Spec. Conf. (2016) 3425–3430.

[17] D.L. King, J.A. Kratochvil, W.E. Boyson, Photovoltaic array performance model, Online. 8 (2004) 1–19.

[18] C. Reise, B. Müller, Uncertainties in PV System Yield Predictions and Assessments, 2018.

[19] G.S. Hernandez, A. Serrano, M.L. Cancillo, J.A. Garcia, Pyranometer thermal offset: Measurement and analysis, J. Atmos. Ocean. Technol. 32 (2015) 234–246.

II

Paper II

Performance evaluation of monitoring algorithms for photovoltaic systems

M.B. Øgaard, Å. Skomedal, and J.H. Selj

In: *Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition* (2019), pp. 1632-1636.
DOI: 10.4229/EUPVSEC20192019-5CV.4.30

PERFORMANCE EVALUATION OF MONITORING ALGORITHMS FOR PHOTOVOLTAIC SYSTEMS

Mari B. Øgaard*[1,2], Åsmund Skomedal[2] and Josefine H. Selj[1,2]
[1]Department of Technology Systems, University of Oslo, Gunnar Randers vei 19, 2007 Kjeller, Norway
[2]Department of renewable energy systems, Institute for Energy Technology, Instituttveien 18, 2007 Kjeller, Norway
*Corresponding author. E-mail: mari.ogaard@ife.no, Tel.: +47 976 356 08

ABSTRACT: Monitoring solutions for commercial photovoltaic (PV) systems are becoming increasingly widespread, but often performs poorly, especially in locations with varying weather conditions. In this work two standard performance metrics commonly used in PV system monitoring, temperature corrected performance ratio and specific yield, have been calculated and evaluated for real-world conditions. The data is collected from eight inverters of 13-18 kW_p each, installed at a commercial large-scale PV system in Norway. The results show that naïve use of the tested performance metrics give unreliable monitoring with high variation in the PV system performance estimation, often resulting in false alarms. Very low solar elevation and irradiance, snow and technical irregularities in the installation are the primary causes of false alarms in the monitoring. It is shown that for certain climates standard filtering approaches are not sufficient to solve these problems, and that site-specific filtering of data gives more stable monitoring output, entailing more data and less variation.
Keywords: PV systems, Monitoring, Performance, Rooftop

1 INTRODUCTION

With the recent year's increased focus on operation and maintenance of photovoltaic (PV) systems, an extensive number of algorithms and performance metrics have been proposed to improve the PV system monitoring solutions [1]. From very basic to more advanced – the aim of the algorithms is to detect when the PV system is deviating from normal operation and identify faults. The more advanced solutions are also targeting failure diagnosis. Despite that the demand for PV monitoring solutions is growing rapidly, the algorithms are still not sufficiently sophisticated to handle the noise and variations in real-world data in a satisfactory manner, resulting in noise also in the monitoring output. The noise originates from different issues that are difficult to capture in generalized algorithms, like certain weather conditions and differences in e.g. installation configurations, data quality and measurement availability. Consequently, analysis and estimations based on real-world data in commercial systems often conceal faults and degradation, and lead to frequent false alarms when used in monitoring. From an operational point of view, false alarms are just as problematic as undetected faults, as it reduces the trust in the monitoring system.

Common approaches to handle the noise in PV system performance estimates are filtering, such as clear sky filtering or irradiance value filtering [2–5], or lowering the time resolution. Although this can be useful for some applications, information which may be necessary to do advanced fault diagnosis (e.g. detecting faults impacting the low light performance of the PV modules [6]) or day to day monitoring in areas with challenging weather conditions may be lost. Lowering the time resolution by aggregating over longer periods of time introduce unknown uncertainties and increase the reaction time of the algorithm.

In this work, we evaluate two standard performance metrics commonly used in PV system monitoring: temperature corrected performance ratio (PR_{TC}) and specific yield (Y_f) inverter comparison. This is done by testing the methods on data from a commercial PV system located in Norway, where the PV modules are exposed to diverse types of challenging weather conditions (e.g. snow, high frequency of cloudy weather), and large variations in irradiation conditions throughout the year. The evaluation is conducted by calculating the metrics and assessing the periods where there are large deviations from the expected constant values. The effect of removing the main issues identified in the evaluation of the unstable periods is compared to standard filtering approaches. To efficiently remove the main issues, a new snow detection method was developed. As discussed in our previous work [7], there is a lack of methods for robust data-based snow detection in PV systems in periods with partial melting.

The aim of the described analysis is to improve the monitoring methods for commercial PV systems. This is done by providing an understanding of the current limitations, particularly with respect to noise and applicability in climates with large variations in weather. The evaluation allows for a further assessment of how these methods can be improved, and how they eventually should be modified for different types of PV installations in different climates to work more efficiently. This lays a foundation and identify a direction for the development of improved methods and efficient filtering strategies in performance analysis and fault detection for PV systems.

2 METHODS

2.1 Dataset

The data is collected from a 135 kW_p PV system, located in the South-Eastern part of Norway (59.9 °N / 10.8 °E). The PV modules are East oriented, with an azimuth of 112° and a tilt of around 10°, and they are installed on an approximate flat roof. The roof has a tilt of 1-2° in the North-South direction, meaning half of the PV modules has the same tilt North, and the other half has the same tilt South. The module type is IBC Solar PolySol 250 CS. The PV modules are connected to eight different inverters, and the PV capacity for each inverter varies from 13 to 18 kW_p. Plane of array (POA) irradiance is measured by a crystalline silicon reference cell. The temperature of the reference cell is measured, and it is used as an estimate of the PV module temperature.

Data from September 2014 to April 2018 is used, logged with 5 minutes averages. Night time values, i.e. logged values of 0 for current or irradiance, are not included in the analysis.

2.2 Performance metrics

Two basic performance metrics commonly used in monitoring are tested: Specific yield (Y_f) inverter comparison:

$$Y_{f\ comparison} = Y_{fDC\ inverter\ x} \ / \ Y_{fDC\ median\ all\ inverters}$$

and temperature corrected performance ratio (PR_{TC}):

$$PR_{TC} = \ (Y_{fDC} / (1 + \gamma(T_{mod} - T_{STC}))) \ / \ Y_r$$

Y_f is the specific yield – the energy generated in a given time interval, divided by the rated power of the system. Y_r is the POA insolation in the same time interval divided by the reference irradiance 1000 W/m^2 [8]. γ is the material dependent maximum power temperature coefficient. For the given technology this coefficient is -0.43%/ °C. T_{mod} is the estimated PV module temperature, and T_{STC} is the reference temperature 25°C. In the specific yield comparison, the inverter energy output is compared to the median inverter energy. In this way, weather conditions are inherently accounted for, and sensor data quality is not an issue. Using the median instead of the mean reduces the influence of faulty inverters in the comparison, should there be any.

2.3 Evaluation of performance metrics

The performance metrics are tested on the dataset by calculating the parameters on an hourly basis. Hourly averaged performance parameters are commonly used to provide a balance between resolution and stability. Here it is also used to enable separation between different effects influencing the behavior of the performance metrics. The assumption is that the metrics are stable under normal operation, while changes in the performance will lead to a decrease. However, this is not always a correct assumption: In some periods the metrics are unstable, giving very varying or unexpected results that are not caused by faults. These periods are qualitatively assessed to explain the large variations.

The standard deviation (σ) of the performance metrics can be used to quantify the variation in the metric under normal operation for a given system, as discussed in our previous work [9]. With lower variation in the metrics during normal operation conditions, the performance metric has a higher sensitivity for detecting abnormal situations. The standard deviation can hence be used to measure the stability and accuracy of the performance metrics.

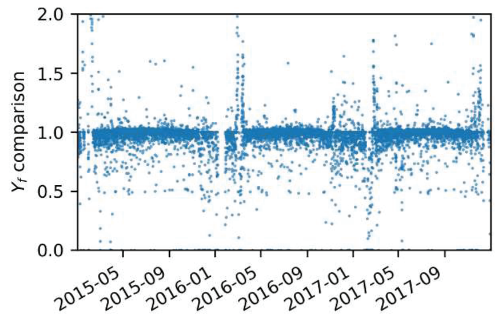
To quantify the impact of the different effects causing periods with large variation in the hourly performance metrics, the standard deviation in the metrics is calculated before and after filtering out the effects. This is compared to the change in standard deviation after applying standard filtering to the metrics. The standard filtering approaches used is low irradiance and clear sky filter. The clear sky detection algorithm described in [10] as implemented in pvlib [11] is used for clear sky filtering. The python version of pvlib is also used in the estimation of the POA clear sky irradiance used in the clear sky detection algorithm, and for the estimation of solar elevation.

To evaluate if there are any differences in irradiance conditions between the inverter strings and between the inverter strings and the irradiance sensor due to e.g. slightly different installation angles or hard shadowing, the clear sky signal was estimated for the irradiance sensor and for each string using the statistical clear sky fitting algorithm proposed by [12]. Using this algorithm, the clear sky current and irradiance for each day through the year was estimated using the measured current and irradiance data. For the inverters, the current values were used instead of the power values to focus on the irradiance signal and exclude temperature effects.

3 RESULTS AND DISCUSSION

3.1 Performance evaluation using unprocessed data

The specific yield inverter comparison and the temperature corrected performance ratio for one inverter, using unfiltered hourly data, are presented in Figure 1. The trends are similar for all the inverters. The variation in the specific yield comparison and the temperature corrected performance ratio is large, both relatively (Figure 1) and absolutely (Figure 2). The average standard deviation of the Y_f inverter comparison of the 8 inverters is 0.38. For PR_{TC} it is 0.25. These large variations in the estimation of the normal state of the PV system challenge efficient use of these performance metrics for fault detection and performance evaluation. Fault detection is normally based on detecting when a system is operating outside normal conditions, such large variations will hence produce false alarms and result in low sensitivity [9].



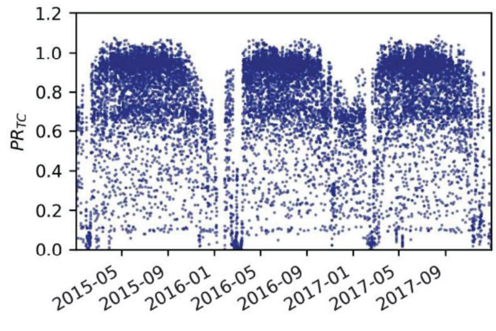


Figure 1: Variation during normal operation in Y_f comparison (top) and in PR_{TC} (bottom) using hourly data from one inverter.

3.2 Performance evaluation using standard filtering

To reduce the variation and increase the accuracy in PV performance analysis, it is common to filter out the low irradiance and/or applying a clear sky filter. In [2] a low irradiance threshold of 200 W/m^2 and a clear sky filter is proposed to remove time periods of poor or variable solar resource conditions to get a stable degradation estimate. The same irradiance threshold is also applied by [3] for fault detection, and also in this work it is observed that clear sky days have lower variation in the estimates of the current and power under normal conditions. The average standard deviation for all the inverters and the remaining data after applying the same irradiance threshold and a clear sky filter on the calculated Y_f comparison and the PR_{TC}, are given in Table I. The filtered results for the PR_{TC} are also visualized in Figure 3.

Filtering the data with the standard approach reduces the standard deviation of the data. However, the number of data points are also drastically reduced and not all large variations are removed. Adding the clear sky filter in addition to the low irradiance threshold increases the variation due to the large reduction in data points – also the ones that are stable. Hence, naïve filtering is not a global solution for all monitoring. Here the methods are both imprecise and too strict, leaving too little data to base the monitoring on.

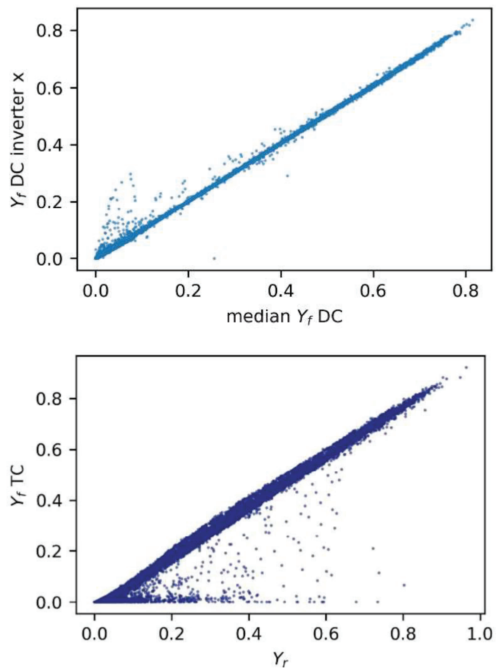


Figure 2: Absolute comparison between the inverter Y_f and the median Y_f (top) and between $Y_{f\,TC}$ and Y_r (bottom).

Table I: The average standard deviation (σ) of the two metrics for all inverters, without filters, and after consecutively removing low irradiance (< 200 W/m^2) and cloudy periods [11]. Remaining data after filtering is also given.

	Avg σ Y_f comparison	Avg σ PR_{TC}	Remaining data
Raw data	0.38	0.25	100 %
Low irradiance	0.16	0.17	38 %
Cloudy periods	0.21	0.21	6 %

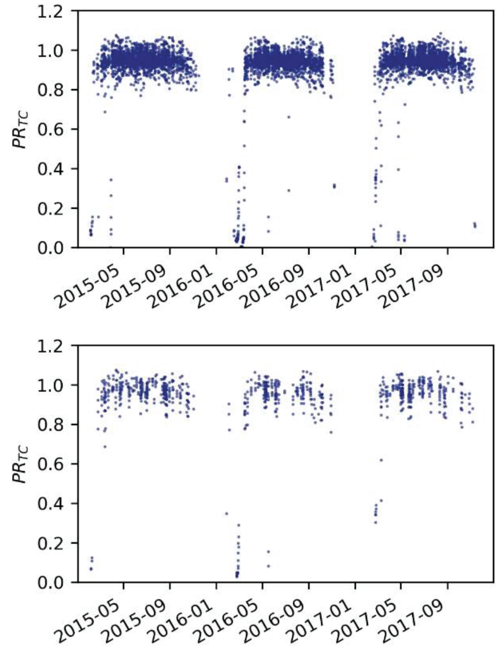


Figure 3: PR_{TC} using hourly data from one inverter, after consecutively removing (top) low irradiance (< 200 W/m^2) and (bottom) cloudy periods.

3.3 Evaluation of time periods with large variations

To better understand when the monitoring methods do not work, the time periods with large variations have been analyzed. The explanations for the largest variations can be divided into three major categories, discussed in the following subsections.

3.3.1 Snow

Snow is a well-known challenge in PV system monitoring in Northern climates. For the tested performance metrics, a full snow cover is unproblematic. With zero production, there are no variation between the inverters and consequently no variations in relative inverter performance. When the irradiance sensor is covered in snow, no low PR values will be calculated. The main challenge in PV system monitoring, is the melting period. When the snow is melting, the inverters and the irradiance sensor might receive different irradiance. Additionally, the inverters might have partial snow covers, giving signatures similar to faults.

To remove data from periods with snow covered PV modules, a new snow detection method was developed. Using local snow depth estimation from the Norwegian Water Resources and Energy Directorate [13], and power and irradiance data for the system, the variation in DC voltage for the system under normal conditions and for snow melting periods was found. In periods with partial snow cover, the DC voltage of each string has increased variation compared to normal operation, and there is larger variation between different inverters. A threshold for DC voltage variation was determined empirically. The periods with full snow cover and partial snow cover was accordingly removed based on a combination of snow depth data and the DC voltage variation limit for normal operating conditions.

3.3.2 Morning/evening effects

As expected, there were large relative variations in Y_f and PR_{TC} in the morning and evening. One of the main explanations for this is variations in low light behavior of the PV modules and inverters. Both the low irradiance and the increased share of diffuse light in the morning and evening will influence variation in PV module behavior. Additionally, small variations in PV module tilts, as discussed in depth in the next section, can lead to significant differences in the angle of incidence of the incoming light, and consequently a variation in reflected and received irradiance. By relating the Y_f and PR_{TC} to irradiance level and solar elevation, it was found that for this specific system, these effects were most prominent for irradiance values < 50 W/m^2 and solar elevation < 10°. A general algorithm for estimating the optimal filtering threshold of these values for different locations is proposed in [9].

3.3.3 Physical irregularities in the installation

Due to physical limitations in PV system installations such as variations in roof inclination, topography, objects shadowing the PV modules, different PV modules/inverter strings might receive different irradiance, resulting in different energy output. This can also affect the irradiance sensor. Also, other technical irregularities in the installation and variations in local climate can lead to variation for a PV system in e.g. temperature and soiling patterns. For this system, particularly two installation specific irregularities influence the monitoring output: the modules in one of the strings had a different tilt angle from the rest, and there was a difference in the tilt of the modules and the POA irradiance sensor. The effect of each of these aspects of the installation are explained in the following.

The variation in received irradiance on the different inverter strings are illustrated in Figure 4, using the DC current. As shown in this figure, inverter 6 has a current curve with a clearly different shape than the other inverters. This is due to the 1-2° tilt in the North and the South direction of the roof (while the PV modules are faced East). Where the rest of the inverters have PV modules that is both tilted slightly towards South and North, inverter 6 has only South tilted modules. This leads to significant variation in irradiance conditions, also on an hourly basis, between inverter 6 and the rest of the inverters and weakens the basis for comparison.

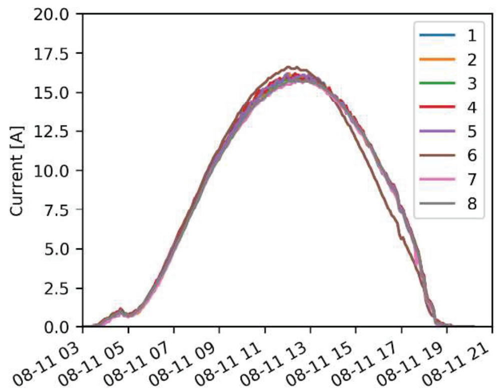


Figure 4: The DC current for each inverter during one clear day (5 minute averages), illustrating the variation in received irradiance for the inverter strings.

For the PR_{TC} values, it was observed especially high values in the morning, and very low values in the afternoon. This was found to be because of the tilt of the reference cell, which was 1-2° lower than the average tilt of PV-modules. Additionally, it had a 1-2° tilt towards South. Consequently, there are several hours the reference cell is not measuring a representative irradiance for the PV system. Difference in tilt between reference cell and the PV modules is an issue that will influence most irradiance based performance metrics.

These effects were filtered out based on deviations between the estimated clear sky behavior [12] for each inverter and between the inverters and the reference cell.

3.4 Effect of the identified issues on the performance metric variation

The effect of consecutively removing the issues identified and described in Section 3.3, are shown for each inverter in Figure 5, and for the average of all the inverters in Table II. The percentage of remaining data after removing the effects is also given in the table.

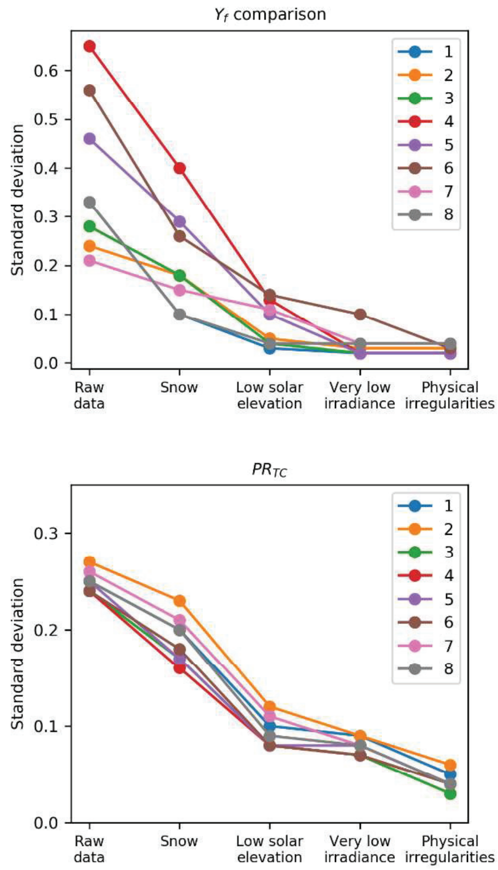


Figure 5: The standard deviation of the two metrics for each inverter, where the effects leading to unstable periods are consecutively removed.

Compared to the results of the standard filtering approach presented in Table I, the variation is significantly decreased and at the same time less data is removed. In the comparison of the specific yield, removing periods where there were large variations in incoming irradiance because of physical deviations was only relevant for inverter 6, as this is the only inverter that has significantly different installation configurations compared to the other inverters. For the PR_{TC}, removing this effect influence the variation for all the inverters because the irradiance sensor has different tilt angles than all the PV module strings.

Table II: The average standard deviation of the two metrics for all inverters, where the effects leading to unstable periods are consecutively removed. Remaining data after filtering is also given.

	Avg σ Y_f comparison	Avg σ PR_{TC}	Remaining data
Raw data	0.38	0.25	100 %
Snow	0.21	0.19	84 %
Low solar elevation	0.08	0.09	63 %
Very low irradiance	0.04	0.08	58 %
Physical irregularities	0.03	0.04	Inverter specific

4 CONCLUSIONS

The results show that naive use of standard performance metrics such as specific yield and temperature corrected performance ratio in a monitoring system for PV installations, give unreliable results with high variation in the PV system performance estimation. This will both reduce the sensitivity and the fault detection ability of the monitoring system and typically result in false alarms. Very low solar elevation and irradiance, snow and technical irregularities in the installation are the primary causes of the high variation in the monitoring output. It is shown that for certain climates standard filtering is not sufficient to solve these problems, and that site-specific filtering of data gives more stable monitoring output, entailing more data and less variation.

REFERENCES

[1] A. Triki-Lahiani, A.B.-B. Abdelghani, I. Slama-Belkhodja, Fault detection and monitoring systems for photovoltaic installations: A review, Renew. Sustain. Energy Rev. (2017). doi:https://doi.org/10.1016/j.rser.2017.09.101.

[2] D.C. Jordan, C. Deline, S.R. Kurtz, G.M. Kimball, M. Anderson, Robust PV Degradation Methodology and Application, IEEE J. Photovoltaics. (2017) 1–7. doi:10.1109/JPHOTOV.2017.2779779.

[3] S. Silvestre, L. Mora-López, S. Kichou, F. Sánchez-Pacheco, M. Dominguez-Pumar, Remote supervision and fault detection on OPC monitored PV systems, Sol. Energy. 137 (2016) 424–433. doi:10.1016/j.solener.2016.08.030.

[4] G. Belluardo, P. Ingenhoven, W. Sparber, J. Wagner, P. Weihs, D. Moser, Novel method for the improvement in the evaluation of outdoor performance loss rate in different PV technologies and comparison with two other methods, Sol. Energy. 117 (2015) 139–152. doi:10.1016/j.solener.2015.04.030.

[5] K. Kiefer, D. Dirnberger, B. Müller, W. Heydenreich, A. Kröger-Vodde, A Degradation Analysis of PV Power Plants, Proc. 25th Eur. Photovolt. Sol. Energy Conf. (2010) 5032–5037. doi:10.4229/25thEUPVSEC2010-5BV.4.26.

[6] D.C. Jordan, S.R. Kurtz, The dark horse of evaluating long-term field performance-Data filtering, IEEE J. Photovoltaics. 4 (2014) 317–323. doi:10.1109/JPHOTOV.2013.2282741.

[7] M.B. Øgaard, H. Haug, J. Selj, Methods for Quality Control of Monitoring Data from Commercial PV Systems, in: Eur. Photovolt. Sol. Energy Conf. Exhib. METHODS, 2018: pp. 2083–2088. doi:10.1093/annonc/mdy039/4835470.

[8] B. Marion, J. Adelstein, K. Boyle, H. Hayden, B. Hammond, T. Fletcher, D. Narang, A. Kimber, L. Mitchell, S. Richter, Performance parameters for grid-connected PV systems, Photovolt. Spec. Conf. 2005. Conf. Rec. Thirty-First IEEE. 31 (2005) 1601–1606. doi:10.1109/PVSC.2005.1488451.

[9] Å. Skomedal, M.B. Øgaard, J. Selj, H. Haug, E.S. Marstein, General, robust, and scalable methods for string level monitoring in utility scale PV systems, 36th Eur. Photovolt. Sol. Energy Conf. Exhib. (2019).

[10] M.J. Reno, C.W. Hansen, Identification of periods of clear sky irradiance in time series of GHI measurements, Renew. Energy. 90 (2016) 520–531. doi:10.1016/j.renene.2015.12.031.

[11] W. F. Holmgren, C. W. Hansen, M. A. Mikofski, pvlib python: a python package for modeling solar energy systems, J. Open Source Softw. (2018). doi:10.21105/joss.00884.

[12] B. Meyers, M. Tabone, E.C. Kara, Statistical Clear Sky Fitting Algorithm, Conf. Rec. IEEE Photovolt. Spec. Conf. (2018) 1–6. http://www.wcpec7.org/eWCPEC/manuscripts/MeWCPEC930_0521112519.pdf.

[13] NVE, seNorge, (n.d.). www.senorge.no (accessed September 1, 2019).

III

Paper III

Photovoltaic system monitoring for high latitude locations

M.B. Øgaard, H.N. Riise, H. Haug, S. Sartori, and J.H. Selj

In: *Solar Energy,* Vol. 207 (2020), pp. 1045-1054.
DOI: 10.1016/j.solener.2020.07.043

Solar Energy 207 (2020) 1045–1054

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Photovoltaic system monitoring for high latitude locations

Mari B. Øgaard[a,b,*], Heine Nygard Riise[b], Halvard Haug[a,b], Sabrina Sartori[a], Josefine H. Selj[a,b]

[a] Department of Technology Systems, University of Oslo, Gunnar Randers vei 19, 2007 Kjeller, Norway
[b] Renewable Energy Systems Department, Institute for Energy Technology, Instituttveien 18, 2007 Kjeller, Norway

ARTICLE INFO

Keywords:
Photovoltaic systems
Monitoring
Filtering
Performance metric testing
Machine learning
High latitude climates

ABSTRACT

Reliable monitoring of PV systems is essential to establish efficient maintenance routines that minimize the levelized cost of electricity. The existing solutions for affordable monitoring of commercial PV systems are however inadequate for climates where snow and highly varying weather result in unstable performance metrics. The aim of this work is to decrease this instability to enable more reliable monitoring solutions for PV systems installed in these climates.

Different performance metrics have been tested on Norwegian installations with a total installed capacity of 3.3 MW: (i) comparison of specific yield, (ii) temperature corrected performance ratio, and (iii) power performance index based on both physical modelling and machine learning. The most influential effects leading to instability are identified as snow, low light, curtailment, and systematic irradiance differences over the system. The standard deviation of all the performance metrics is reduced when filters targeting these four effects are applied. Compared to general low irradiance or clear sky filtering, a greater reduction in the variation of the metrics is achieved, and more data remains in the useful dataset. The most suitable performance metrics are comparison of specific yield and performance index based on machine learning modelling.

The analysis highlights two paths to accomplish increased reliability of PV monitoring systems without increased hardware costs. First, better reliability can be achieved by selecting a suitable performance metric. Second, the variability of the performance metric can be reduced by utilizing filters that specifically target the origin of the variability instead of using standard literature thresholds.

1. Introduction

1.1. PV system monitoring

With recent years' increased focus on operation and maintenance of photovoltaic (PV) systems related to its more important role in cost reduction (Klise et al., 2014; Whaley, 2016), numerous algorithms and performance metrics have been proposed to improve monitoring of PV installations (Daliento et al., 2017; Livera et al., 2019; Triki-Lahiani et al., 2018). The aim of these algorithms is to detect periods when the PV system is deviating from normal operation and identify faults. The existing solutions for affordable monitoring of commercial PV systems are however often inadequate for high latitude climates, as snow and highly varying weather result in unstable performance metrics.

PV system monitoring is typically based on a comparison between the production data directly acquired from the inverter and a yield target (Daliento et al., 2017). Examples of this is yield comparison of similar units (Skomedal et al., 2019), performance ratio (*PR*) with or without temperature correction (Dierauf et al., 2013; IEC, 2017; Woyte et al., 2014), and comparison with physical or data driven models of the system (Daliento et al., 2017). Due to the many parameters influencing PV energy generation (Fouad et al., 2017) which challenge accurate physical modelling, and the increasing amount of acquired data, machine learning have gained increased attention in PV system monitoring research in recent years (Daliento et al., 2017; Rodrigues et al., 2017; Triki-Lahiani et al., 2018).

Competitive solutions for automated monitoring of PV systems must have high sensitivity and fast detection, and at the same time minimize false alarms. This demands frequent and accurate performance estimation. Exact performance estimation is however challenging. Certain weather and irradiance conditions, seasonal soiling, shading, early system degradation, clipping or intentional curtailment, problems with data quality or lack of measurement availability (Jordan et al., 2017; Kurtz et al., 2013) can lead to errors or noise in the performance assessment. Noise will reduce the useful information that may be extracted from monitoring of commercial PV systems. It may lead to false alarms or conceal faults because the sensitivity of the fault detection is reduced. This causes challenges for implementation of generalized

* Corresponding author at: Department of Technology Systems, University of Oslo, Gunnar Randers vei 19, 2007 Kjeller, Norway.
E-mail address: mari.ogaard@its.uio.no (M.B. Øgaard).

https://doi.org/10.1016/j.solener.2020.07.043
Received 15 May 2020; Received in revised form 26 June 2020; Accepted 13 July 2020

monitoring algorithms for commercial PV systems without extensive documentation and/or comprehensive instrumentation. This challenge is even greater for high latitude locations where seasonal soiling and variable irradiance conditions are especially prominent. Due to the drastic cost reductions for PV installations in recent years, the PV installation rate is rapidly increasing also in continental and subarctic climate zones. Consequently, the need for reliable monitoring solutions for these conditions is also increasing. For development of monitoring algorithms with high performance under difficult conditions, it is necessary to study the effect of data quality, snow cover, low light conditions and applied filtering more closely and subsequently investigate how the various monitoring methods need to be adjusted to different installation and weather conditions.

1.2. Data quality

In the PV monitoring standard IEC 61724-1, it is recommended to check for unphysical and missing values, and compare similar measurements to detect and remove erroneous PV system data caused by poor data quality (IEC, 2017). In the literature, more advanced methods are suggested, like assessment of expected relationships between measurements. Examples are calculations of the nominal operating cell temperature (NOCT) to evaluate irradiance and temperature measurements (Ransome, 2008) and relating temperature measurements to irradiance to detect detachment of module temperature sensors (Woyte et al., 2014). Øgaard et al. (2018) suggest comparison of irradiance measurements to modelled clear sky irradiance.

Dong et al. (2017) suggest to evaluate if fault characteristics are influenced by external factors, such as solar position, to detect variation caused by shading. For a location with a high share of clear days and direct light, this approach could be used to identify both deviations caused by shading and irradiance variations caused by differences in module tilt angles due to topography. For locations with larger share of diffuse light and cloudy weather, the correlation of these deviations with angle of incidence will be less clear.

1.3. Snow cover

Snow cover is a common seasonal soiling issue in high latitude locations, with high impact on PV system data. Handling snow covers in a monitoring system is important, as a full snow cover looks like an inverter breakdown, whereas a partial snow cover, leading to partial shading, can give power losses and changes in the maximum power point (Belhachat and Larbes, 2015) similar to serious PV module failures (Tsanakas et al., 2016). Snow cover is easy to identify by visual inspection, but it is not easily detected by an automated monitoring system without extra sensors. When analyzing historical data, snow cover is usually easily detected, as it can be characterized as transitory periods in the wintertime with no power production or low efficiency. In real time automated monitoring, it is challenging to predict the onset, duration and shading effect of the snow cover. To our knowledge, there are no reliable algorithm for detection of full and partial snow covers for real time monitoring presented in the literature.

1.4. Low light conditions

It is well known that low light conditions lead to noise in PV system analysis (Belluardo et al., 2015; Jordan et al., 2017; Reich et al., 2012). Typically, calculations of performance metrics result in high levels of noise in the morning and evening. Both the module and the inverter efficiencies are unstable under these conditions, as they are very sensitive to changes at low irradiance and power, and reflection losses are highly sensitive to changes in angle at high angle of incidence. Additionally, at lower irradiance levels a higher variation in module efficiencies could be anticipated, due to a potential variation in cell shunt resistance, as shown by Grunow et al. (2004). As discussed by Louwen et al. (2017), higher air mass, increased amount of diffuse irradiance and resulting spectral variation also impact efficiency, and hence increased instability at low light conditions. Reduced accuracy in irradiance measurements (Reise and Müller, 2018) and inverter measurements at lower irradiance/power levels, could give additional contributions to noise

1.5. Filtering

To reduce the uncertainty caused by noise at low irradiance conditions, filtering based on irradiance and clearness level is commonly used (Belluardo et al., 2015; Camus et al., 2018; Jordan et al., 2017; Reich et al., 2012; Silvestre et al., 2016). Moser et al. (2014) additionally filter out periods with high wind speed to ensure uniform temperature conditions, to further reduce uncertainties. To remove periods with clipping, Jordan et al. (2017) recommend to remove data where the power is $>$ 99% of the maximum, and Meftah et al. (2019) suggest an upper irradiance limit.

Optimal filtering thresholds will depend on the system technology and the purpose of the analysis. Although methods to find optimal filtering thresholds to reduce noise in different datasets have been suggested, (Jordan and Kurtz, 2014; Skomedal et al., 2019) there are no standardized methods for data handling and filtering in PV system analysis. This is unfortunate, as recent studies on PV system performance loss (Curran et al., 2019) and degradation (Jordan et al., 2019; Jordan and Kurtz, 2014), show that different filtering and data handling methods can lead to differences in the loss estimate of more than one percentage point. This emphasizes a strong need for standardized methods. The effect of the applied filters will also depend on the specific dataset and operating environment of the PV system. In high latitude locations, typical filtering approaches like irradiance thresholds and clear sky filters can remove too much data rendering day-to-day monitoring difficult, while still not adequately reducing the noise (Øgaard et al., 2019).

1.6. Aim and approach

To identify areas of improvement of PV monitoring solutions in high latitude climates, we have tested four different PV system monitoring approaches (Y_f comparison, PR'_{STC}, *PPI* with physical and machine learning based modelling) using data from six commercial PV systems in Norway. The locations of the systems span various climate zones (Beck et al., 2018), but they are generally exposed to highly variable weather including long periods with low light conditions and snow. It is well known that snow (Andrews and Pearce, 2012; Marion et al., 2013) and low light conditions (Westbrook et al., 2012) constitute a challenge for accurate modelling of the PV energy generation. In the presented work, the challenges of PV system monitoring at high latitude locations have been evaluated, and the effect of applying tailored filters to remove specific conditions that generate noise is studied and compared to standard, more general filters used in PV monitoring. To enable detection of shading and irradiance differences caused by topography variations in locations with few clear sky days, we propose to use the statistical clear sky fitting algorithm suggested by Meyers et al. (2018). Based on this, improved methods for monitoring of PV systems in climates with highly variable weather and irradiance conditions is suggested.

2. Methodology

2.1. Dataset

The data in this study are taken from commercial PV systems on approximately flat roofed buildings, representing many of the larger systems in the Nordic countries. The modules are mounted with a tilt angle of $\sim 10^{\circ}$ in an east/west configuration. A such low tilt angle is not optimal for high latitude locations with respect to maximum production, but is typically used in installations on flat roof systems also in these types of locations to maximize roof coverage and to achieve a more even production distribution through the year. The exact orientation depends on the orientation of the building. Production data is collected from the inverters, and temperature and irradiance data are collected from the monitoring system of the installations. The sampling interval of the data logging is 5 min, and DC inverter data measured at maximum power point tracker level (MPPT) is used to increase granularity. The effective irradiance incident on the PV modules, i.e. the irradiance the modules can utilize (King et al., 2004; Stein and Farnung, 2017), is measured with a crystalline silicon reference cell in the plane of the PV modules. The irradiance measurements are controlled for shifts and degradation by comparison to modelled clear sky irradiance and the statistical clear sky method described in Section 2.3.1. Measurements of the reference cell temperature is used as an approximation for solar cell temperature. The cell temperature measurements are validated against module back sheet temperature measurements and thermography of the modules. The measurement uncertainty of the irradiance is ± 5 W/m^2 ± 2.5% of measured value, and for the temperature it is 1 K. The inverters register curtailment and clipping events. Snow depth data is collected for each location from seNorge.no (NVE, 2019).

Relevant technical and geographical information about the systems are given in Table 1. For each location, the typical meteorological year (TMY) irradiance and ambient temperature is presented in Fig. 1a and b, respectively. System 4 and the west oriented part of systems 1a and 1b lack temperature measurements and are excluded when temperature measurements are required to estimate performance metrics. As typical for systems at these coordinates, the yearly global horizontal irradiation is below 1000 kWh/m^2, most of the energy is generated at temperatures around 10–20 °C, and the total share of diffuse light can reach 50% (Imenes and Selj, 2017).

2.2. Performance metrics

Specific yield, performance ratio and power performance index are well known metrics to evaluate the performance of a PV system. The power performance index is the ratio of measured to modelled power, and both physical and statistical models can be used. In this work, the physical PVWatts model, the single diode model, and as an example of a machine learning model, a commonly used random forest regressor (Pedregosa et al., 2011) is selected.

To identify the conditions and effects which lead to the highest level of noise in the performance metrics, environmental and inverter parameters in the "noisy" time periods are evaluated. Sequentially, the identified noise-generating effects are removed from the dataset, and the procedure is repeated. Unphysical values are utilized to identify effects leading to bias in the performance metrics. Finally, a set of recommended filtering parameters is found. The effect of the suggested filters is quantified by the standard deviation of the metrics and by the percentage of remaining data points and energy. This will be compared to more general approaches: irradiance thresholding and clearness filtering. The filtering values from Jordan et al. (2017) (apparent clear sky conditions and irradiance threshold of 200 W/m^2) is used.

2.2.1. Yield comparison

The specific yield (Y_f) is the energy generated over a given time interval, divided by the rated power of the system (Woyte et al., 2014). For systems with multiple power measurements of units with identical configuration, comparing the specific yield of the different units allows for a monitoring system where weather conditions are inherently accounted for. The specific yield for each unit (MPPT measurement) in the system, $Y_{f\,DC}$, is compared to the median specific yield of all units, $\widetilde{Y}_{fDC}$:

$$Y_{f\,rel} = Y_{\,fDC}/\widetilde{Y}_{fDC}. \quad (1)$$

Using the median instead of the mean reduces the influence of outliers (*i.e.* anomalous MPPTs) in the comparison, reducing the risk of a biased result.

2.2.2. Temperature corrected performance ratio

To evaluate the yield for one unit over time, or to compare the yield of systems in different locations, the performance ratio corrected to standard test conditions (*STC*) temperature (PR'_{STC}) can be used (Daliento et al., 2017). PR'_{STC} is defined as the specific yield normalized to irradiance and temperature at STC (IEC, 2017), and is given by:

$$PR'_{STC} = (Y_{fDC}/(1+\gamma(T_{cell}-T_{STC})))/(G_{POA}/G_{ref}) \quad (2)$$

Here, G_{POA} is the measured plane of array irradiation in the same time interval as the specific yield, G_{ref} the reference irradiance 1000 W/m^2, T_{cell} is the estimated PV module temperature, T_{STC} is the reference temperature of 25 °C, and γ is the material dependent module power temperature coefficient. For the module technologies used at the five sites studied (different generations of mono- and multi-crystalline silicon) γ varies from −0.423 to −0.43%/°C.

2.2.3. Power performance index

The power performance index (*PPI*) is a comparison between expected and measured power (IEC, 2017):

$$PPI = P_{DCmeasured}/P_{DCexpected} \quad (3)$$

The expected DC power output is simulated using both physical modelling and machine learning based modelling.

For the physical modelling, the single diode and the PVWatts models are used. The single diode model (Corkish et al., 2013) is commonly employed in PV modelling for monitoring (Daliento et al., 2017). The PVWatts DC power model (Dobos, 2014) is a simpler model with fewer module specific inputs. Similarly to PR'_{STC}, it uses G_{POA}, T_{cell}, γ and the nominal capacity (P_{DC0}) to model the expected DC power output:

$$P_{DC} = (G_{POA}/G_{ref}) \times P_{DC0}(1+\gamma(T_{cell}-T_{STC})) \quad (4)$$

For the single diode model, module datasheet values is used as input to the System Advisor Model (SAM) (Blair et al., 2011) to estimate the diode ideality factor, light generated current, dark reverse saturation current, shunt resistance and series resistance at reference conditions, and the parameter for adjustment to temperature coefficient for short circuit current. These parameters are used as inputs to the CEC model described by Dobos (2012), together with the measured effective irradiance and cell temperature to estimate the photocurrent, saturation current, shunt resistance and thermal cell voltage under the different

Table 1
Detailed information of the various PV plants investigated in this study. (#MPPTs = number of MPPTs in each direction).

Plant	Region	Coordinates [°N, °E]	Altitude [m a.s.l.]	Azimuth angles[°]	Installed capacity [kW]	# MPPTs	Time period
1a	Eastern Norway	59.59, 10.74	80	112/292	371	11	09-2014 – 09-2019
1b	Eastern Norway	59.59, 10.74	80	112/292	222	7	11-2016 – 09-2019
2	Eastern Norway	59.94, 10.87	126	128/308	471	15	12-2016 – 09-2019
3	Central Norway, inland	60.89, 10.92	158	122/302	421	12	06-2017 – 09-2019
4	Central Norway, coast	63.34, 10.37	154	111/291	931	7	07-2017 – 09-2019
5	Western Norway	60.40, 5.47	62	105/285	886	28	09-2017 – 09-2019

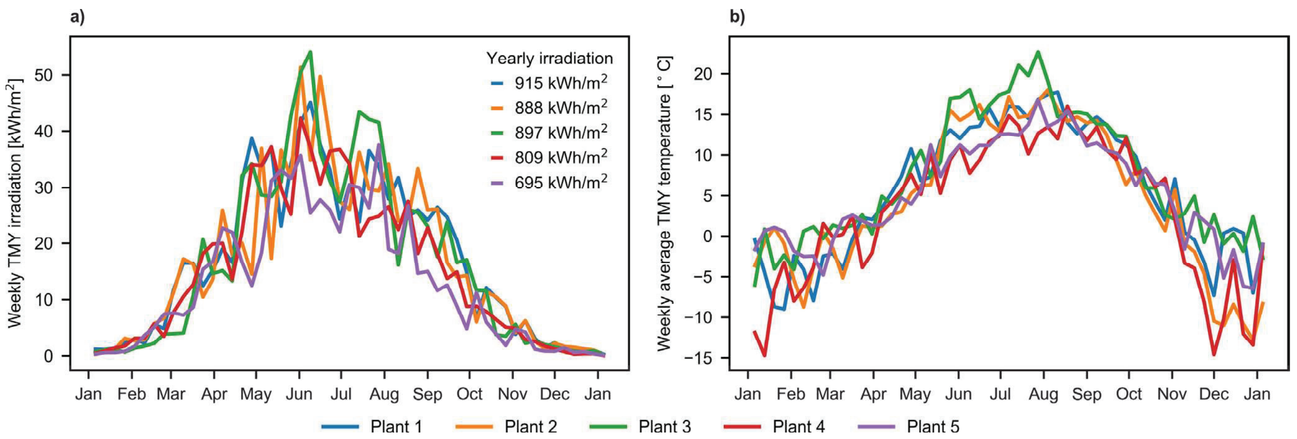


Fig. 1. Yearly variation in a) weekly global horizontal TMY irradiation, and b) weekly average TMY ambient temperature. TMY data from PVGIS (Huld et al., 2012).

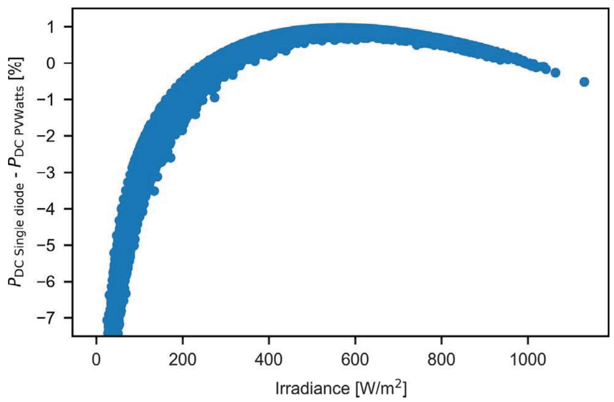


Fig. 2. The irradiance dependency of the difference between the modeled power using the single diode model and the PVWatts model.

measured conditions. The expected power output for each module is estimated by solving the single diode equation based on the parameters estimated with the CEC model, as implemented in pvlib (F. Holmgren et al., 2018). An important difference between these models is that the single diode model includes the effect of the irradiance intensity on the PV efficiency. Fig. 2 reports the values of the difference in modeled power using the two models against irradiance. The single diode model estimates a lower power value at low irradiance, and up to 1% higher at irradiance levels around 500 W/m^2. This is in agreement with the results presented by Dobos et al. (2019). The total constant losses (L_{total}) for the physical models are estimated using the PVWatts system loss model (Dobos, 2014):

$$L_{\text{total}}(\%) = 100(1 - \Pi_i(1 - L_i) \tag{5}$$

where L_i is the contribution from the individual loss mechanism (i) in percent. The modelled results are used for instant comparison with measured power values, and varying losses such as soiling, shading, snow and availability are therefore not included. The following loss mechanism values are used: Wiring = 2%, connections = 0.5% and light induced degradation = 1.5%.

For the machine learning based model, a Random Forest (RF) regressor based on the sklearn library in Python (Pedregosa et al., 2011) is trained on historical irradiance and, when available, solar cell temperature data with power output per MPPT as model targets. According to Rodrigues et al. (2017), irradiance and temperature are the most common condition describing input features used in PV system modeling with machine learning. Only times with logged production data are used, and the data is split into a training set and a test set (80% training set, 20% test set). The regressor performance is evaluated on the test set using normalized root mean square error (*nRMSE*) as a metric:

$$nRMSE = \sqrt{\sum\nolimits_{t=1}^{T} (\hat{y}_t - y_t)^2} \Big/ \bar{y}. \tag{6}$$

2.3. Methods used in filtering

2.3.1. Clear sky modelling and detection

The clear sky detection algorithm described by Reno and Hansen (2016) as implemented in pvlib python (F. Holmgren et al., 2018) is used for clear sky filtering. The pvlib python library is also used in the estimation of the POA clear sky irradiance used in the clear sky detection algorithm, and for the estimation of solar elevation, using the system configuration data as input. The clear sky curves for the irradiance sensors and each inverter string are also estimated with a more empirical approach: using the statistical clear sky fitting algorithm proposed by Meyers et al. (2018). With this algorithm, the clear sky current and irradiance for each day through the year are estimated based on the measured current and irradiance data. For the inverters, the current values are used instead of the power values in order to focus on the irradiance signal and exclude temperature effects. The fitted clear sky curves are used to detect systematic irradiance differences between the different inverters and the irradiance sensor. These differences are quantified by the ratio of the clear sky curve of the inverter to the median inverter clear sky curve or to the scaled irradiance curve. Additionally, the degradation application (Meyers et al., 2019) of the statistical clear sky fitting algorithm was used to control the irradiance measurements for drift.

2.3.2. Snow filtering

External snow depth measurements, often available from local weather measurements, are used for filtering out periods with snow on the modules. To ensure that periods with partial snow covers are removed from the data, all periods with snow on the ground (snow depth > 0 m) are removed. This removes more data than necessary as PV modules typically get snow free before the ground, but noise and potential false alarms are significantly reduced.

3. Results and discussion

3.1. Identification of conditions and effects leading to noise in performance evaluation

Based on the analysis procedure described in Section 2.2., effects and conditions leading to noise, bias or difficulty in performance metric implementation are identified. The filters and filtering thresholds used to remove the different effects, and their impact on the dataset are presented and discussed in Section 3.2. The effects are categorized in the following categories: invalid data, data quality and availability, and unstable conditions. The invalid data category contains data where performance metric calculations yield illegitimate results. Effects leading to bias or systematic errors in the performance metrics are placed in the data quality and availability category. Noise generating situations are classified as unstable conditions.

3.1.1. Invalid data

The output considered as invalid in the performance evaluation, are zero output and an output wrongfully suggesting failures. Zero output is typically caused by downtime in monitoring measurements or communication. The main conditions giving false failures signatures in the production data, are snow and inverter induced power reduction. The power reduction is caused by both curtailment and inverter power clipping.

For monitoring algorithms based on yield comparisons on inverter/string level, full snow cover is not problematic as this will give the same behavior for all inverters/strings. Partial snow covers, however, may yield large relative differences due to (random) partial shading of the system, often with hard and uneven shading. These snow covers are difficult to predict, since they are influenced by several parameters and snow melting is stochastic by nature. This might be even more relevant for irradiance-based monitoring metrics: When the irradiance sensor is experiencing the same snow cover as the PV modules, the monitoring system will not report anomalous behavior. If the irradiance sensor is not covered, the resulting deviation between expected and actual production will lead to alarms. The PR'_{STC} values from one of the systems in a period of snow melting are presented in Fig. 3, showing the large variations partial snow covers may inflict, especially between irradiance sensor and modules. On 21 March, both the PR'_{STC} and the irradiance are zero due to full snow cover. When the snow melts, the PR'_{STC} values increase until normal operation is reached, while the difference between the inverters only is large in the periods with substantial melting.

For performance evaluation based on machine learning, snow and curtailment are especially challenging. If not removed from the training data, these effects will perturb the correlations between irradiance, temperature and production. When the model attempts to accommodate these perturbations, the model might lose nuances at normal operation, potentially reducing the accuracy of the machine learning algorithm also in periods without snow or curtailment. As reported in Table 2, the *nRMSE* is indeed reduced when time periods with snow and curtailment are removed from the training and testing data set.

3.1.2. Data quality and availability

We find that systematic differences in irradiance level between different inverter strings, or between inverter strings and irradiance sensor, lead to solar position dependent errors for all the systems. This is especially prominent under clear conditions. The differences in irradiance are caused by shading and different tilt angles of the PV modules and/or the irradiance sensor. Such local and varying differences in irradiance affect the basis for accurate comparison for specific time periods during clear days. Both the variation in tilt of the modules and the sensor vary mainly with the roof topography. Minor errors in the installation of the irradiance sensor could also give similar results. Fig. 4 shows how such local differences in irradiance impact the DC current of certain strings over a time period of one day. "System 1a" is a system where the modules in the strings of one the inverters have a slightly different tilt than the other strings. "System 1b" is a system with inhomogeneous shading conditions, where one inverter is experiencing more shading than the other inverters. For both systems, inverter 1 is an example of a normal inverter, and inverter 2 is the deviating inverter. The relative DC current of the inverters is also shown in Fig. 4. For the deviating inverters in the two systems, a difference in the absolute and relative current compared to the other inverters depending on time of the day is clearly visible. This reflects the difference in irradiance of the inverters.

The estimated clear sky current signal based on the statistical clear sky algorithm for the inverters is plotted together with the measured clear sky current in Fig. 4. We see that the measured and modelled values overlap well, except for the beginning and the end of the day. Hence, for most of the day, we can use these estimated curves to identify time periods where there are systematic irradiance differences between the modules. Fig. 5 shows that even with large share of diffuse conditions, like at the tested systems, we get deviations through the year because of irradiance differences. The figure shows a boxplot with weekly values of relative current for one year for two normal inverters (top) and the two deviating inverters (bottom) in system 1a and 1b. Even though the weekly median relative current values of the deviating inverters are close to one in both cases, we see that the variation in the values is much larger than for the normal inverters. This implies that the deviations we see in Fig. 4 are present through the year.

Comparison with estimated clear sky signal also enables identification of systematic irradiance differences between the inverter strings

Table 2
The nRMSE of the machine learning model for the six systems for three different training datasets: raw data, dataset filtered for snow, and dataset filtered for snow and curtailment.

nRMSE [%]	1a	1b	2	3	4	5
Raw data	12.8	15.1	21.8	44.2	26.6	24.3
Snow filter	8.8	12.5	18.9	33.7	20.5	16.7
Snow + curtailment filter	8.8	12.5	10.9	13.7	18.7	16.3

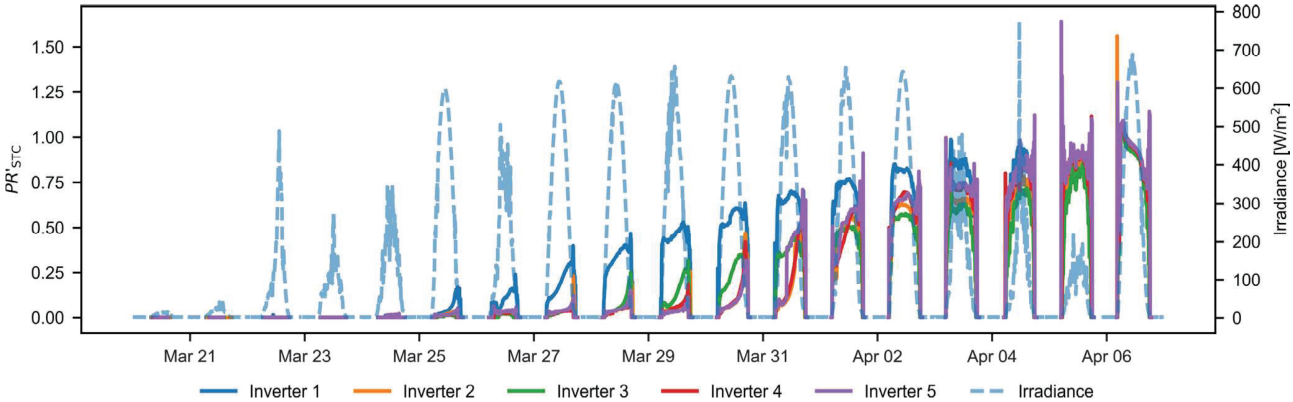


Fig. 3. Measured irradiance and PR'_{STC} for different inverters for a system where the snow cover on the modules is melting during a time period of 18 days.

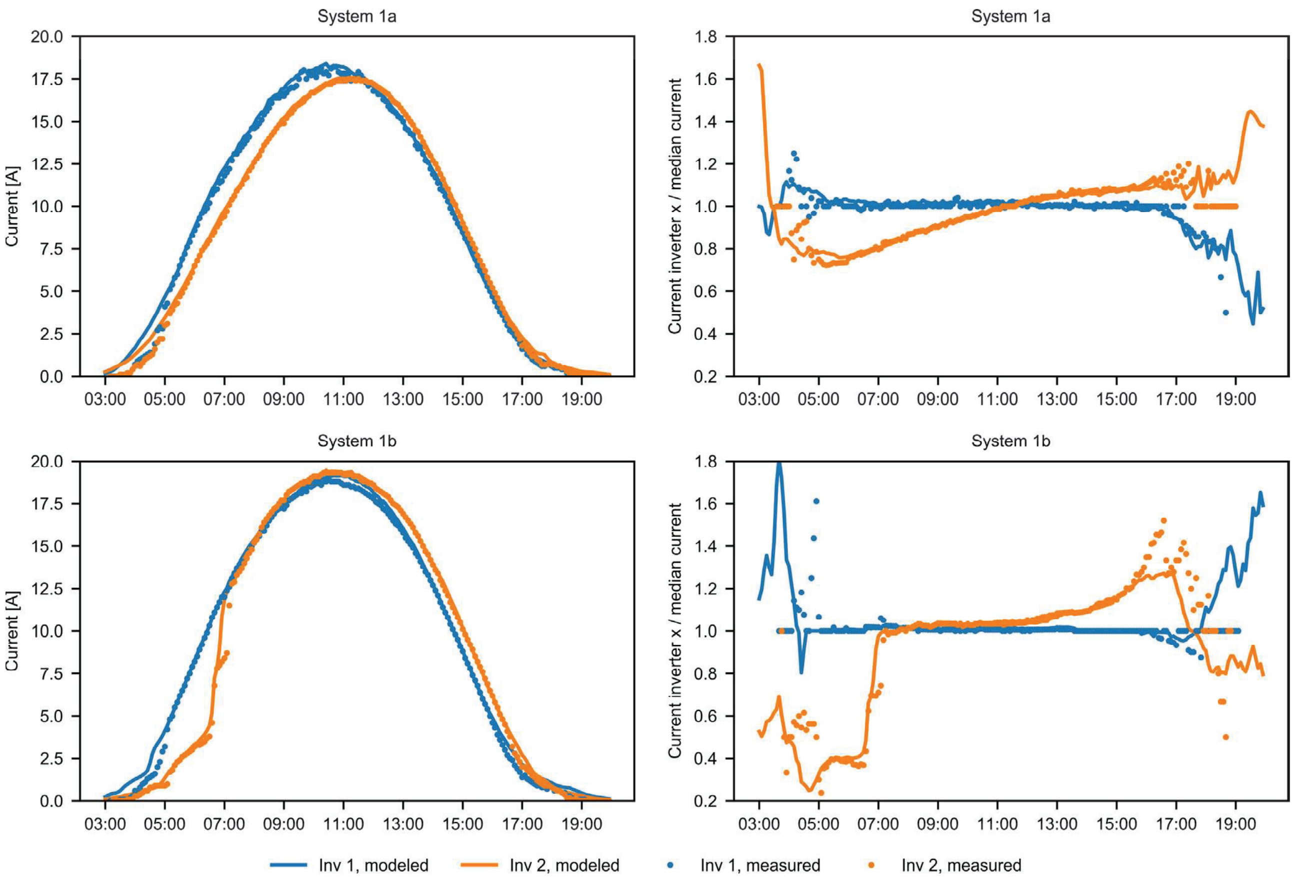


Fig. 4. Absolute (left) and relative (right) clear sky current values for one day, representing the clear sky irradiance for different inverter strings, for a system where one string has modules with a slightly different tilt than the rest of the strings (1a), and a system where one string is differently shaded than the other two (1b). Modeled (straight line) and measured (dots) values.

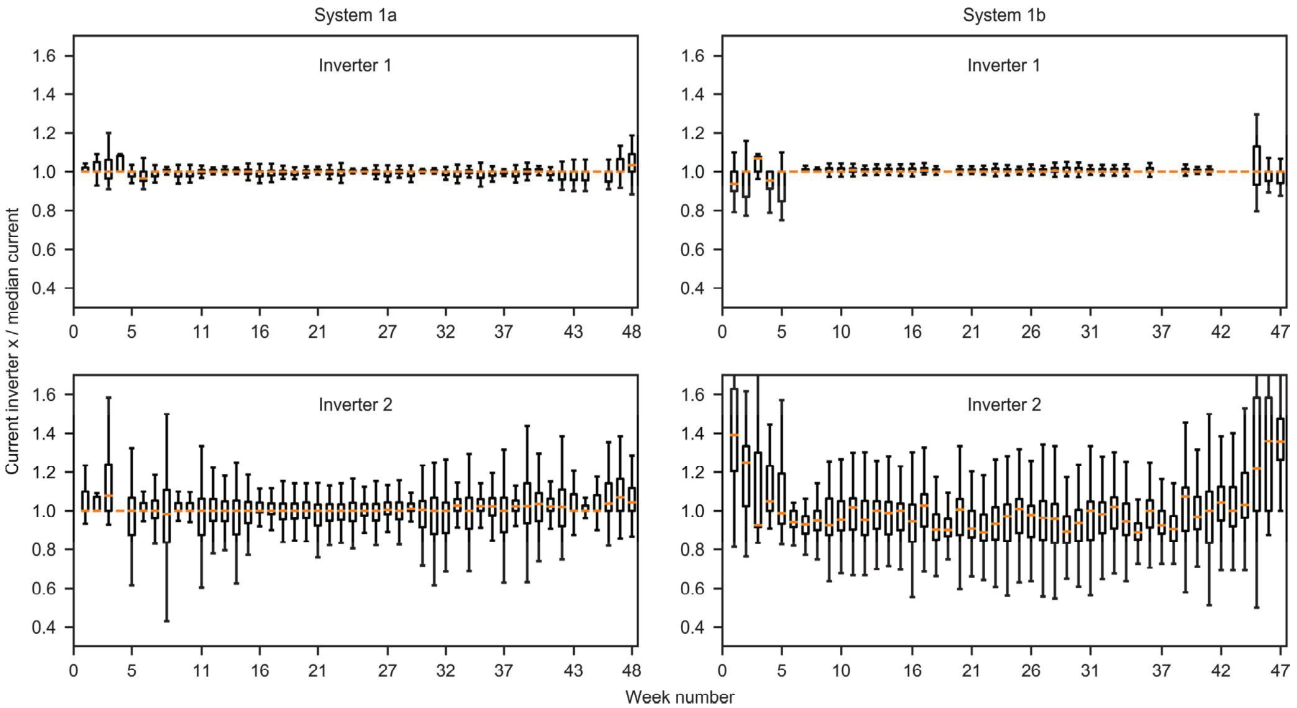


Fig. 5. Boxplot of weekly relative current values for one year for two normal inverters in system 1a and 1b (top) compared to the two deviating inverters in the same systems (bottom). The box extends from the first to the third quartile values of the data, with a line on the median. The whiskers extend to maximum 1.5 multiplied the interquartile range, and outliers are not included.

and the irradiance sensors, which give unstable results for the *PR'*$_{STC}$ and the *PPI* based on physical modelling. *PR'*$_{STC}$ and *PPI* are typically also impacted by other irradiance data sensor issues, like shifts, soiling and degradation. When the degradation application of the statistical clear sky algorithm was applied to the longer time series in the tested datasets however, no significant degradation in the irradiance sensors was detected in these cases. Temperature data quality issues may also have negative impact on these metrics. A general example of this, is that temperature measurements typically are point measurement and may not be representative for the whole system. Physical system modelling is additionally influenced by availability of system data challenging accurate estimations of balance of systems (BOS) losses, i.e. constant losses in cables and connectors, or the production dependent losses of the inverter. These losses vary between installations due to different configurations. The machine learning model automatically takes these losses into account, and is to a larger degree able to handle the effect of non-representative temperature measurements. Degradation in the irradiance sensor will however also impact the quality of the machine learning based model. Immunity to sensor data quality issues are one of the strengths of monitoring systems based on comparison of specific yield.

3.1.3. Unstable periods

As previously mentioned, it is well known that low light conditions lead to noise in PV system analysis. The resulting noise in the morning and evening is normally filtered out by removing low irradiance periods. The noise in this period is however not only due to low irradiance: high angles of incidence may cause additional instability in this period. Fig. 6, shows how high (> 0.1, noisy data) and low (< 0.1) standard deviation in the specific yield comparisons of system 1a correlate to irradiance and solar elevation angle. We use solar elevation as an indicator of angle of incidence instead of using the actual angle of incidence. Calculating the angle of incidence explicitly, and using it as a basis for filtering, requires accurate definitions in the system configuration input data, specifically tilt and azimuth angles. The histograms show all the data except snow periods, with the scatterplot showing a randomly selected sample of these data. It is clear that a typical low irradiance filter (< 200 W/m^2) would remove data points that do not have particularly high standard deviation. Replacing in this case the 200 W/m^2 low irradiance filter with a 50 W/m^2 filter *and* a solar elevation angle threshold of 20° will remove most of the noisy data, while a larger set of useful data remains. Therefore, a more detailed filtering procedure, specifically removing the conditions leading to instability, can lead to an increase in data points without increasing the noise. The trend is representative for all the tested systems, although there are some differences in the optimal filtering thresholds. The differences can

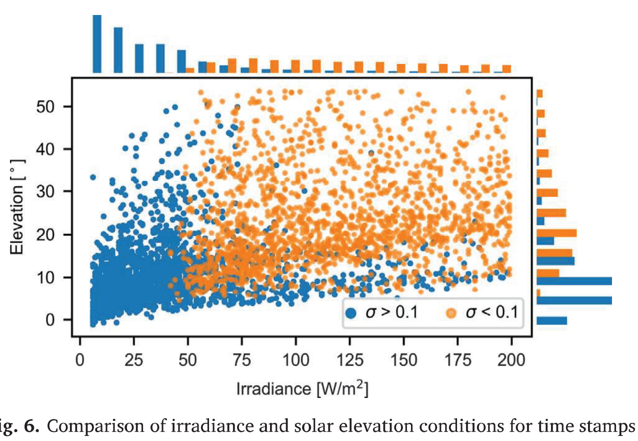


Fig. 6. Comparison of irradiance and solar elevation conditions for time stamps where there is a standard deviation in the yield comparison in system 1a of > 0.1 and < 0.1. All data are shown in the histograms, a randomly selected sample of the data is shown in the scatterplot.

Table 3
Overview of the different identified noise generating effects and their influence on different performance metrics.

	Challenge	Yield comparison	*PR'*$_{STC}$	*PPI*: Physical modelling	*PPI*: Statistical modelling
Invalid data	Snow	Noise if partial snow cover	Low/zero values, potentially over longer time periods (weeks)		
	Clipping, Curtailment	Noise if not equal for all inverters	Low values, potentially for multiple hours a day		
Data quality and availability	Constant system losses	Inaccurate estimations if not equal for all inverters		Estimation for every system necessary	–
	Systematic differences in irradiance	Inaccurate estimation if differences between inverters	Inaccurate estimation if differences between inverters and irradiance sensor		–
	Degradation in irradiance data	–	Bias/shifts depending on degradation type		
Unstable periods	Low light conditions	Noise			
	Rapid, large irradiance changes	Noise, short time period			

be partially explained by inverter sizing. The inverter efficiency is not directly related to irradiance, but with power and inverter capacity ratio, hence, undersized inverters would reach a stable efficiency at lower irradiance conditions than oversized inverters.

Other types of conditions observed to give unstable performance metrics in the datasets, are fast and large changes in irradiance transitions caused by moving clouds. With rapidly changing irradiance, the efficiency of the maximum power point tracker (MPPT) may decrease (Sanchis et al., 2007), or the moving clouds may lead to uneven shading and mismatch in the system (Lappalainen and Valkealahti, 2017), resulting in unpredictable losses. These losses will however last for very short time periods.

3.2. Impact of the identified effects on the performance evaluation methods

An overview of all the identified effects which generate noise or errors in the performance estimations discussed in Section 3.1 is given in Table 3. A qualitative assessment of their influence on the different performance metrics is provided.

The impact of consecutively filtering out the effects with largest contribution to noise (snow, low light conditions, curtailment and periods with systematic differences in irradiance levels) on the standard deviation of the performance metrics is presented in Fig. 7. Rapid irradiance changes give less contribution to noise, because it only affects short time periods. The constant system losses introduce no variability, but a bias in the metrics, especially for the *PPI* based on physical modelling. This can be corrected for by learning from the data, but not filtered out. The filters presented in Table 4 are used for removing data, based on the findings presented in previous sections. For the irradiance difference filtering, a ± 2.5% difference between the inverters clear sky current and the clear sky median inverter current or the scaled clear sky measured irradiance is set. As the filtering is performed consecutively and the parameters are not independent, this does not quantify the noise generation of the different parameters. The variation in all the metrics is significantly reduced when the filters for removing snow periods and low light conditions are applied. The performance index based on machine learning modelling especially experienced large variations at low light conditions due to frequent underestimation of the power in these periods. The results are compared to more standard filtering approaches based on irradiance and clearness level in Table 5. We see that a filter for low light conditions based on the parameters in Table 4 in combination with a snow filter gives lower variation and more data than a standard irradiance or clear sky filter.

Table 4
Filters used in the standard filtering approach and the system specific filtering approach.

Parameter	Filters
Standard	
Irradiance	$G_{POA} > 200\ W/m^2$
Clear sky	pvlib detect_clearsky
System specific	
Snow	Snow depth > 0 m
Irradiance difference	$I_{Clear\ sky,\ Inverter}/I_{Clear\ sky,\ reference} < \pm 0.025$
Low light	
Irradiance	$G_{POA} > 50\ W/m^2$
Solar elevation	Solar elevation angle $> 20^o$

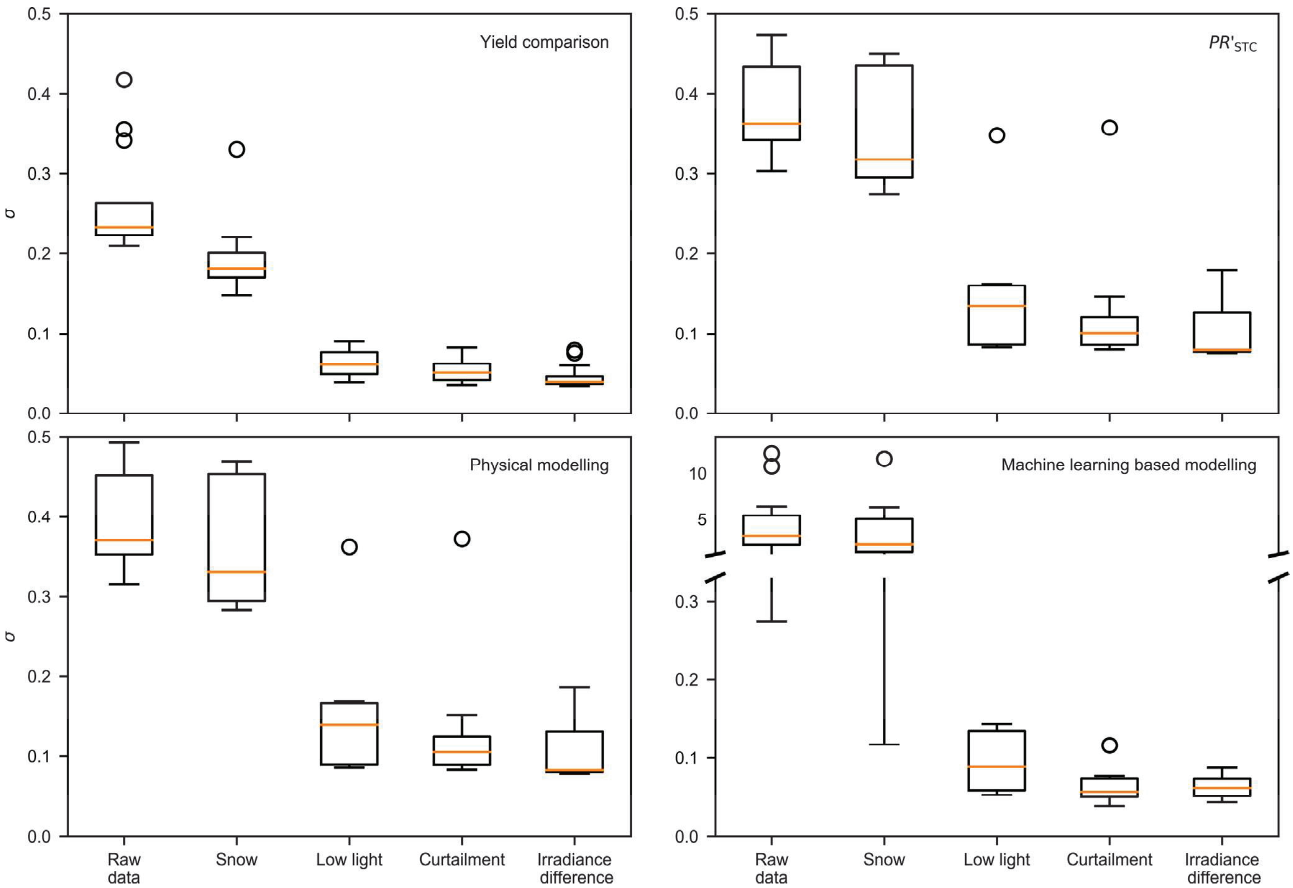


Fig. 7. Variation in median standard deviation in different performance metrics for the different MPPTs in all the systems, showing the effect when different filters (snow, low light conditions, curtailment, irradiance difference) is consecutively applied. The box extends from the first to the third quartile values of the data, with a line on the median. The whiskers extend to maximum 1.5 multiplied the interquartile range, and outliers are shown as circles.

Table 5
Median standard deviation and share of remaining data points and energy for the system specific filtering (consecutively applied) compared to more standard filtering strategies: 200 W/m^2 irradiance cut off and clear sky filtering. For the irradiance difference filter there is especially large system variations in how much of the data that are removed.

Filters	*Median standard deviation*					*Median*	
	$Y_{f\ rel}$	PR'_{STC}	*PPI:* PVWatts	*PPI*: Single diode	*PPI:* Machine learning	Remaining data points	Remaining energy
None (raw data)	0.23	0.36	0.38	0.39	3.44	100%	100%
Standard filtering							
Irradiance > 200 W/m^2	0.14	0.17	0.18	0.18	0.17	34%	76%
Clear sky	0.22	0.36	0.38	0.38	0.45	14%	20%
Specific filtering							
Snow	0.18	0.32	0.33	0.34	2.57	94%	98%
Low light	0.06	0.13	0.14	0.14	0.09	47%	83%
Curtailment	0.05	0.10	0.10	0.11	0.06	43%	69%
Irradiance difference	0.04	0.08	0.08	0.08	0.06	–	3–58%

The clear sky filter especially removes large amounts of data under the given climatic conditions.

The variations in the metrics are further reduced by adding filters for systematic irradiance differences, curtailment and inverter clipping. Some of the systems had poor tilt angle match between irradiance sensor and the modules, resulting in removal of a large share of the data when the irradiance difference filter is applied for the irradiance-based performance metrics. Only the machine learning model can handle the systematic irradiance differences caused by shading and the difference in tilt between different PV modules and the irradiance sensor. Fig. 7 shows that the irradiance difference filter does not have the same impact on reduction of the standard deviation for the machine learning based *PPI* as for the other performance metrics. Traditional low irradiance filtering or clear sky filtering is not sufficient in situations of irradiance differences and curtailment. In some cases, irradiance differences caused by shading and topography variations will be most prominent at low irradiance levels, and inadvertently be removed at low irradiance filtering. These effects will however be more prominent at clear sky, and curtailment and inverter clipping are more likely to occur at high energy generation and irradiance.

The yearly PR'_{STC} values using data filtered for irradiance difference vary from 0.68 to 0.93. In the winter months we see significant losses for all the systems due to both snow and low light condition, and in the summer some of the systems have significant curtailment losses.

4. Conclusions

Five effects that reduce the stability of the monitoring systems are identified: (i) Snow, (ii) curtailment & clipping, (iii) systematic irradiance differences over the system, (iv) low light conditions and (v) rapid changes in irradiance. The four first are most influential, as they might impact longer time periods. The standard deviation of all the performance metrics are significantly reduced when narrowly targeted filters for these four effects are applied. Compared to general low irradiance or clear sky filtering, the reduction in standard deviation of the metrics is greater, while more data remains in the useful dataset. Difficulty in estimating the constant system losses for different installation and sensor data quality issues may also introduce bias in some of the performance metrics.

More efficient filtering of low light conditions can be achieved by filtering with respect to both solar elevation and irradiance to directly address the issues leading to noise. To detect periods with systematic differences in irradiance between different units, statistical clear sky fitting can be employed.

The solutions of the discussed challenges and the specific filtering approaches are relevant for analysis and monitoring of most PV systems, when recognizing different effects in the data and efficient filtering is necessary. Automatic detection of periods with irradiance differences caused by e.g. shading or topography is particularly important for robust PV system analysis. The suggested methods will however be especially important for systems at higher latitudes exposed to more unstable conditions.

Filtered values of specific yield comparison and power performance index based on machine learning modelling yielded best results in terms of a stable metric. *PPI* based on physical modeling gives inaccurate results because of insufficient input data to estimate losses and dependency of accurate irradiance measurements. Machine learning based modeling handled both these challenges and periods with systematic differences in irradiance more efficiently, showing that data driven methods might be particularly suitable for challenging weather conditions and systems with lower grade of uniformity.

The analysis highlights two paths to accomplish increased reliability of PV monitoring systems without increased hardware costs. First, better reliability can be achieved by evaluating the availability and the quality of the input data, then based on this choose a suitable performance metric. Second, the variability of the chosen performance metric can be reduced by utilizing filters that specifically target the origin of the variability instead of using typical literature thresholds. As this analysis focuses on challenges and limitations met in monitoring of commercial PV systems, implementing the suggested solutions in monitoring software would improve the performance analysis and fault detection in the system. This could increase the PR and reduce the levelized cost of electricity.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge funding from the University of Oslo, Innovation project number 803801 (Autonomous monitoring, control and protection of renewable energy infrastructure).

References

Andrews, R.W., Pearce, J.M., 2012. Prediction of energy effects on photovoltaic systems due to snowfall events. In: 2012 38th IEEE Photovoltaic Specialists Conference. IEEE, pp. 3386–3391. https://doi.org/10.1109/PVSC.2012.6318297.

Beck, H.E., Zimmermann, N.E., McVicar, T.R., Vergopolan, N., Berg, A., Wood, E.F., 2018. Present and future köppen-geiger climate classification maps at 1-km resolution. Sci. Data 5. https://doi.org/10.1038/sdata.2018.214.

Belhachat, F., Larbes, C., 2015. Modeling, analysis and comparison of solar photovoltaic array configurations under partial shading conditions. Sol. Energy 120, 399–418. https://doi.org/10.1016/j.solener.2015.07.039.

Belluardo, G., Ingenhoven, P., Sparber, W., Wagner, J., Weihs, P., Moser, D., 2015. Novel method for the improvement in the evaluation of outdoor performance loss rate in different PV technologies and comparison with two other methods. Sol. Energy 117,

139–152. https://doi.org/10.1016/j.solener.2015.04.030.
Blair, N., Dobos, A.P., Freeman, J., Neises, T., Wagner, M., Ferguson, T., Janzou, S., 2011. System advisor model, sam 2014.1. 14: General description, No. NREL/TP-6A20-61019. Golden, CO (United States).
Camus, C., Hüttner, M., Lassahn, D., Kurz, C., Hauch, J., Brabec, C.J., 2018. Data-filtering-dependent variability of long-term degradation rates of MW-scale photovoltaic power plants from "non-ideal" monitoring and weather data. In: Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition, pp. 2069–2074 https://doi.org/10.3130/aijsaxx.152.0_69.
Corkish, R., Green, M.A., Watt, M.E., Wenham, S.R., 2013. Applied Photovoltaics. Routledge. https://doi.org/10.4324/9781849770491.
Curran, A.J., Jones, C.B., Lindig, S., Stein, J., Moser, D., French, R.H., 2019. Performance loss rate consistency and uncertainty across multiple methods and filtering criteria. In: 2019 IEEE 46th Photovoltaic Specialists Conference. IEEE, pp. 1328–1334.
Daliento, S., Chouder, A., Guerriero, P., Pavan, A.M., Mellit, A., Moeini, R., Tricoli, P., 2017. Monitoring, diagnosis, and power forecasting for photovoltaic fields: A review. Int. J. Photoenergy 2017. https://doi.org/10.1155/2017/1356851.
Dierauf, T., Growitz, A., Kurtz, S., Hansen, C., 2013. Weather-corrected performance ratio, No. NREL/TP-5200-57991. Golden, CO (United States).
Dobos, A.P., 2012. An improved coefficient calculator for the California energy commission 6 parameter photovoltaic module model. J. Sol. Energy Eng. 134. https://doi.org/10.1115/1.4005759.
Dobos, A.P., 2014. PVWatts Version 5 Manual, No. NREL/TP-6A20-62641. Golden, CO (United States). https://doi.org/10.2172/1158421.
Dobos, A.P., Freeman, J.M., Blair, N.J., 2019. Improvements to PVWatts for Fixed and One-Axis Tracking Systems, No. NREL/CP-6A20-74097. Golden, CO (United States). https://doi.org/10.1109/pvsc40753.2019.8981312.
Dong, A., Zhao, Y., Liu, X., Shang, L., Liu, Q., Kang, D., 2017. Fault diagnosis and classification in photovoltaic systems using scada data. In: 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control. IEEE, pp. 117–122 https://doi.org/10.1109/SDPC.2017.31.
Fouad, M.M., Shihata, L.A., Morgan, E.S.I., 2017. An integrated review of factors influencing the performance of photovoltaic panels. Renew. Sustain. Energy Rev. 80, 1499–1511. https://doi.org/10.1016/j.rser.2017.05.141.
Grunow, P., Lust, S., Sauter, D., Hoffmann, V., Podlowski, L., 2004. Weak light performance and annual yields of PV modules and systems as result of the basic parameter set of industrial solar cells. In: Proceedings of the 19th European Photovoltaic Solar Energy Conference, pp. 2190–2193.
Holmgren, F.W.W., Hansen, C.A., Mikofski, M., 2018. pvlib python: a python package for modeling solar energy systems. J. Open Source Softw. 3, 884. https://doi.org/10.21105/joss.00884.
Huld, T., Müller, R., Gambardella, A., 2012. A new solar radiation database for estimating PV performance in Europe and Africa. Sol. Energy 86, 1803–1815. https://doi.org/10.1016/j.solener.2012.03.006.
IEC, 2017. Photovoltaic system Performance - Part 1: Monitoring (IEC 61724-1.0, 2017-03). Geneva, Switerland.
Imenes, A.G., Selj, J., 2017. Irradiance and temperature distributions at high latitudes: Design implications for photovoltaic systems. In: 2017 IEEE 44th Photovoltaic Specialist Conference. IEEE, pp. 619–625 https://doi.org/10.1109/PVSC.2017.8366376.
Jordan, D.C., Kurtz, S.R., 2014. The dark horse of evaluating long-term field performance-Data filtering. IEEE J. Photovoltaics 4, 317–323. https://doi.org/10.1109/JPHOTOV.2013.2282741.
Jordan, D.C., Deline, C., Kurtz, S.R., Kimball, G.M., Anderson, M., 2017. Robust PV degradation methodology and application. IEEE J. Photovoltaics 8, 525–531. https://doi.org/10.1109/JPHOTOV.2017.2779779.
Jordan, D.C., Deline, C., Deceglie, M.G., Nag, A., Kimball, G.M., Shinn, A.B., John, J.J., Alnuaimi, A.A., Elnosh, A.B.A., Luo, W., Jain, A., Saleh, M.U., von Korff, H., Hu, Y., Jaubert, J.-N., Mavromatakis, F., 2019. Reducing interanalyst variability in photovoltaic degradation rate assessments. IEEE J. Photovoltaics 10, 206–212. https://doi.org/10.1109/jphotov.2019.2945191.
King, D.L., Kratochvil, J.A., Boyson, W.E., 2004. Photovoltaic array performance model, No. SAND2004-3535. https://doi.org/10.2172/919131.
Klise, G.T., Balfour, J.R., Keating, T.J., 2014. Solar PV O&M standards and best practices - existing gaps and improvement efforts, No. SAND2014-19432. Albuquerque, NM (United States).
Kurtz, S., Newmiller, J., Kimber, A., Flottemesch, R., Riley, E., Dierauf, T., McKee, J., Krishnani, P., 2013. Analysis of photovoltaic system energy performance evaluation method, No. NREL/TP-5200-60628. Golden, CO (United States).
Lappalainen, K., Valkealahti, S., 2017. Output power variation of different PV array configurations during irradiance transitions caused by moving clouds. Appl. Energy 190, 902–910. https://doi.org/10.1016/j.apenergy.2017.01.013.
Livera, A., Theristis, M., Makrides, G., Georghiou, G.E., 2019. Recent advances in failure diagnosis techniques based on performance data analysis for grid-connected photovoltaic systems. Renew. Energy 133, 126–143. https://doi.org/10.1016/j.renene.2018.09.101.
Louwen, A., de Waal, A.C., Schropp, R.E.I., Faaij, A.P.C., van Sark, W., 2017. Comprehensive characterisation and analysis of PV module performance under real operating conditions. Prog. Photovoltaics 25, 218–232. https://doi.org/10.1002/pip.
Marion, B., Schaefer, R., Caine, H., Sanchez, G., 2013. Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations. Sol. Energy 97, 112–121. https://doi.org/10.1016/j.solener.2013.07.029.
Meftah, M., Lajoie-Mazenc, E., Van Iseghem, M., Perrin, R., Boublil, D., Radouane, K., 2019. A less environment-sensitive and data-based approach to evaluate the performance loss rate of PV power plants. In: Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition, pp. 1554–1559. https://doi.org/10.1017/CBO9781107415324.004.
Meyers, B., Deceglie, M., Deline, C., Jordan, D., 2019. Signal Processing on PV Time-Series Data : Robust Degradation Analysis without Physical Models. 46th IEEE Photovolt. Spec. Conf. PP, 1–8. https://doi.org/10.1109/JPHOTOV.2019.2957646.
Meyers, B., Tabone, M., Kara, E.C., 2018. Statistical Clear Sky Fitting Algorithm, in: 2018 45th IEEE Photovoltaic Specialist Conference. IEEE.
Moser, D., Pichler, M., Nikolaeva-Dimitrova, M., 2014. Filtering procedures for reliable outdoor temperature coefficients in different photovoltaic technologies. J. Sol. Energy Eng. 136. https://doi.org/10.1115/1.4024931.
NVE, 2019. seNorge [WWW Document]. URL www.senorge.no (accessed 9.1.19).
Øgaard, M.B., Haug, H., Selj, J., 2018. Methods for quality control of monitoring data from commercial PV systems. In: Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition, pp. 2083–2088. https://doi.org/10.4229/35thEUPVSEC20182018-6DV.1.53.
Øgaard, M.B., Skomedal, Å., Selj, J.H., 2019. Performance evaluation of monitoring algorithms for photovoltaic systems. In: Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition, pp. 1632–1636.
Pedregosa, F., Weiss, R., Brucher, M., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine learning in python. J. Mach. Learn. Res. 12, 2825–2830.
Ransome, S., 2008. Array performance analysis using imperfect or incomplete input data. In: Proceedings of the 23rd European Photovoltaic Solar Energy Conference and Exhibition, pp. 3187–3191.
Reich, N.H., Goebel, A., Dirnberger, D., Kiefer, K., 2012. System performance analysis and estimation of degradation rates based on 500 years of monitoring data. In: 2012 38th IEEE Photovoltaic Specialists Conference. IEEE, pp. 1551–1554. https://doi.org/10.1109/PVSC.2012.6317890.
Reise, C., Müller, B., 2018. Uncertainties in PV system yield predictions and assessments. Report IEA-PVPS T13–12.
Reno, M.J., Hansen, C.W., 2016. Identification of periods of clear sky irradiance in time series of GHI measurements. Renew. Energy 90, 520–531. https://doi.org/10.1016/j.renene.2015.12.031.
Rodrigues, S., Ramos, H.G., Morgado-Dias, F., 2017. Machine learning in PV fault detection, diagnostics and prognostics: A review. In: 2017 44th IEEE Photovoltaic Specialist Conference. IEEE, pp. 3178–3183. https://doi.org/10.1109/PVSC.2017.8366581.
Sanchis, P., López, J., Ursúa, A., Gubía, E., Marroyo, L., 2007. On the testing, characterization, and evaluation of PV inverters and dynamic MPPT performance under real varying operating conditions. Prog. Photovoltaics Res. Appl. 15, 541–556. https://doi.org/10.1002/pip.
Silvestre, S., Mora-López, L., Kichou, S., Sánchez-Pacheco, F., Dominguez-Pumar, M., 2016. Remote supervision and fault detection on OPC monitored PV systems. Sol. Energy 137, 424–433. https://doi.org/10.1016/j.solener.2016.08.030.
Skomedal, Å., Øgaard, M.B., Selj, J.H., Haug, H., Marstein, E.S., 2019. General, robust and scalable methods for string level monitoring in utility scale. In: PV systems, in: Proceedings of the 36th European Photovoltaic Solar Energy Conference and Exhibition, pp. 1283–1287.
Stein, J.S., Farnung, B., 2017. PV performance modeling methods and practices results from the 4th PV performance modeling collaborative workshop. Report IEA-PVPS T 13–06, 201.
Triki-Lahiani, A., Abdelghani, A.B.-B., Slama-Belkhodja, I., 2018. Fault detection and monitoring systems for photovoltaic installations: A review. Renew. Sustain. Energy Rev. 82, 2680–2692. https://doi.org/10.1016/j.rser.2017.09.101.
Tsanakas, J.A., Ha, L., Buerhop, C., 2016. Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: A review of research and future challenges. Renew. Sustain. Energy Rev. 62, 695–709. https://doi.org/10.1016/j.rser.2016.04.079.
Westbrook, O.W., Copanas, B.H., Collins, F.D., 2012. New approaches for characterizing photovoltaic system performance, in: 2012 38th IEEE Photovoltaic Specialists Conference. IEEE, pp. 1529–1534. https://doi.org/10.1109/PVSC.2012.6317886.
Whaley, C., 2016. Best practices in photovoltaic system operations and maintenance, No. NREL/TP-7A40-67553. Golden, CO (United States).
Woyte, A., Richter, M., Moser, D., Green, M., Mau, S., Beyer, H.G., 2014. Analytical monitoring of grid-connected photovoltaic systems, Report IEA-PVPS T13-03. https://doi.org/10.13140/2.1.1133.6481.

Paper IV

Identifying snow in photovoltaic monitoring data for improved snow loss modeling and snow detection

M.B. Øgaard, B.L. Aarseth, Å.F. Skomedal, H.N. Riise, S. Sartori, and J.H. Selj

In: *Solar Energy,* Vol. 223 (2021), pp. 238-247.
DOI: 10.1016/j.solener. 2021.05.023

IV

Solar Energy 223 (2021) 238–247

Contents lists available at ScienceDirect

Solar Energy

journal homepage: www.elsevier.com/locate/solener

Identifying snow in photovoltaic monitoring data for improved snow loss modeling and snow detection

Mari B. Øgaard [a,b,*], Bjørn L. Aarseth [a,b], Åsmund F. Skomedal [b], Heine N. Riise [b], Sabrina Sartori [a], Josefine H. Selj [a,b]

[a] Department of Technology Systems, University of Oslo, Gunnar Randers vei 19, 2007 Kjeller, Norway
[b] Solar Power Systems Department, Institute for Energy Technology, Instituttveien 18, 2007 Kjeller, Norway

A R T I C L E I N F O

Keywords:
Photovoltaic systems
Snow
Snow loss modeling
Snow detection
PV performance
Soiling

A B S T R A C T

As cost reductions have made photovoltaics (PV) a favorable choice also in colder climates, the number of PV plants in regions with snowfalls is increasing rapidly. Snow coverage on the PV modules will lead to significant power losses, which must be estimated and accounted for in order to achieve accurate energy yield assessment and production forecasts. Additionally, detection and separation of snow loss from other system losses is necessary to establish robust operation and maintenance (O&M) routines and performance evaluations.

Snow loss models have been suggested in the literature, but developing general models is challenging, and validation of the models are lacking. Characterization and detection of snow events in PV data has not been widely discussed.

In this paper, we identify the signatures in PV data caused by different types of snow cover, evaluate and improve snow loss modeling, and develop snow detection. The analysis is based on five years of data from a commercial PV system in Norway. In an evaluation of four snow loss models, the Marion model yields the best results. We find that system design and snow depth influence the natural snow clearing, and by expanding the Marion model to take this into account, the error in the modeled absolute loss for the tested system is reduced from 23% to 3%. Based on the improved modeling and the identified data signatures we detect 97% of the snow losses in the dataset. Endogenous snow detection constitutes a cost-effective improvement to current monitoring systems.

1. Introduction

Due to a substantial decline in the price of photovoltaic (PV) installations in recent years, large scale PV plants are increasingly common in cold climates with wintertime snowfalls (Burnham et al., 2020; Hashemi et al., 2020; IEA, 2020; Jäger-Waldau, 2020). This development necessitates robust methods for analyzing PV yield and performance, as well as flexible monitoring and forecasting solutions in snowy conditions. Thus, accurate snow loss modeling and snow detection are required.

Snow losses are expected to vary significantly with climate, system configuration and from year to year. At its maximum, it might give monthly losses up to 100% in the winter season and annual losses above 30% (Pawluk et al., 2019). Consequently, it is an important parameter to consider in simulation and yield assessment of future PV systems in locations with snowfalls, as well as in production forecasts and performance and loss analysis of historical PV data. Snow losses will also introduce significant challenges in monitoring, giving signatures in the production data which resemble failures. A full snow cover gives an electrical response similar to an inverter breakdown. A partial snow cover leading to partial shading can give electrical losses (Schill et al., 2015) similar to serious PV module failures (Tsanakas et al., 2016). When using empirical or machine learning based methods for PV modeling, snow events in the training data will perturb the correlations between irradiance, temperature and production. These perturbations can increase the uncertainty of the models (Øgaard et al., 2020).

Recent research has demonstrated that uncertainty in yield estimations (Bosman and Darling, 2018; Marion et al., 2013; Ryberg and Freeman, 2017; Townsend and Powers, 2011) and forecasting (Lorenz et al., 2011) can be reduced if snow loss models are included. Despite this, snow loss models are often not implemented in PV simulation software. The System Advisor Model (SAM) has implemented the model

* Corresponding author at: Department of Technology Systems, University of Oslo, Gunnar Randers vei 19, 2007 Kjeller, Norway.
E-mail address: mari.ogaard@its.uio.no (M.B. Øgaard).

https://doi.org/10.1016/j.solener.2021.05.023
Received 13 January 2021; Received in revised form 4 May 2021; Accepted 7 May 2021
Available online 28 May 2021

suggested by Marion et al. (Marion et al., 2013; Ryberg and Freeman, 2017), but in other software, snow is either not considered (PVGIS, 2020) or estimated by constant soiling values (PVsyst, 2020; Solargis, 2016), typically not related to the climatic conditions.

In PV monitoring, if at all considered, detection of snow is a more common approach than snow loss modeling. In the literature, snow detection methods based on dedicated or external sensors like weight sensors, web cameras and satellite data have been proposed (Aarseth et al., 2018; Andrews et al., 2013; Wirth et al., 2010). Ambient temperature (Lorenz et al., 2007) and module temperature (Øgaard et al., 2018) have been suggested as measurements that can be used to identify snow-related losses in PV monitoring and failure diagnosis. Except for this, identifying and characterizing the effects of snow in PV monitoring data, a prerequisite to separate snow losses from failures and a method to cost effectively detect snow, is not widely discussed.

Accurate snow loss modeling and robust snow detection are challenging, because the parameters influencing the snow cover and resulting PV system loss are manifold. The influential parameters range from weather conditions (irradiance, temperature, wind, etc.), to installation and technology specific configurations (tilt, module technology, ground/roof mounted, etc.) and type of snow. This is challenging for both physical and empirical models due to the amount of required input data. Existing snow loss models, use weather data and technical system configuration to either estimate (i) snow coverage or (ii) the losses directly (Pawluk et al., 2019). Most of the suggested methods are based on empirical approaches, including both simple linear relationships (Pawluk et al., 2019) and machine learning (Bashir et al., 2020; Hashemi et al., 2020). Validation of the models on other PV systems is typically lacking (Ryberg and Freeman, 2017). While the uncertainty for monthly and annual losses often are low compared to the size of the loss, the uncertainty on daily and higher time resolutions is high (Andrews and Pearce, 2012; Marion et al., 2013).

In particular, it is the process of natural snow clearing that is difficult to model. The main mechanisms of natural snow clearing are melting and sliding, both effects typically connected to ambient temperatures larger than 0 ℃ (Pawluk et al., 2019), but sliding at −10 °C has also been observed (Becker et al., 2006). Friction and adhesion between the snow and the solar panels are parameters that contribute to the complexity of natural snow clearing, as both are expected to vary with type of snow (Andrews et al., 2013; Pawluk et al., 2019). While wet snow has lower friction, it is also more likely to freeze to the module (Andenæs et al., 2018; Ross, 1995). Natural snow clearing is thus dependent on how temperature evolves with time. Additionally, system configurations like tilt and elements obstructing the path of snow sliding (e.g. the module frame (Riley et al., 2019), or little empty space below the modules giving ground/roof interference (Heidari et al., 2015)) will impact natural snow clearing. Technical system aspects might also impact the heat transfer to the system and thus the snow melting. Increased melting can e.g. be caused by absorbed reflected irradiance on the rear side for a ground mounted bifacial system (Burnham et al., 2019), or by poor roof insulation for a roof mounted system.

Because different types of modules have different shading response, snow losses and the signatures in the electrical data will also depend on type of modules (thin film or crystalline silicon, full or half cells, monofacial or bifacial), and for the most typical crystalline silicon (c-Si) module with three bypass diodes: whether the modules are installed in portrait or landscape orientation. When the snow slides down the tilted module, it typically shades the lower part, as shown in Fig. 1. This gives shading orthogonal to the substring current for modules installed in portrait orientation, and parallel to the substring current for modules installed in landscape. In the first case, all the substrings in the modules are impacted, in the second case, the shaded area can be bypassed by the bypass diodes. This can lead to significantly higher snow-related losses for modules installed in portrait orientation than modules installed in landscape orientation under similar partial snow covers (Andenæs et al., 2018; Andrews et al., 2013; van Noord et al., 2017). On the other hand,

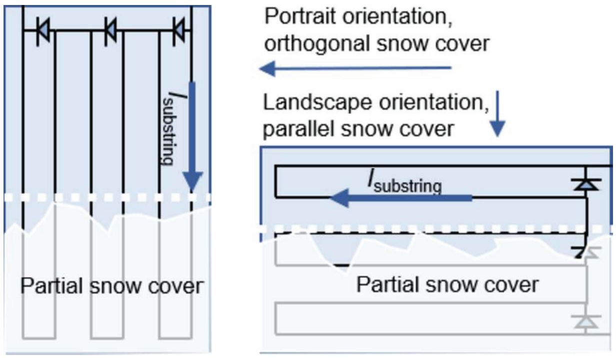


Fig. 1. Illustration of portrait and landscape module orientation and how a partial snow cover typically shades the tilted modules when it is sliding off the modules. A module installed in portrait is shaded orthogonal to the module substring current, and a module installed in landscape is shaded in parallel with the module substring current.

natural snow clearing has been observed to happen faster for modules in portrait orientation than for modules in landscape orientation (Burnham et al., 2020). One of the suggested explanations is that the frame impedes sliding more for modules in landscape. Additionally, if the modules are irradiated but generate no or low current (compared to the irradiance level), they are expected to be warmer because the energy is not converted to electricity (Teubner et al., 2019). Because modules in portrait will generate less current under partial snow cover, they will become warmer than modules in landscape, aiding the melting process.

The aim of this work is to (i) characterize the effect of snow in PV monitoring data, (ii) to assess and improve on existing PV snow loss models, and (iii) to develop snow detection methods for PV monitoring data. The main focus is on monofacial c-Si technology. To characterize the impact of different types of snow covers on the measured variables of a PV system, we have analyzed data from two PV systems in Norway with regular snow cover in the winter. The identified signatures in PV monitoring data caused by snow, are assessed by using simulations of shaded modules and transmittance measurements. The PV monitoring data is further used for evaluation and improvement of snow loss models, and both the improved snow loss models and the signatures are used in development of snow detection.

2. Methodology

2.1. PV monitoring data

The PV monitoring data utilized in this study is primarily from a commercial 185 kW_p roof top PV system. The data from the commercial system is complemented by detailed studies and experimental data obtained from a 4 kW_p ground mounted test system at the Institute for Energy Technology research facility. The commercial PV system is installed on a flat roof, the lowest part of the modules nearly touching the rooftop. The multicrystalline silicon PV modules are oriented East-West with a tilt of 10°, and installed in landscape orientation. This system configuration is typical for many of the larger PV systems in the Nordic countries. A tilt angle of 10° is not optimal for high latitude locations with respect to maximum production, but is typically used in installations on flat roof systems also at high latitudes to maximize roof coverage and to, in combination with the East-West orientation, achieve a more even production distribution through the year and the day. The test system has South-oriented crystalline silicon modules with a tilt of 28° installed in portrait orientation (FME Susoltech, 2020). The array height is two modules for the test system, and one module for the commercial system. Both systems are located approximately 60° North and 11° East.

For both systems, 5 years of data are collected. For the commercial system, current, voltage and power data are collected from the inverters, and temperature and irradiance data are collected from the monitoring system of the installation. The data is logged at 5 min interval, and the DC inverter data are measured at maximum power point tracker level (MPPT). Three strings of 24 modules are connected in parallel to one MPPT. In the test system, the electrical data is measured at module level by power optimizers, and the recording interval of the data logging is 15 min. The effective in plane irradiance incident on the PV modules, i.e. the irradiance the modules can utilize (Stein and Farnung, 2017), is in both cases measured with a crystalline silicon reference cell in the plane of the PV modules. The measurement uncertainty of the irradiance is ±5 W/m^2 ± 2.5% of reading. The module temperature is measured by a sensor attached to the rear side of the modules. Wind and humidity data was collected from nearby weather stations (Norsk Klimaservicesenter, 2020). To identify time periods with snow on the reference cells, the reference cell irradiance measurements are compared to measurements from irradiance sensors at the systems that are observed to be less effected by snow: a heated horizontal pyranometer at the commercial system, and a vertical pyranometer at the test system. To reduce the effect of snow-covered irradiance sensors on the analysis, the measurements from the heated horizontal pyranometer are used as a replacement of the in plane irradiance at the commercial system for days where the daily irradiation measured by the pyranometer irradiation is more than twice the daily reference cell irradiation.

At the test system, the transmittance of the snow cover on the modules was measured for different snow cover thicknesses, by measuring the irradiance on the front and rear side of a full size module glass with the same tilt as the PV modules, using a spectroradiometer (Spectral Evolution, PSR-1100F). The measurements were conducted over 7 days with different snow and irradiance conditions. Observations of the snow coverage on the modules were collected at the same time by sample images. Daily estimated snow data for the two locations, based on interpolated observational data are collected from seNorge.no (NVE, 2019).

2.2. Identification of snow signatures in PV monitoring data

2.2.1. Observed snow signatures

To study the signatures of snow in PV monitoring data, deviations compared to snow free production for electrical DC data (power, voltage, current) and module temperature in time periods after snow falls are evaluated. The evaluation is performed for both modules installed in landscape and portrait orientation.

The expected power, voltage and current for snow free conditions is modeled by a single diode model. Module datasheet values and PySAM (NREL, 2020) are used to estimate the diode ideality factor, light generated current, dark reverse saturation current, shunt resistance and series resistance at reference conditions, and the parameter for adjusting the short circuit current temperature coefficient, as described by Dobos (2012). These parameters together with the measured effective irradiance and cell temperature are used as inputs to the CEC model (Dobos, 2012), which estimates the photocurrent, saturation current, shunt resistance and thermal cell voltage. The expected electrical output for each module is estimated by solving the single diode equation based on the parameters estimated with the CEC model, as implemented in pvlib python (Holmgren et al., 2018). The constant system losses are estimated by comparing the modeled power to the measured power under snow free conditions. Based on this, some differences were observed in the angular response between the reference cell and the module strings, giving seasonal variation in the system losses. To compensate for this, the additional reflection loss of the modules, was modeled with the ASHRAE IAM model (Holmgren et al., 2018; Souka and Safwat, 1966) with an IAM adjustment parameter of 0.03. The expected PV module temperature is modeled by the cell temperature model from the Sandia Array Performance Model (SAPM) (Holmgren et al., 2018; King et al., 2004), where the module temperature is estimated based on global irradiance, ambient temperature, and wind speed.

Uncertainty in PV modeling is typically higher at lower irradiance and high angles of incidence, as it is challenging to capture all loss effects under these conditions. This can give a small absolute, but high relative, overestimation of the expected power and current in the wintertime, and thus overestimation of the snow losses in these parameters. Snow on the irradiance sensor can on the other hand lead to underestimation of both absolute and relative losses. On a monthly basis for the periods without snow, the mean absolute error in the daily modeled energy generation for the commercial system is up to 0.1 kWh/kW$_p$ in the summer months and down to 0.02 kWh/kW$_p$ in the winter months. For both systems, the mean absolute percentage error in daily modeled energy is 2% for most months, but in the darkest winter months when the energy generation can be <1 kWh/kW$_p$ per day, small deviations in the model can give high relative errors, up to 20%. When the expected energy generation is aggregated for longer time periods, the days with highest production and lowest uncertainty will dominate and reduce the relative uncertainty.

2.2.2. Simulated electrical snow signatures

The expected electrical signatures in PV module data for different snow covers are modeled using circuit simulations in MATLAB Simulink. A system with the same configuration as the commercial system described in Section 2.1 is modeled, with 60 cell modules having 3 bypass diodes each. A variable voltage source is used to trace the full IV curve of the modeled system. Solar cell blocks in Simulink are modeled by solving the single diode equation, and piecewise linear diodes are utilized as bypass diodes. The Simulink solar cell single diode parameters are fitted so that 60 cells in series match the IV characteristics of the commercial system. The simulations are performed for a case where the cell temperature is 25 °C and the irradiance is 450 W/m^2. Snow covers are simulated as a reduction in irradiance for the covered part of the modules, and the resulting power, current and voltage from the simulated IV trace are used to calculate electrical losses for different shading situations. The loss is calculated by comparing the yield of the snow-covered system with an unshaded, identical system. The covers are varied in size and transmittance, and the partial covers are modeled both for portrait and landscape module orientation, i.e. orthogonal and parallel to the substring current, respectively. The simulations are not validated through field data because we have no accurate measures of snow covers. This means we have no estimates of the performance of the simulation at low light conditions and what error is introduced by using the same cell temperature for all simulations, thereby not including the temperature differences caused by snow cover. The efficiency of the inverter at low irradiances, the MPPT voltage range of the inverter, and how the MPPT handles partial shading will also influence the loss in electrical parameters. We do, however, still believe that the simulations capture the general behavior at these conditions and help us understand how different snow covers impact electrical PV measurements.

2.3. Snow loss model evaluation

The calculated snow power loss, i.e. the deviation between the measured power and the modeled power (Section 2.2.1), is used to validate snow loss models. The data from the commercial system is used in the evaluation, as it has multiple identical arrays and is thus expected to give an insight into eventual loss variations for similar configurations. The tested models are the models suggested by Andrews and Pearce (2012), Powers et al. (2010), Townsend and Powers (2011), and Marion et al. (2013) as implemented in pvlib python (Holmgren et al., 2018; Ryberg and Freeman, 2017). The three first models aim to estimate the snow losses based on empirical correlations with different environmental parameters. Andrews and Pearce estimate daily losses based on a correlation between snow losses and irradiance, temperature, and the change in snow depth for the two last days. Powers et al. use a correlation between annual snow losses, snow depth and module tilt.

Townsend and Powers estimate monthly losses using a correlation between snow losses and humidity, temperature, irradiance, snow fall, and a ground interference parameter. The Marion model initially estimates the snow cover, and subsequently calculates the snow loss based on the snow cover estimate. The model assumes that when the snow starts to melt, it is cleared by sliding off the modules. Snowfall data are used to identify the presence of snow, and irradiance and module temperature are used to identify conditions where snow slides off the modules. Snow sliding is assumed to happen when:

$$T_{\text{amb}} > G_{\text{POA}}/m, \tag{1}$$

where T_{amb} is the ambient temperature, G_{POA} is the in plane irradiance and m is an empirically defined value of -80 W/(m^2 °C). How much the snow will slide, measured in fractions of the total row height, is determined by the tilt of the modules and an empirical sliding coefficient (*sc*):

$$Snow\ \ slide\ \ amount = sc^{*}\sin(tilt) \tag{2}$$

For roof mounted systems *sc* was found to be 0.20 (Holmgren et al., 2018; Marion et al., 2013). The snow loss is subsequently estimated from the calculated snow coverage and the number of parallel connected strings (including module substrings) along the row height, taking into account whether the modules are installed in portrait or landscape orientation (Holmgren et al., 2018). If a module substring is partially covered by snow, the capacity is assumed to be zero (Gilman et al., 2018). All snowfalls greater than 0 cm are included in the snow loss modeling.

3. Results and discussion

The impact of full and different levels of partial snow cover on the PV monitoring data is presented in Section 3.1.1. The results are assessed using simulations of shaded strings (Section 3.1.2) and transmittance measurements (Section 3.1.3). The signatures in the monitoring data caused by snow are summarized in Section 3.1.4. Evaluation and improvement of snow loss modeling is presented in Section 3.2. Based on the improved model and the snow signatures, a method for snow detection is proposed in Section 3.3.

3.1. Snow signatures

3.1.1. Observed snow signatures in PV monitoring data

Fig. 2 shows the daily losses in voltage, current and power for a time period with snow melting where the modules gradually are going from fully snow covered, through different levels of partial cover, to snow free. The event is in March/April, in a period with high irradiance, giving low relative uncertainty in the modeled expected value. The boxplot shows the variation in the measurements. In the beginning of the period, when the snow cover is assumed to be full and opaque, the losses in all electrical parameters are 100%. When the snow cover starts to melt, the first development is an increase in voltage. For some of the modules in the test system, voltage gain is registered. As the snow continues to melt, a stepwise reduction in voltage losses is observed, while the losses in current and power are gradually reduced. The variation in losses between different modules/inverters is large for partial snow cover, reflected in a large spread in the measured loss. For the test system (portrait orientation), where the loss is measured at module level and not aggregated for larger subarrays as for the commercial system, particularly large variations in both current and voltage are seen.

The module temperature is also significantly influenced by snow cover. Fig. 3 shows how the measured module temperatures in the test system develop compared to the ambient temperature and the modeled module temperature during the same melting period as in Fig. 2. The module temperature is quite stable at full snow cover with less pronounced diurnal variations than the ambient temperature. As the snow cover melts, the measured module temperatures are more impacted by irradiance and ambient temperature, and there are large variations between different module temperature sensors, due to the local variations

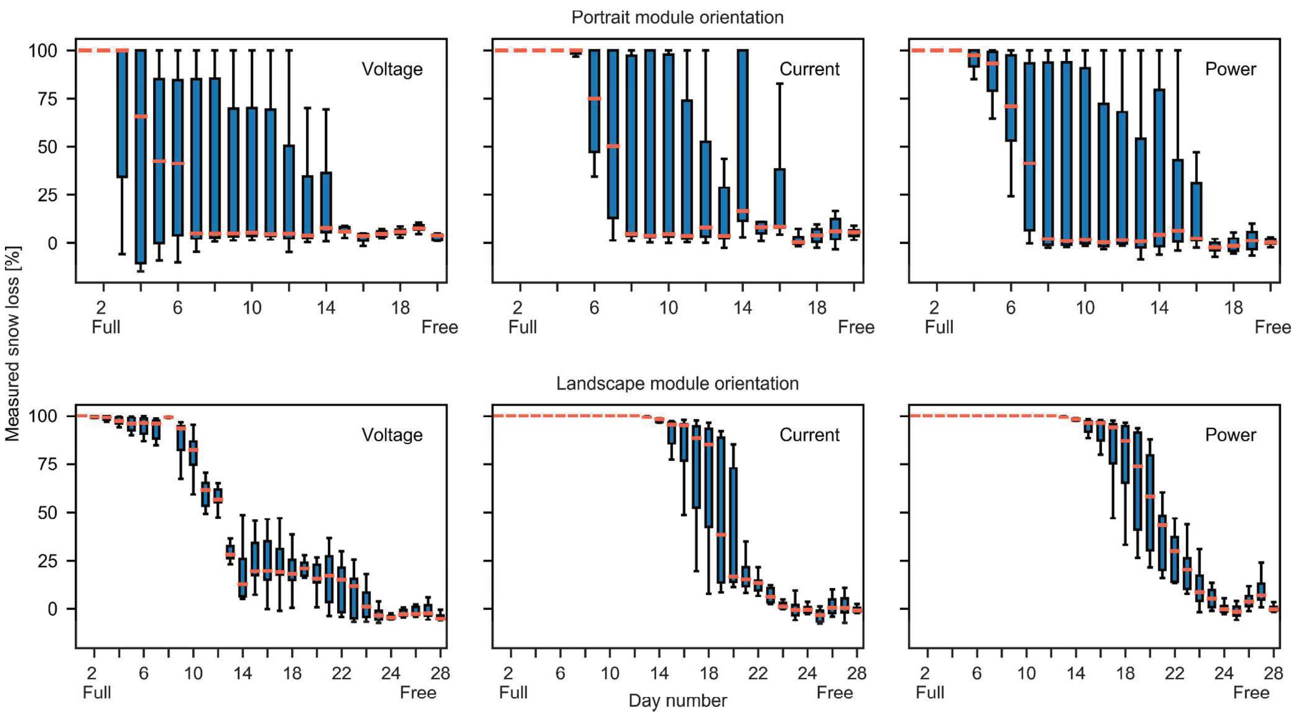


Fig. 2. The measured daily snow losses in voltage, current and power for the two different module orientations in a period where the modules go from fully covered, through different levels of partial cover, to snow free. The boxplot shows the variation in loss between different inverters/modules. The boxes extend from the first to the third quartile values of the data, with a line on the median. The whiskers extend to the maximum or minimum value within 1.5 times the interquartile range, and outliers are not included.

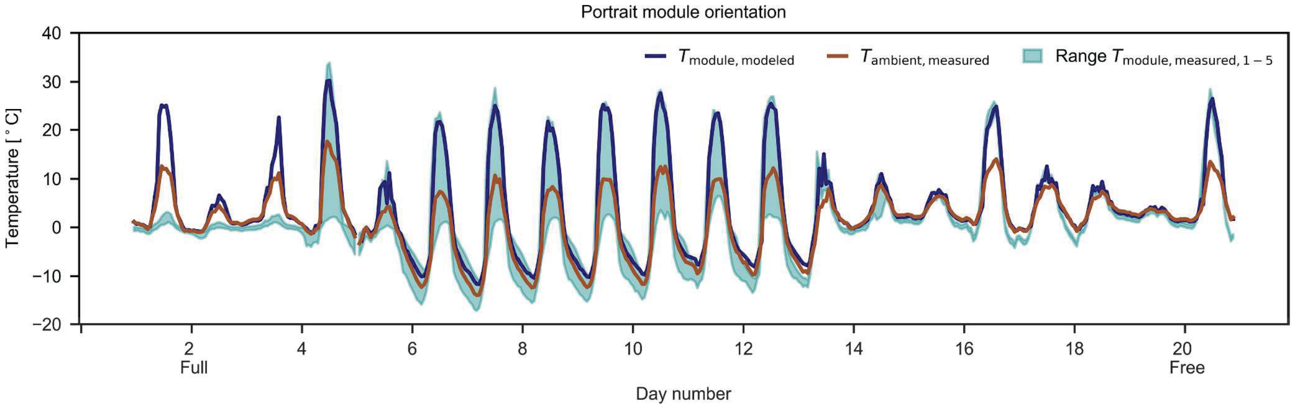


Fig. 3. Development of the modeled module temperature ($T_{module,\ modeled}$), the measured ambient temperature ($T_{ambient,\ measured}$), and the range of the measured module temperatures ($T_{mod,\ measured\ 1\text{-}5}$), in a snow melting period for the modules installed in portrait orientation.

in snow cover. The observed signatures illustrated in Figs. 2 and 3 are representative for the time periods after snow falls for the whole dataset. The electrical signatures for the days where exact snow coverage is measured through imaging, also support these observations.

3.1.2. Simulated electrical snow signatures

Fig. 4 shows the simulated losses in current, voltage and power as a function of snow coverage for modules installed in landscape and portrait orientation. The transmittance of the snow cover is 0 (opaque). As a function of snow coverage, the loss in current follows a simple relationship: If at least one cell in all module substrings is covered, the loss is 100%, and if there is at least one snow-free module substring, the loss is zero. The voltage loss is dependent on the snow free area (as seen for modules in portrait orientation with shading orthogonal to the substring current), but also on activation of bypass diodes (as seen for the modules in landscape orientation with shading in parallel with the substring current). As discussed in Section 2.2.2, we can expect additional electrical losses in the measured data, depending on the inverter and MPPT efficiency at low irradiance, low voltage and partial shading. While the trend in the simulated voltage losses looks similar to what is seen for the measured data in Fig. 2, the losses in current recovers more gradually in Fig. 2 than what is seen in Fig. 4.

Snow covers with increasing transmittance could explain the gradual recovery of the current and power losses seen in Fig. 2. Fig. 5 shows the simulated electrical losses for fully covered modules as a function of snow transmittance. While the current loss is linearly dependent on the transmittance, the losses in voltage are almost recovered as soon as the cells are irradiated.

Fig. 6 gives an example of the combined impact of snow transmittance and coverage, showing how the electrical losses vary with snow transmittance when half of the module is covered. For modules in portrait orientation, the current is still linearly dependent on the transmittance, but in voltage a gain is observed because 50% of the cells are fully irradiated. For the modules in landscape orientation, it is seen that at low transmittance, the shaded module substrings are bypassed giving zero loss in current and 66% loss in voltage. When the transmittance increases and the current loss in the snow-covered module substrings are reduced, the bypass diodes are no longer active resulting in loss in current and zero voltage loss.

While the simulations might explain the trends seen in Fig. 2, they do not explain the variation in losses between different system units. This can, however, be explained by nonuniformity of the snow cover on the system. It is observed that during the process where the snow clears of the modules, there can be variation in both snow coverage and thickness. The total losses are therefore also influenced by the distribution of shading and the configuration of series and parallel connections in the system, as this will affect the maximum power points of the different subarrays.

The impact of both the snow coverage and transmittance on the losses, illustrated in Figs. 4–6, together with the potential nonuniformity of these parameters, explain the trends and the large variations in measured electrical losses during melting shown in Fig. 2. This shows that the assumption of an opaque snow cover in all situations, as is often done in snow loss modeling, is a simplification.

3.1.3. Snow transmittance measurements

To investigate if the transmittance of the snow cover can be high enough to explain the field data observations shown in Fig. 2 as suggested in Section 3.1.2, the transmittance of the snow cover at different thicknesses was measured at the test site. As shown in Fig. 7 and discussed in (Andenæs et al., 2018; Perovich, 2007; Skomedal, 2017), at snow depths less than about 2 cm, transmittance of more than 10%

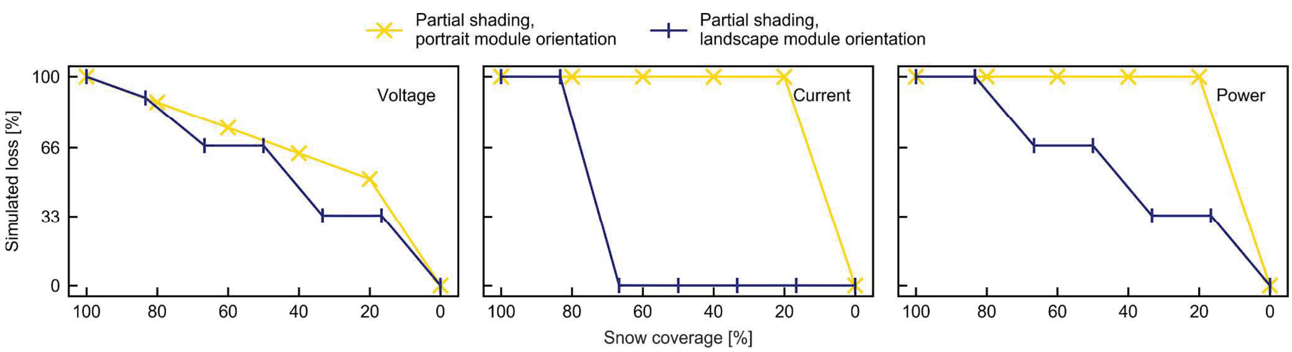


Fig. 4. Simulated losses in voltage, current and power at irradiance of 450 W/m^2 for varying snow coverage with zero transmittance, shown for modules installed in both portrait and landscape orientation.

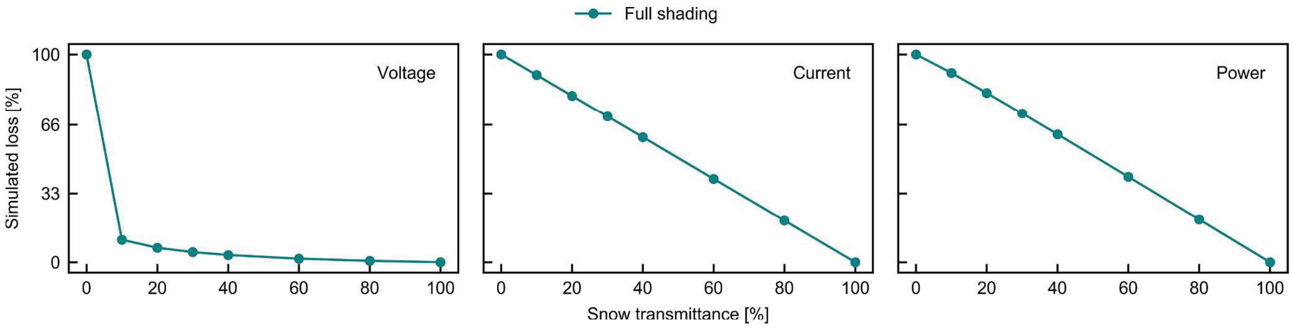


Fig. 5. Simulated losses in voltage, current and power at irradiance of 450 W/m^2 for a full snow cover with varying transmittance.

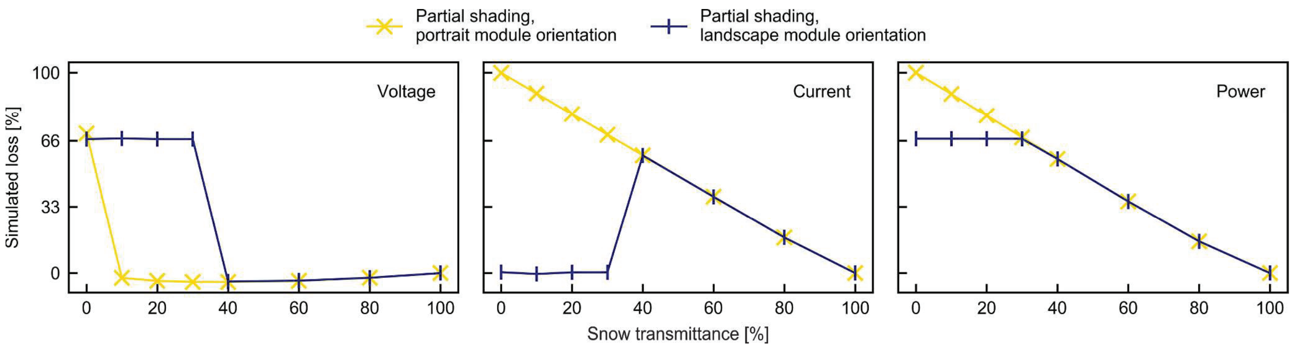


Fig. 6. Simulated losses in voltage, current and power at irradiance of 450 W/m^2 for 50% snow cover with varying transmittance, shown for modules installed in both portrait and landscape orientations.

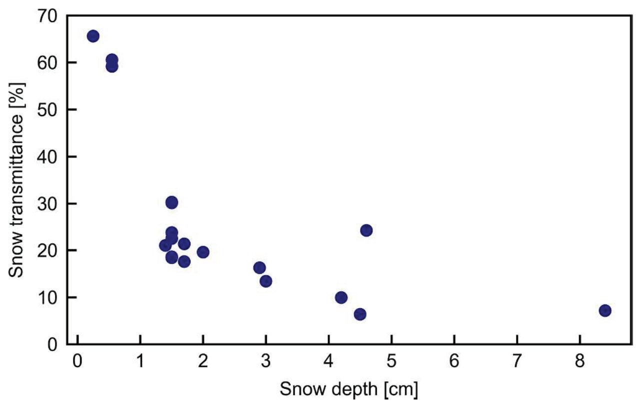


Fig. 7. Measured transmittance of the snow cover on a full size test module glass as a function of snow depth.

might occur. The optical properties of snow is depending on type of snow (Andenæs et al., 2018), so some variations will be expected, but combined with high irradiance, generation of voltage and current is possible at thin snow covers. This is also observed in the electrical measurements of the modules. In situations where the measured snow cover transmittance and irradiance are high, normal voltage values and high current losses are measured. Typically, this is seen in the voltage losses, which for days with snow cover can be very high in the morning and afternoon, and down to zero in the middle of the day when the irradiance is high.

3.1.4. Snow signature overview

To summarize the results discussed in Sections 3.1.1–3.1.3, an overview of the impact of different types of snow cover on measured system variables and the overall PV plant behavior, is given in Table 1.

A full opaque snow cover leads to 100% loss in all the electrical parameters. This could be interpreted as an inverter breakdown. To separate the two cases, development in additional parameters such as snow depth and module temperature should be utilized.

When the snow cover is semitransparent and/or partial the situation is more complex, a wider range of outcomes are possible, and larger

Table 1
Overview of different PV parameters for a c-Si monofacial system, and how they are affected by different types of snow covers (full or partial, opaque or semi-transparent) for modules in portrait and landscape orientation.

	Full, opaque	Full, semitransparent	Partial, opaque		Partial, semitransparent	
			Portrait	*Landscape*	*Portrait*	*Landscape*
Module temperature	<< Normal operating temperature		Typically < Normal operating temperature			
DC current	0	Low*	0	Normal	Low	Low-normal*
DC voltage	0	Normal	Low	Low-medium*	High	Low-normal*
Power	0	Low*	0	Low-medium*	Low	Low-medium*
PV plant	No production	All inverters have low/0* power	Many or all of the inverters have low power. There may be large variations in power, current, voltage, and module temperatures.			

* Depending on snow coverage and/or transmittance.

variation within the PV plant is seen. The response in current, voltage and module temperature will depend on both transmittance, size and nonuniformity of the snow coverage, and also on module orientation for crystalline silicon modules. The voltage will be recovered as soon as the cells in the modules are irradiated, either because of clearing of the snow cover or increased transmittance. The maximum current in each module substring will be limited by the least irradiated cell. If partial snow covers lead to variations in irradiance for different module substrings, the bypass diodes in the modules can activate. Because of this, voltage losses during snow covers are characteristic for systems where the modules are installed in landscape orientation, and the typical snow shading is in parallel to the module substrings. No loss in voltage would require snow covers with both high uniformity and high transmittance, which is possible, but not very common. How many bypass diodes that are active depends on the transmittance and uniformity of the snow cover. For module substrings with snow covers with high transmittance, active bypass diodes and voltage losses can lead to larger power losses than when the diodes are not active, as illustrated in Fig. 6. If one shaded module substring is not bypassed, this gives a loss in current in addition to the voltage loss, as observed for the commercial system in Fig. 2.

3.2. *Snow loss model evaluation*

The snow loss models described in Section 2.3 were evaluated for the commercial system (landscape orientation). The loss was measured for all the inverters in the system to capture eventual variations in snow losses for identical system configurations. Due to its small size and severe system shading for some parts of the winter, the test system was not found suitable for model evaluation.

3.2.1. *Evaluation of empirical snow loss models*

The models built on empirical correlations between ambient conditions and losses, failed to estimate snow losses satisfactorily, particularly when there were differences in ambient conditions between the tested dataset and the dataset the model was based on. For the model suggested by Andrews and Pearce (2012), the R^2 of the relationship between the power loss and the suggested explanatory parameters was 0.24, showing a low correlation. The snow data in this model is limited to snow fall data from the two previous days. For the dataset in this study, however, snow covers can in some cases last longer than a month. For the simple model for yearly relative losses suggested by Powers et al. (2010), the modeled losses were 2.3–5% compared to measured losses of 2.2–11.2%. For most years, the difference between measured and modeled losses was below 1 percentage point, but for the year with largest losses, the difference was 6.2 percentage points. Different ambient conditions might also here be influential: in Truckee, California, where the model is developed, the difference in total irradiation from summer to winter is lower than for the data in this study because of the difference in latitude. For the Norwegian location, the irradiance changes a lot through the year, and the time of the snow cover also influences the total losses, as snow cover in the middle of the winter will have less impact on the annual losses than a springtime snow cover. The second model developed by Townsend and Powers (2011), had a mean absolute error in the estimation of relative monthly snow losses of 23%.

3.2.2. *The Marion snow loss model*

The empirical models can be used to give rough estimates of the losses, but for models based on a few datasets, it appears to be difficult to capture all aspects of snow covers and resulting PV losses and develop accurate and transferable models. Modeling different aspects of snow covers and losses separately and aim for modeling of absolute losses, like in the Marion model, was shown to be a more robust and flexible approach, yielding more accurate loss estimations. The threshold defined in Eq. (1). to identify sliding events caused by snow melting, correlated well with melting events found in the snow data. Most melting events, and all large melting events, could be predicted by the conditions defined in Eq. (1). The default sliding coefficient in pvlib (0.20), estimated for roof mounted systems, was however observed to be too high. This coefficient is expected to depend on different system and module designs, because technical aspects can either promote or obstruct snow sliding (Burnham et al., 2020). Frameless modules (Riley et al., 2019), empty space below modules (Heidari et al., 2015), and heating on the rear side of the module (Ross, 1995) (e.g. from reflected irradiance – in particular for bifacial modules (Burnham et al., 2019), or the building if roof mounted) will promote sliding, for instance. In the studied case, where the modules are installed on a flat, well-insulated roof, and there is no empty space where sliding snow can accumulate below the modules, high roof interference and a low sliding coefficient is expected (Heidari et al., 2015). Generally, when the snow depth is increasing, the empty space below the modules will decrease, giving increased ground/roof interference. In this case, because the modules are not elevated, how much the snow can slide down the module surface will also decrease with increasing snow depth. The top of the modules is approximately 30 cm above the rooftop. With snow depths above 30 cm, the system will be fully submerged in snow and there will be no sliding. Snow depths above 30 cm are rare for the tested system, as shown in Fig. 10, but the observed snow depths do often lead to situations where the system is partly submerged in snow, reducing the possibility for snow sliding. Melting is therefore most likely an important snow clearing mechanism in the tested system, a process that typically is slower than sliding for thick snow layers. Fig. 8 shows for different *sc* values, for periods with snow depth $>$ 3 cm, how measured snow loss correlate with modeled snow cover, and the correlation between measured and modeled daily snow loss. Fig. 8 a) shows how a larger fraction of the timestamps with measured snow loss correlate with timestamps with modeled snow cover, both for high and lower measured snow loss, when using a lower sliding coefficient. As shown in Fig. 8 b), reducing the sliding coefficient gives a better fit between measured and modeled daily losses. Here, because the Marion model assumes zero production from partly covered module substrings and a uniform snow cover, the modeled loss is stepwise, and the only possible outcomes are 0, 33%, 66% or 100% loss. As shown in Fig. 2, the measured power loss has a wider range of outcomes. Some of the variations in the measured losses, can also be caused by the high relative errors in the modeled daily expected power for parts of the winter periods.

The data show that for thin snow covers, however, snow clearing happened significantly faster. There is more room for snow to slide down the module surface, and thin snow covers are also more likely to melt directly on the module, a process that for thin snow covers is faster than sliding (Andrews et al., 2013; Pawluk et al., 2019). Additionally, as thin snow covers have higher transmittance, heating of the module that can aid the melting is expected (Pawluk et al., 2019). As shown in Fig. 9, when the measured snow depth is low, the sliding coefficient that most consistently models 100 or 66% loss in periods with high losses, and 0 or 33% in periods with low losses, is 0.4, which is higher than what was seen in Fig. 8. For the test system it is also observed that the sliding coefficient seems to be influenced by the snow conditions. In Fig. 9, it is also seen that the modeled losses for thin snow covers shows a poorer fit with measured loss compared to thicker snow covers. Thin snow covers have also previously been shown to introduce noise in loss modeling (Andrews and Pearce, 2012). In addition to the challenge of exact estimation of snow coverage, snow transmittance is, as previously discussed, playing a role for thin snow covers and might challenge the loss estimation.

As shown in Fig. 10, reducing the sliding coefficient to 0.06 compared to the default sliding coefficient in pvlib of 0.20, gave a better fit between measured and modeled losses for most years. The exception is 2017, a year with very low snow depths. Also shown in Fig. 10, introducing separate sliding coefficients (or more general: snow *clearing* coefficients) for snow depths above and below 3 cm yields an even better fit with the total measured losses. With the default *sc* the total modeled absolute snow loss for the five years of data was underestimated by 23%,

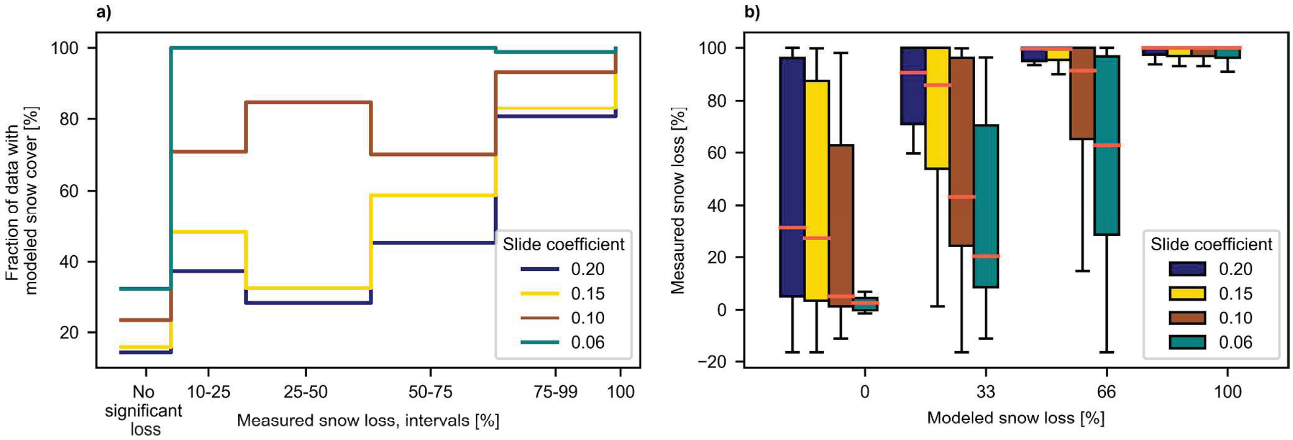


Fig. 8. For periods with snow depths > 3 cm and with four different sliding coefficients used in snow cover/loss modeling, a) the fraction of data for different power loss intervals with modeled snow cover > 0 (due to the uncertainty in the modeling of the expected power, 10% is set as the lower limit of significant snow loss), and b) the variation in daily measured loss at the different modeled loss values (four possible outcomes: 0, 33, 66 and 100%). The boxes extend from the first to the third quartile values of the data, with a line on the median. The whiskers extend to the maximum or minimum value within 1.5 times the interquartile range, and outliers are not included.

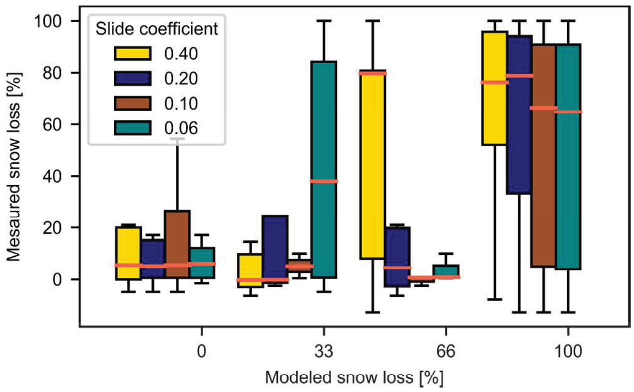


Fig. 9. For periods with snow depths < 3 cm and with four different sliding coefficients used in snow cover/loss modeling, the variation in daily measured loss at the different modeled loss values (four possible outcomes: 0, 33, 66 and 100%). The boxes extend from the first to the third quartile values of the data, with a line on the median. The whiskers extend to the maximum or minimum value within 1.5 times the interquartile range and outliers are not included.

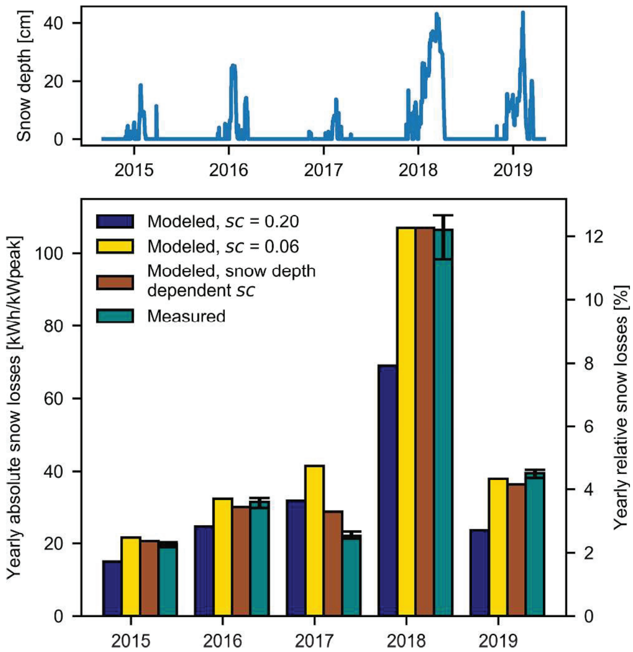


Fig. 10. Measured (plant median) and modeled yearly absolute and relative snow losses. The relative losses are calculated based on the mean yearly expected energy generation in the analysis period. The error bar shows the range of the measured losses in the PV plant. The snow losses are modeled in pvlib with default sliding coefficient (0.20), reduced sliding coefficient (0.06) and a snow depth dependent coefficient: 0.4 for snow depth < 3 cm, and 0.06 for snow depths > 3 cm. The corresponding snow depth measurements are also shown.

with the reduced *sc* (0.06) the losses were overestimated by 11%, and with the snow depth dependent *sc* the model overestimated by 3%, yielding a significant improvement to the model. Relative to the mean yearly energy generation in the analysis period, the differences in measured and modeled losses when using the model with snow depth dependent *sc*, was between −0.8 and 0.3 percentage points.

It would still be expected that snow loss modeling is still not exact on high time resolutions even with improved sliding coefficients, both due to challenges with estimating the snow coverage, the transmittance and non-uniformity of the snow coverage, and the difficulty of accurately quantifying the effect the snow has on the PV production. It can, however, be used to assess the probability of snow cover on the modules and give reasonable snow loss estimates for yield estimations which are typically aggregated to lower time resolutions.

3.3. Snow detection

The observed snow signatures in the data and the improved snow model are promising starting points for building snow detection algorithms for monitoring purposes, failure diagnosis and performance loss analysis. While snow loss modeling has too low accuracy on high time resolutions to directly model losses in monitoring, the improved snow cover model suggested in Section 3.2 can be used to indicate the possibility of snow-covered modules, as shown in Fig. 8 a). For the tested commercial system, the loss in voltage is the signature that in most cases is connected to snow loss, as discussed in Section 3.1.4. Fig. 11 shows, for different power loss intervals, how large share of the data that would be labeled as snow, given a snow detection criterion of: 1) voltage loss between 10% and 100%, 2) modeled snow cover larger than 0, 3) either criterion 1 or 2. The data has 5-minute resolution and is taken from

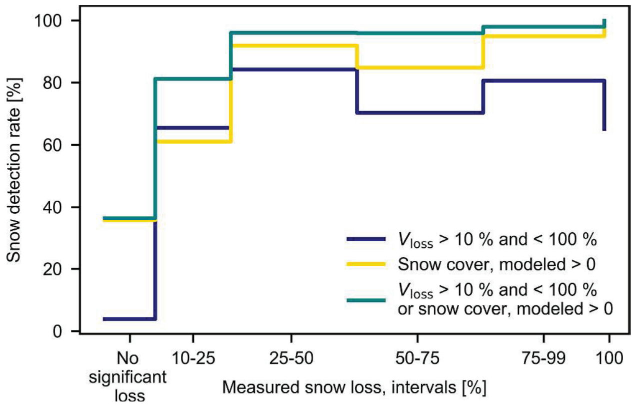


Fig. 11. Snow detection rate for different intervals of measured snow loss for the following criteria: 1) voltage loss between 10% and 100%, 2) modeled snow cover greater than zero, and 3) either 1 or 2. Only data points from periods with snow on the ground and irradiance above 50 W/m^2 are considered. Due to the uncertainty in the modeling of the expected power, 10% is set as the lower limit of significant snow loss.

periods with irradiance above 50 W/m^2 and snow on the ground. For the third criterion, 97% of the snow losses above 10% is labeled as snow. The detection rate is higher at high snow loss. At full snow covers giving 100% loss, the detection rate is 100%. In periods with a measured power loss smaller than 10%, i.e. no significant snow loss, 38% of the data points are labeled as snow, which we interpret as false positives. These false positives are mostly related to the uncertainty in the snow loss modeling for thin snow covers and during melting, causing the model to indicate snow cover in periods where the snow has been cleared. A consequence of false positives in snow detection could be that actual system faults are falsely labeled as snow losses. Snow loss modeling is consequently best used to indicate the probability of snow cover.

To improve snow detection, more of the snow signatures described in Section 3.1 could be included. The module temperature measurements and the duration and evolution of the snow signatures could e.g. be taken into account. Snow losses, especially during snow melting, can change significantly from day to day and within a day, in a different way than typical system faults. The results suggest that due to the high rate of the data with losses that correctly are identified as snow, the snow detection method will improve fault detection and diagnosis as well as loss analysis, and that further improvements could be achieved by including more of the identified snow signatures and by using snow loss modeling to indicate probability of snow cover.

4. Conclusions

In this paper we describe the effect of different types of snow cover on PV energy generation, and snow related signatures in PV monitoring data are identified. In addition to snow coverage and system configuration, transmittance and nonuniformity of the snow cover influence the total snow losses, increasing the complexity in snow loss modeling. Existing snow loss models are evaluated. Three of the models are purely empirical, and power loss is directly estimated based on system and weather data. In the Marion model empirical correlations are used to model different effects causing natural snow clearing, and snow coverage and the resulting loss are modeled separately. We find that the purely empirical models are less general and flexible than the Marion model.

For the evaluated system with low tilt modules on a flat roof, the natural snow clearing rate is observed to be much faster for thin snow covers (<~2–3 cm) than for thicker snow covers. This difference in the snow clearing process between thin and thick snow covers is assumed to be especially large for the evaluated system. This is because there is little available space where the snow can slide away, leading to slow snow clearing for thick snow covers. The evaluated system design is, despite this, very common in the Nordic countries because it gives increased roof coverage and more even energy generation throughout the year. By including the effect of snow depth dependent snow clearing in the Marion model, we achieve reduced uncertainty in the modeled snow losses, allowing more accurate energy yield assessment for new PV systems. We also find that the identified snow data signatures and the improved Marion snow model can be used to detect and separate snow losses from other phenomena that affects PV production, such as faults. This is important for the development of PV in cold climate areas that are prone to snow.

We discuss how different system designs can promote or obstruct snow clearing, and we find that for the tested system the snow clearing rate is lower than for the systems the snow sliding/clearing coefficients in the Marion model is based on. Future work should therefore include further validation of the snow clearing coefficients for different system designs. Additionally, as snow is a complex weather phenomenon, validation of the improved model and the snow detection for larger datasets and different environments is necessary. The aspect of the evolution of the snow data signatures with time should also be further investigated to improve snow detection.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors acknowledge funding from the Innovation project number 803801 (Autonomous monitoring, control and protection of renewable energy infrastructure) and RCN 282357 - Smart styring av næringsbygg med lokal strømproduksjon.

References

Aarseth, B.B., Øgaard, M.B., Zhu, J., Strömberg, T., Tsanakas, J.A., Selj, J.H., Marstein, E. S., 2018. Mitigating snow on rooftop PV systems for higher energy yield and safer roofs. In: Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition, pp. 1630–1635. https://doi.org/10.4229/35thEUPVSEC2018-6CO.3.5.

Andenæs, E., Jelle, B.P., Ramlo, K., Kolås, T., Selj, J., Foss, S.E., 2018. The influence of snow and ice coverage on the energy generation from photovoltaic solar cells. Sol. Energy 159, 318–328. https://doi.org/10.1016/j.solener.2017.10.078.

Andrews, R.W., Pearce, J.M., 2012. Prediction of energy effects on photovoltaic systems due to snowfall events. In: 2012 38th IEEE Photovoltaic Specialists Conference. IEEE, pp. 3386–3391. https://doi.org/10.1109/PVSC.2012.6318297.

Andrews, R.W., Pollard, A., Pearce, J.M., 2013. The effects of snowfall on solar photovoltaic performance. Sol. Energy 92, 84–97. https://doi.org/10.1016/j.solener.2013.02.014.

Bashir, N., Irwin, D., Shenoy, P., 2020. DeepSnow: Modeling the impact of snow on solar generation. In: Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. pp. 11–20. https://doi.org/10.1145/3408308.3427620.

Becker, G., Schiebelsberger, B., Weber, W., Vodermayer, C., Zehner, M., Kummerle, G., 2006. An approach to the impact of snow on the yield of grid connected PV systems. Proceedings of the 21st European Photovoltaic Solar Energy Conference and Exhibition.

Bosman, L.B., Darling, S.B., 2018. Performance modeling and valuation of snow-covered PV systems: examination of a simplified approach to decrease forecasting error. Environ. Sci. Pollut. Res. 25, 15484–15491. https://doi.org/10.1007/s11356-018-1748-1.

Burnham, L., Riley, D., Braid, J., 2020. Design considerations for photovoltaic systems deployed in snowy climates. In: Proceedings of the 37th European Photovoltaic Solar Energy Conference and Exhibition, pp. 1626–1631.

Burnham, L., Riley, D., Walker, B., Pearce, J.M., 2019. Performance of bifacial photovoltaic modules on a dual-axis tracker in a high-latitude, high-albedo environment. In: 2019 46th IEEE Photovoltaic Specialists Conference. IEEE, pp. 1320–1327. https://doi.org/10.1109/PVSC40753.2019.8980964.

Dobos, A.P., 2012. An improved coefficient calculator for the California energy commission 6 parameter photovoltaic module model. J. Sol. Energy Eng. 134 https://doi.org/10.1115/1.4005759.

FME Susoltech, 2020. Interactive map of monitored PV installations in Norway, IFE test site, Kjeller. [WWW Document]. URL https://susoltech.no/solar-panel-map/ (accessed 12.22.20).
Gilman, P., Dobos, A., DiOrio, N., Freeman, J., Janzou, S., Ryberg, D., 2018. SAM photovoltaic model technical reference update, No. NREL/TP-6A20-67399. Golden.
Hashemi, B., Cretu, A.M., Taheri, S., 2020. Snow loss prediction for photovoltaic farms using computational intelligence techniques. IEEE J. Photovoltaics 10, 1044–1052. https://doi.org/10.1109/JPHOTOV.2020.2987158.
Heidari, N., Gwamuri, J., Townsend, T., Pearce, J.M., 2015. Impact of snow and ground interference on photovoltaic electric system performance. IEEE J. Photovoltaics 5, 1680–1685. https://doi.org/10.1109/JPHOTOV.2015.2466448.
Holmgren, W.F., Hansen, C.W., Mikofski, M.A., 2018. pvlib python: a python package for modeling solar energy systems. J. Open Source Softw. 3, 884. https://doi.org/10.21105/joss.00884.
IEA, 2020. PVPS Trends in photovoltaic applications 2020.
Jäger-Waldau, A., 2020. Snapshot of photovoltaics—February 2020. Energies 13.
King, D.L., Kratochvil, J.A., Boyson, W.E., 2004. Photovoltaic array performance model, No. SAND2004-3535. Albuquerque. https://doi.org/10.2172/919131.
Lorenz, E., Betcke, J., Drews, A., de Keizer, A.C., Stettler, S., Scheider, M., Bofinger, S., Beyer, H.G., Heydenreich, W., Wiemken, E., van Sark, W., Toggweiler, P., Heilscher, G., Heinemann, D., 2007. Intelligent performance check of PV system operation based on satellite data (PVSAT-2), final technical report.
Lorenz, E., Heinemann, D., Kurz, C., 2011. Local and regional photovoltaic power prediction for large scale grid integration: Assessment of a new algorithm for snow detection. Prog. Photovoltaics Res. Appl. 20, 760–769.
Marion, B., Schaefer, R., Caine, H., Sanchez, G., 2013. Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations. Sol. Energy 97, 112–121. https://doi.org/10.1016/j.solener.2013.07.029.
Norsk Klimaservicesenter, 2020. Observasjoner og værstatistikk [WWW Document]. URL https://seklima.met.no/observations/ (accessed 12.22.20).
NREL, 2020. PySAM.
NVE, 2019. seNorge [WWW Document]. URL www.senorge.no (accessed 9.1.19).
Øgaard, M.B., Haug, H., Selj, J., 2018. Methods for quality control of monitoring data from commercial PV systems. In: Proceedings of the 35th European Photovoltaic Solar Energy Conference and Exhibition, pp. 2083–2088. https://doi.org/10.4229/35thEUPVSEC20182018-6DV.1.53.
Øgaard, M.B., Riise, H.N., Haug, H., Sartori, S., Selj, J.H., 2020. Photovoltaic system monitoring for high latitude locations. Sol. Energy 207. https://doi.org/10.1016/j.solener.2020.07.043.
Pawluk, R.E., Chen, Y., She, Y., 2019. Photovoltaic electricity generation loss due to snow – A literature review on influence factors, estimation, and mitigation. Renew. Sustain. Energy Rev. 107, 171–182. https://doi.org/10.1016/j.rser.2018.12.031.
Perovich, D.K., 2007. Light reflection and transmission by a temperate snow cover. J. Glaciol. 53, 201–210. https://doi.org/10.3189/172756507782202919.
Powers, L., Newmiller, J., Townsend, T., 2010. Measuring and modeling the effect of snow on photovoltaic system performance. In: 2010 35th IEEE Photovoltaic Specialists Conference. IEEE, pp. 973–978. https://doi.org/10.1109/PVSC.2010.5614572.
PVGIS, 2020. Data sources and calculation methods [WWW Document]. URL https://ec.europa.eu/jrc/en/PVGIS/docs/methods (accessed 12.22.20).
PVsyst, 2020. Soiling loss [WWW Document]. URL https://www.pvsyst.com/help/index.html (accessed 12.22.20).
Riley, D., Burnham, L., Walker, B., Pearce, J.M., 2019. Differences in snow shedding in photovoltaic systems with framed and frameless modules. In: 2019 46th IEEE Photovoltaic Specialists Conference. IEEE, pp. 558–561. https://doi.org/10.1109/PVSC40753.2019.8981389.
Ross, M., 1995. Snow and ice accumulation on photovoltaic arrays: An assessment of the TN conseil passive melting technology, Division Report EDRL 95-68 (TR). Varennes.
Ryberg, D., Freeman, J., 2017. Integration, validation, and application of a PV snow coverage model in SAM, NREL/TP-6A20-68705. Golden.
Schill, C., Brachmann, S., Koehl, M., 2015. Impact of soiling on IV-curves and efficiency of PV-modules. Sol. Energy 112, 259–262. https://doi.org/10.1016/j.solener.2014.12.003.
Skomedal, Å., 2017. The transmittance of light through snow; an initial study for solar energy systems. https://doi.org/10.13140/RG.2.2.10539.54568.
Solargis, 2016. Solargis pvPlanner User Manual.
Souka, A.F., Safwat, H.H., 1966. Determination of the optimum orientations for the double exposure flat-plate collector and its reflections. Sol. Energy 10, 170–174.
Stein, J.S., Farnung, B., 2017. PV performance modeling methods and practices, Report IEA-PVPS T13-06:201.
Teubner, J., Buerhop, C., Pickel, T., Hauch, J., Camus, C., Brabec, C.J., 2019. Quantitative assessment of the power loss of silicon PV modules by IR thermography and its dependence on data-filtering criteria. Prog. Photovoltaics Res. Appl. 27, 856–868. https://doi.org/10.1002/pip.3175.
Townsend, T., Powers, L., 2011. Photovoltaics and snow: An update from two winters of measurements in the Sierra. In: 2011 37th IEEE Photovoltaic Specialists Conference. IEEE, pp. 3231–3236. https://doi.org/10.1109/PVSC.2011.6186627.
Tsanakas, J.A., Ha, L., Buerhop, C., 2016. Faults and infrared thermographic diagnosis in operating c-Si photovoltaic modules: A review of research and future challenges. Renew. Sustain. Energy Rev. 62, 695–709. https://doi.org/10.1016/j.rser.2016.04.079.
van Noord, M., Berglund, T., Murphy, M., 2017. Snöpåverkan på solelproduktion om snöförluster på takanläggningar i Norra Sverige., Rapport 2017:382.
Wirth, G., Schroedter-Homscheidt, M., Zehner, M., Becker, G., 2010. Satellite-based snow identification and its impact on monitoring photovoltaic systems. Sol. Energy 84, 215–226. https://doi.org/10.1016/j.solener.2009.10.023.

Paper VII

Estimation of snow loss for photovoltaic plants in Norway

M.B. Øgaard, H.N. Riise, and J.H. Selj

In: *Proceedings of the 38th European Photovoltaic Solar Energy Conference and Exhibition* (2021), pp. 1081-1087.
DOI: 10.4229/EUPVSEC20212021-5DO.4.5

VII

ESTIMATION OF SNOW LOSS FOR PHOTOVOLTAIC PLANTS IN NORWAY

Mari B. Øgaard[1,2*], Heine N. Riise[2], and Josefine H. Selj[1,2]
[1]Department of Technology Systems, University of Oslo, Kjeller, Norway
[2]Department of Renewable Energy Systems, Institute for Energy Technology, Kjeller, Norway
*Corresponding author. E-mail: mari.ogaard@ife.no, Tel.: +47 976 356 08,
Adr.: Instituttveien 18, 2007 Kjeller, Norway

ABSTRACT: Large PV plants are increasingly common in locations with colder climates where snow can lead to significant PV power loss. For these locations, estimates of snow loss is necessary for accurate PV yield modeling. Robust estimation of snow loss is, however, challenging. Snow-induced loss is expected to vary with climate, weather, and PV plant design. In this work, we estimate snow loss from historical data for a set of PV plants in Norway. To extend the snow loss dataset, 12 years of weather data and a modified adaption of the Marion snow loss model are used to simulate snow loss for the analyzed PV plants over time. For the historical data, we observe variations in annual losses for the same system of more than 10 percentage points. For some of the systems, we find losses in a range from 0 to 100 % for the same month. As expected, systems with colder climates have higher loss than systems in warmer climates, and systems with higher tilt has lower loss than systems with lower tilt. With snow loss modeling we get improved understanding of typical and extreme values, and the potential inter-annual variation in monthly and annual snow loss.

Keywords: PV System, System Performance, Modeling, Soiling, Snow

1 INTRODUCTION

As cost reductions have made photovoltaics (PV) a favorable choice also in colder climates, deployment rates in regions with snow falls are rapidly increasing [1–3]. Snow on PV modules may lead to significant power loss. For certain locations snow fall can result in zero electricity production in the winter season and more than 30 % annual loss [4]. Consequently, it is an important loss mechanism to consider in PV system models to get accurate assessments of the expected energy generation from PV plants in snow-affected locations. Snow-induced PV power loss is expected to vary from year to year, between different system configurations and between different locations. To get accurate snow losses for a specific system, a model taking into account the different influential parameters is therefore necessary. Recent research has demonstrated that for snow-affected locations the uncertainty in yield estimations [5–8] and forecasting [9] can be reduced if snow loss models are included. Despite this, snow loss models are often not implemented in PV simulation software. The System Advisor Model (SAM) has implemented the model suggested by Marion et al. [5,6], but in other software, snow is either not considered [10] or estimated by constant soiling values [11] with little guidance on how these constant values should be obtained.

Accurate snow loss modeling is, however, challenging, because the parameters influencing the snow cover and resulting PV system loss are manifold. The influential parameters range from weather conditions (precipitation, temperature, irradiance, wind, etc.), to installation and technology specific configurations (tilt, module technology/orientation, objects obstructing snow sliding etc.) [1,12] and type of snow [4]. Multiple snow loss models have been suggested [4], but validation is typically lacking [6]. To include all the parameters influencing snow cover and resulting loss in a physical model is challenging, and most suggested models for PV snow loss are based on empirical correlations [4].

In our previous work [13], we show that the snow loss model suggested by Marion et al. [3], where empirical correlations are used to model natural snow clearing, performs better than models where snow loss is directly estimated based on empirical correlations between power loss and system and weather data. Ryberg et al. [6] and van Noord et al. [14] also find acceptable correlation between estimated and modeled snow loss using the Marion model.

To estimate the snow coverage on PV modules, the Marion model aims to predict: 1) presence of snow cover on PV modules, 2) when snow is cleared off the modules, and 3) the snow clearing rate. The separation of these three processes in the model, enables improvement of the model by developing the modeling of each process by either using additional physical modeling or collecting more empirical data. In the model, the snow clearing rate is estimated with an empirical snow clearing coefficient. Many different parameters related to system design and weather/snow conditions are assumed to impact how fast the snow is cleared [13]. Frameless modules [15], empty space below modules [12] will promote sliding, for instance. With more data from different system configurations in different climates, we would get improved understanding of which parameters that impact the snow clearing rate the most, and consequently also get better values for the snow clearing coefficient and its potential variation.

In our evaluation of the model [13], we estimate the snow clearing coefficient from the snow loss data for the analyzed system, and we observe that for thin snow covers, the natural snow clearing rate is faster than the clearing rate of thicker covers [13]. By introducing separate snow clearing coefficients for thin and thick snow covers, reduced error in modeled snow loss is achieved. This also seems to make the model more general: when using snow depth dependent snow clearing coefficients we get better results when we model losses for systems with similar technical configurations with the same coefficients compared to when we use one single coefficient [16]. For transferability, it is important that we can use the same empirical coefficients for systems with similar technical configurations.

In addition to the challenge of accurate snow loss modeling, there is a lack of established guidelines on how to take snow losses into account when used in e.g. PV yield modeling or PV system dimensioning. Input data for the snow loss estimation, temporal resolution of the loss parameter, inter-annual variations and the impact of climate change need to be discussed. As pointed out by Marion et al. [5], typical meteorological year (TMY) values are not sufficient to use as input in snow loss modeling for PV yield assessments. Because snow is not one of the parameters considered in the derivation of TMY, TMY data does not necessarily represent a typical snow year. Using a long time series of meteorological data, enabling quantification of typical values and the inter-annual variability is suggested instead [5]. It is, however, important to use recent data. Because of climate change, historical snow data might not be representative for future snow conditions. In Norway, it is estimated that climate change will lead to reductions in snow depth and length of snow season, and an increase in snowline elevation [17]. Temporal resolution of the snow loss parameter is to our knowledge not much discussed in the literature. In the simulation tool PVsyst, monthly constant snow losses are used for PV simulations [11]. While this can be sufficient in assessments of total yield, this will not sufficiently describe the potential inter- and intraday variation. This variation can be relevant in system dimensioning, in particular for hybrid/battery systems.

In this work, we estimate the snow loss for a set of PV plants in Norway. Two different system designs are evaluated: commercial systems with modules installed with low tilt angles on flat roofs, and residential systems on tilted roofs. The aim of this analysis is to describe the variations in both monthly and annual snow losses, with respect to both time, location and system configuration, and to discuss how this could be included in e.g. PV yield modeling. The losses are estimated using both historical data and simulations based on longer time series of weather data and a modified adaption of the Marion snow loss model.

2 METHODOLOGY

2.1 PV system data

Seven PV installations in Norway with a total installed capacity of 1.6 MWp are analyzed. The evaluated dataset is the same as the dataset used to validate the modified snow loss model in [16], but some of the data series are extended in time. Two different system types are evaluated: residential systems on tilted roofs, and commercial large-scale systems on flat roofed buildings. The commercial systems have modules installed with low tilt and east/west orientation. This configuration is not optimal for total annual production in Norway, but is commonly used on flat roofed buildings to increase the packing density and reduce the seasonality of the production profile. The modules are installed in portrait orientation at the residential systems, and landscape orientation at the commercial system. All the PV modules are crystalline silicon. Apart from some variations in exact orientation, and tilt for the residential systems, the installations of the same type are assumed technically identical. Tilt and length of analysis period for the systems are given in Table I.

Table I: Module tilt and length of analysis period for analyzed systems

System ID	Tilt	Analysis period
Residential systems		
R1	26	Jan 2019 – June 2021
R2	40	Jan 2018 – June 2021
R3	24	Jan 2019 – June 2021
Commercial systems		
C1	10	Jan 2015 – June 2021
C2	10	Jan 2017 – June 2021
C3	10	Jan 2018 – June 2021
C4	10	Jan 2018 – June 2021

The measured energy of the PV systems is collected from the inverters. For the commercial systems, the effective in plane irradiance and the module temperature is measured by reference cells. The residential systems have no on-site sensors. For all the locations, snow depth and snow fall data are collected from seNorge.no [18] and temperature and global horizontal irradiation (GHI) data are collected from nearby weather stations [19].

As illustrated in Figure 1, the analyzed systems are situated in three different geographic regions in Norway (East, West and Central), and in three different Köppen-Geiger (KG) [20] climate zones (Humid continental climate (Dfb), subarctic climate (Dfc) and oceanic climate (Cfb)). This gives variation in snow and weather conditions between the locations. Figure 2 shows 16 years of snow depth data for the four different combinations of geographic region and climate zone.

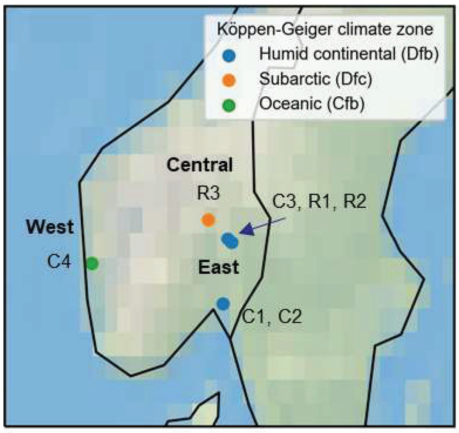


Figure 1: Location on the map for the analyzed systems. The locations are labeled with geographic region and climate zone is given by the marker color.

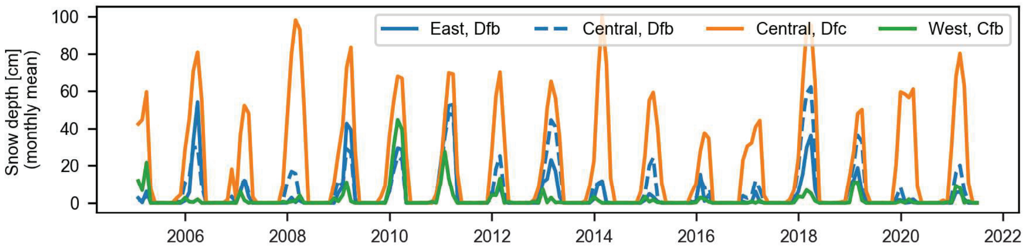


Figure 2: Sixteen years of snow depth data for the four combinations of geographic region and KG climate zone in the analyzed dataset.

2.2 Estimation of snow loss from historical PV data

To estimate historical snow losses from PV monitoring data, it is necessary to get an accurate estimate of what the energy production could have been if there was no snow. This requires an accurate model that considers all other losses of the PV system, and an efficient method to separate snow losses from other losses. To correctly estimate snow losses, it is especially important to take into account other wintertime losses such as losses caused by low irradiance, and high angles of incidence. These types of losses are typical for high latitude locations in the wintertime [21], and can introduce increased uncertainty in PV system modeling if not properly accounted for.

To estimate expected PV module power output for the commercial systems, the effective irradiance measured by the reference cells and the measured module temperature are used as input to a single diode model in pvlib python [22] to model PV module power output, using the procedure described in [13]. For the residential systems, detailed module data and onsite measurements are not available. Effective irradiance and module temperature are modeled in pvlib from measurements of GHI and ambient temperature from nearby weather stations. The GHI measurement is decomposed using the Erbs model [23] to estimate diffuse irradiance, and the Disc [24] model for direct irradiance. When modeling the in plane irradiance for the systems, the Hay and Davies' 1980 model [25] is used to determine the in plane diffuse irradiance from the sky. From the modeled in plane global irradiance, the effective irradiance is calculated by adding reflection losses using an incident angle modifier based on the physical model described in [26]. The module temperature is modeled using the PVsyst temperature model [27]. The expected power output from the modules is modeled using PVWatts [28].

The described PV module power output models do not take into account all the relevant losses (all other losses than snow-induced losses) of the systems. From the energy performance index (*EPI)* of the system, the ratio between measured and modeled energy, we observe that the calculated value is below 1. Additionally, the *EPI* has a systematic seasonal component suggesting higher losses in the winter months, also in periods without snow. We assume that the significant losses not accounted for in the model, can be estimated with a constant and a seasonal component. To accurately find the seasonal components for the analyzed systems, seasonal trend decomposition is performed on the daily *EPI,* after filtering out time periods with snow on the ground (which introduces a non-systematic seasonal component). Seasonal trend decomposition is suggested by [29] as a method to find and correct the seasonal component in PV performance metrics. The deviation between 1 and the median of the seasonally corrected *EPI* is used as an estimate of the constant system losses. These two components are then used to correct the modeled PV module output to find the expected system output. By this way aiming to take all other significant losses into account, the snow loss is then estimated to be the difference in expected system output and measured system output in periods where the snow data suggests snow on the ground.

An additional uncertainty in this methodology is that snow cover on the irradiance sensors can lead to underestimation of snow losses. To reduce this uncertainty, the reference cell measurements from the commercial systems were controlled and corrected by the external GHI data. Pyranometers is expected to have lower risk for full snow cover than reference cells, because of the shape and elevation of the sensor, and better ventilation and maintenance.

2.3 Modeling snow loss with the modified Marion model

In the Marion snow loss model [5] the presence of a new snow cover is assumed to happen after snow fall. The model further assumes that natural snow clearance will happen during melting. Melting is predicted to happen during the following conditions:

$$T_{\text{amb}} > G_{\text{POA}}/m. \qquad (1)$$

T_{amb} is the ambient temperature, G_{POA} is the in plane irradiance and m is an empirically defined value of -80 W/(m^2 °C). During melting, the snow will be cleared by sliding or direct melting on the modules [4]. To estimate the reduction in snow coverage in the melting period, measured in fractions of the system height, the tilt of the modules and an empirical snow clearing coefficient (*sc*) is used:

$$\textit{Snow slide amount} = sc * \sin(tilt). \qquad (2)$$

Based on these assumptions, the snow coverage on the modules is estimated, and the corresponding power loss calculated. If a module substring is partially covered by snow, the power output is assumed to be zero. This way, it is taken into account whether the modules are installed in portrait or landscape orientation. The pvlib python [22] implementation of the Marion model is used in this work to model the relative snow loss. To estimate the absolute energy loss, the modeled relative snow loss is multiplied with the modeled energy output of the system, modeled using the procedure described in Section 2.2.

In the development of the snow loss model, Marion et al. found *sc* to be 0.20 [5] for roof mounted systems. This value is the default *sc* in the implementation of the model in pvlib python [22] and the PV modeling software SAM [6]. The snow clearing coefficient is, as previously discussed, expected to depend on different system and module designs [13], because technical aspects can either promote or obstruct natural snow clearing [1]. In our evaluation of the model, we found that snow clearing is slower for the systems we have analyzed [13,16] compared to the validation systems the Marion model is based on. A possible explanation for the difference is higher roof interference for the systems that we have evaluated. In our evaluation of the model we also find that the rate of snow clearing is influenced by the thickness of the snow cover [13]. We therefore add a small modification to the Marion snow loss model by introducing a snow depth dependent *sc*. Because the dataset in this work is the same as in [16], we use the snow depth dependent snow clearing coefficients from [16] that gave the best modeling results. As also described in [16], we use snow depth data from the ground to separate between thin and thick snow covers for the commercial system where there is little sliding. For the residential systems where there is more sliding and where snow depth data from the ground are less representative, we use cumulative snow fall data as an indicator for snow cover thickness.

2.4. Simulation of snow losses for longer time series

To simulate losses for the analyzed systems over time, to get improved understanding of typical losses, we use long time series of weather and snow data to model snow losses, as proposed by [5]. GHI and ambient temperature data for all the locations from the time period 2005-2016 and the ERA5 database is collected from PVGIS [30]. The expected module power output for all the systems is modeled as described for the residential systems in section 2.2. System loss of 7 % is added using the PVWatts system loss function with default loss values for mismatch, wiring, LID, connections and name plate rating [28]. Snow losses are then modeled using the same procedure as described in 2.3.

3 RESULTS

3.1 Snow loss estimated from historical PV data

Figure 3 shows the annual historical snow loss for the analyzed systems (both system configurations) estimated from historical PV data. The loss is given relative to the mean expected annual yield. The mean value is chosen to avoid variations in the loss caused by variation in the total annual irradiation. We observe large variations in snow losses from year to year, and between different systems.

As expected, we observe that weather, system design and climate on snow losses seem to impact the snow losses. The inter-annual variation in snow losses for the systems, as well as the variation in losses between systems located in the same climate zone, but in different locations (C1 and C3), can be explained by typical variations in weather between different locations and different years. C3, R1 and R2 are located in the same area, but R2 has lower loss than R1 every year, and C3 typically has higher loss than both. This could be explained by the impact of tilt on the snow clearing, as snow clearing is inversely proportional with tilt. C1 and C2 have the same technical configuration and experiences the same weather as they are co-located, and their estimated losses are very similar. C4 located in oceanic climate typically has lower losses than the identical systems (C1-C3) located in humid continental climate. R3, located in a subarctic climate, typically has higher losses than R1, which has almost the same tilt but is located in a humid continental climate.

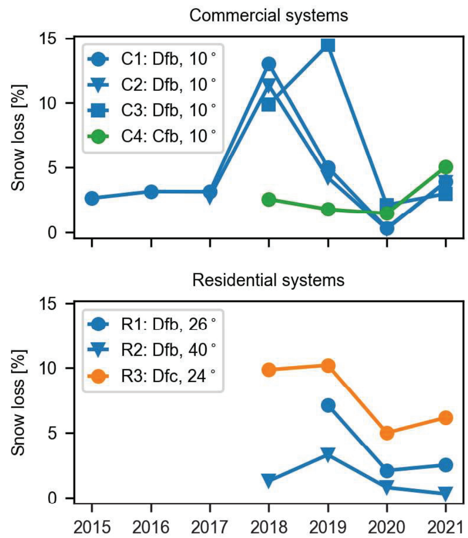


Figure 3: Annual snow loss for the analyzed systems, estimated from PV data. The losses are given relative to mean expected annual yield for the analysis period. The systems in humid continental climate is plotted in blue, green represents oceanic climate and orange represents subarctic climate.

Figure 4 shows the monthly losses for all the full years in the analysis period. Large variations in the monthly loss value are observed for several of the months. For most of the datasets the loss is typically increasing during late autumn, reaching its highest peak in midwinter, before it decreases in the spring. The snow data do, however, not follow the same trend. Typically, the locations have the most snow in the late winter months, but this also corresponds with higher temperatures.

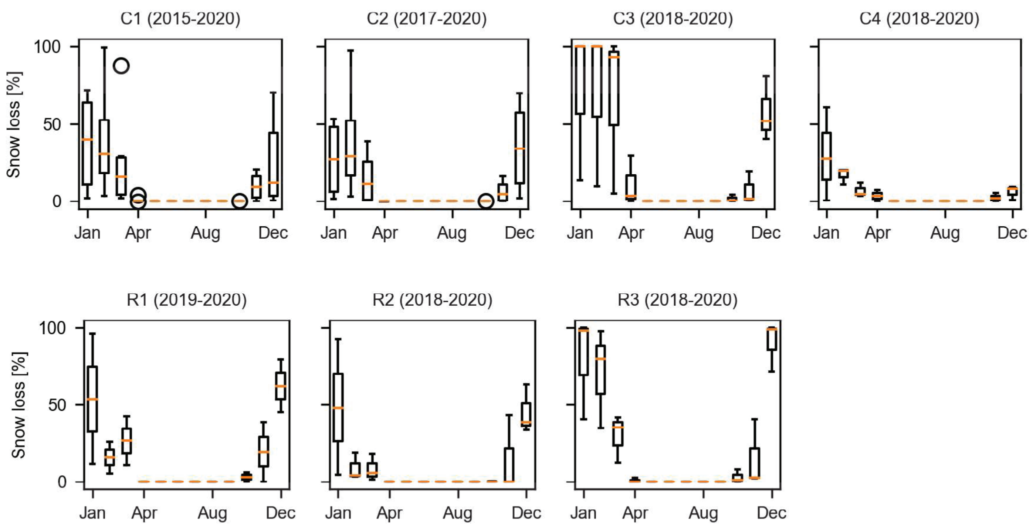


Figure 4: Monthly snow loss for the analyzed systems in the analysis period (given in the subfigure title), estimated from historical PV data. The estimated losses for each month is plotted using a boxplot to show the interannual variation. The box extends from the first to the third quartile values of the monthly loss data, with a line on the median. The whiskers extend to maximum 1.5 multiplied the interquartile range. Outliers are given as circles.

3.2 Simulated snow loss

Based on the results presented in Figure 3 and Figure 4, it is not always clear what would be the best estimate for typical annual and monthly snow losses for the analyzed systems. Especially for the locations with large snow losses, there can be large variations for the same month between different years. With potentially large variations from year to year, estimating typical snow loss for short time series might give an output that is not necessarily representative for the system configurations and the location. Based on this, selecting a representative snow loss value for e.g. a PVSyst simulation seems challenging, as long time series for different system designs and locations would be needed.

Figure 5 and Figure 6 show the correlation between snow loss estimated from PV data and modeled snow loss using the modified Marion snow loss model for respectively annual and monthly losses. Both on the monthly and annual time scales we observe a linear relationship between modeled losses and losses estimated from historical PV data, indicating that the model can be used to predict the losses on both time scales. Some uncertainty in the prediction can, however, be expected. As seen in the figures, there are some deviations between modeled loss and loss estimated from PV data.

In Figure 7 and Figure 8 the simulated monthly and annual snow losses for the analyzed systems using 12 years of irradiation and temperature data from PVGIS is presented. With the longer time series, we get a better understanding of what is typical losses, and what the potential variation and the extreme values could be. With the longer time series, we now see for all of the systems that the losses are highest during mid-winter. Some of the systems get higher monthly median losses than what we observed in Figure 4. This suggests that the years in the analysis period used to estimate losses from historical data are not necessarily years that represent the long-term trend.

Using longer time series and modeling could also enable estimation of snow losses for locations where PV data is lacking. Additionally, future snow losses could be estimated using output data from climate models giving data for the future. To avoid the impact of extreme values, we propose to utilize the median value of the modeled losses as an estimate of the monthly/annual snow losses in yield simulations.

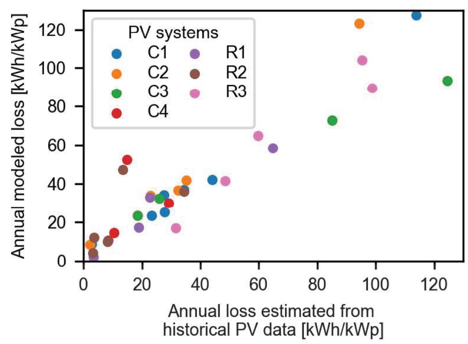


Figure 5: Annual modeled absolute loss compared to loss values estimated from historical PV data.

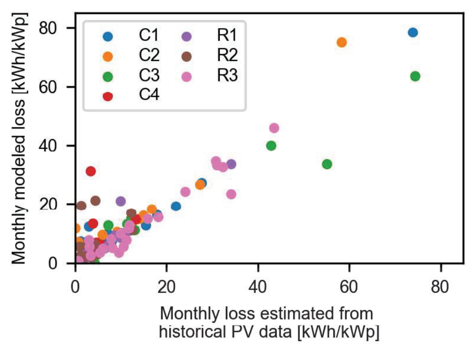


Figure 6: Monthly modeled absolute loss compared to loss values estimated from historical PV data.

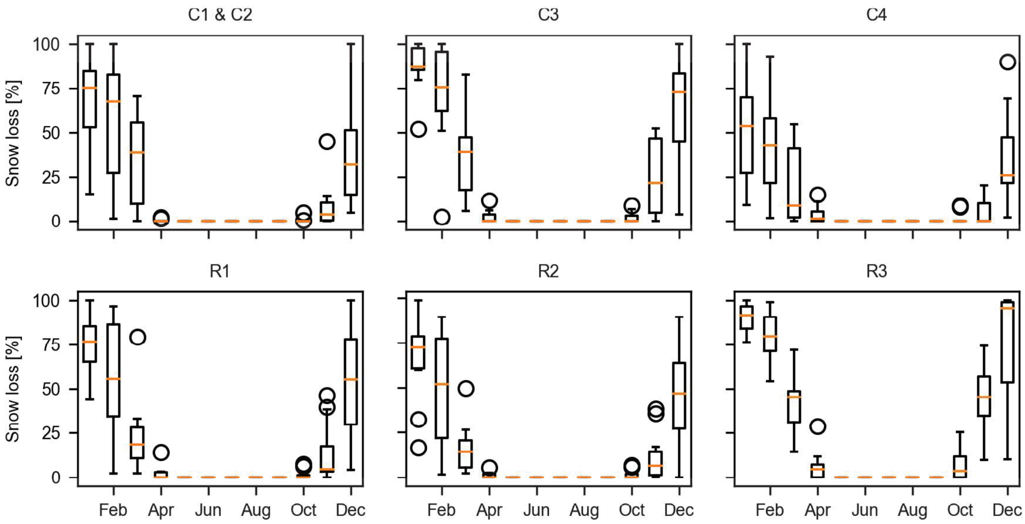


Figure 7: Monthly simulated losses for the analyzed systems, based on 12 years of PVGIS data and the modified Marion snow loss model. The simulated losses for each month is plotted using a boxplot to show the interannual variation. The box extends from the first to the third quartile values of the monthly loss data, with a line on the median. The whiskers extend to maximum 1.5 multiplied the interquartile range. Outliers are given as circles.

Using simulations to estimate PV systems snow loss could in addition to the loss value and estimation on interannual variability, also give realistic production profiles on daily and hourly timescale, which is useful in system size optimization and when building synthetic data series or adding synthetic performance loss for testing of e.g. fault detection algorithms [29]. The uncertainty in the modeling on high time resolutions is likely too high for e.g. monitoring purposes where the modeled PV output should match measured data, but to describe how snow losses vary within a day and from day to day, the modeling is useful.

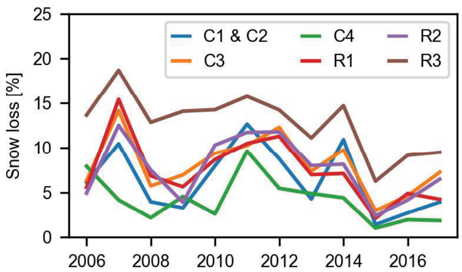


Figure 8: Simulated annual snow loss for the analyzed systems, based on 12 years of PVGIS data and the modified Marion snow loss model. The losses are given relative to mean expected annual yield for the analysis period.

4 CONCLUSIONS

In this work, we estimate annual and monthly snow loss for a set of PV plants in Norway. In both annual and monthly losses, we observe large interannual variations, and we see that systems in colder climates typically have higher losses than systems in warmer climates. We also observe that higher tilt gives reduced losses, confirming previous studies. A modified adaption of the Marion snow loss model where snow depth is considered in the snow clearing modeling is used with 12 years of weather data to simulate losses for a longer time series, to get improved understanding on the potential interannual variation in snow losses. We find that snow loss modeling is a useful tool for estimating monthly or annual snow losses for use in yield modeling when long time series of snow loss data for a given type of system in a given location is not available.

5 REFERENCES

[1] L. Burnham, D. Riley, J. Braid, Design considerations for photovoltaic systems deployed in snowy climates, 37th Eur. Photovolt. Sol. Energy Conf. Exhib. (2020) 1626–1631.
[2] B. Hashemi, A.M. Cretu, S. Taheri, Snow loss prediction for photovoltaic farms using computational intelligence techniques, IEEE J. Photovoltaics. 10 (2020) 1044–1052.
[3] IEA, PVPS Trends in photovoltaic applications 2020, 2020.
[4] R.E. Pawluk, Y. Chen, Y. She, Photovoltaic electricity generation loss due to snow – A literature review on influence factors, estimation, and mitigation, Renew. Sustain. Energy Rev. 107 (2019) 171–182.
[5] B. Marion, R. Schaefer, H. Caine, G. Sanchez, Measured and modeled photovoltaic system energy losses from snow for Colorado and Wisconsin locations, Sol. Energy. 97 (2013) 112–121.
[6] D. Ryberg, J. Freeman, Integration, validation, and application of a PV snow coverage model in SAM, Golden, 2017.
[7] T. Townsend, L. Powers, Photovoltaics and snow: An update from two winters of measurements in the Sierra,

in: 2011 37th IEEE Photovolt. Spec. Conf., IEEE, 2011: pp. 3231–3236.
[8] L.B. Bosman, S.B. Darling, Performance modeling and valuation of snow-covered PV systems: examination of a simplified approach to decrease forecasting error, Environ. Sci. Pollut. Res. 25 (2018) 15484–15491.
[9] E. Lorenz, D. Heinemann, C. Kurz, Local and regional photovoltaic power prediction for large scale grid integration: Assessment of a new algorithm for snow detection, Prog. Photovoltaics Res. Appl. 20 (2011) 760–769.
[10] PVGIS, Data sources and calculation methods, (2020). https://ec.europa.eu/jrc/en/PVGIS/docs/methods (accessed December 22, 2020).
[11] PVsyst, Soiling loss, (2020). https://www.pvsyst.com/help/index.html.
[12] N. Heidari, J. Gwamuri, T. Townsend, J.M. Pearce, Impact of snow and ground interference on photovoltaic electric system performance, IEEE J. Photovoltaics. 5 (2015) 1680–1685.
[13] M.B. Øgaard, B.L. Aarseth, Å.F. Skomedal, H.N. Riise, S. Sartori, J.H. Selj, Identifying snow in photovoltaic monitoring data for improved snow loss modeling and snow detection, Sol. Energy. 223 (2021) 238–247.
[14] M. van Noord, T. Landelius, S. Andersson, Snow-Induced PV Loss Modeling Using Production-Data Inferred PV System Models, Energies. 14 (2021) 1574.
[15] D. Riley, L. Burnham, B. Walker, J.M. Pearce, Differences in snow shedding in photovoltaic systems with framed and frameless modules, in: 2019 46th IEEE Photovolt. Spec. Conf., IEEE, 2019: pp. 558–561.
[16] M.B. Øgaard, H.N. Riise, J.H. Selj, Modeling Snow Losses in Photovoltaic Systems, in: 2021 48th IEEE Photovolt. Spec. Conf., 2021.
[17] T. Saloranta, J. Andersen, Simulations of snow depth in Norway in a projected future climate (2071-2100), 2018.
[18] NVE, seNorge, (2019). www.senorge.no.
[19] Norsk Klimaservicesenter, Observasjoner og værstatistikk, (2020). https://seklima.met.no/observations/.
[20] H.E. Beck, N.E. Zimmermann, T.R. McVicar, N. Vergopolan, A. Berg, E.F. Wood, Present and future köppen-geiger climate classification maps at 1-km resolution, Sci. Data. 5 (2018).
[21] M.B. Øgaard, H.N. Riise, H. Haug, S. Sartori, H. Selj, Photovoltaic system monitoring for high latitude locations, Sol. Energy. 207 (2020) 1045–1054.
[22] W.F. Holmgren, C.W. Hansen, M.A. Mikofski, pvlib python: a python package for modeling solar energy systems, J. Open Source Softw. 3 (2018) 884.
[23] D.G. Erbs, S.A. Klein, J.A. Duffie, Estimation of the diffuse radiation fraction for hourly, daily and monthly-average global radiation, Sol. Energy. 28 (1982).
[24] E.L. Maxwell, A quasi-physical model for converting hourly Global Horizontal to Direct Normal Insolation, Sol. Energy Res. Inst. (1987).
[25] J.E. Hay, J.A. Davies, Calculation of the solar radiation incident on an inclined surface, in: Proc. First Can. Sol. Radiat. Data Work., 1980.
[26] W. De Soto, S.A. Klein, W.A. Beckman, Improvement and validation of a model for photovoltaic array performance, Sol. Energy. 80 (2006) 78–88.
[27] PVsyst, Array Thermal losses, (2021). https://www.pvsyst.com/help/index.html.
[28] A.P. Dobos, PVWatts Version 5 Manual, Golden, CO (United States), 2014.
[29] Å.F. Skomedal, M.B. Øgaard, H. Haug, E.S. Marstein, Robust and fast detection of small power losses in large-scale PV systems, IEEE J. Photovoltaics. 11 (2021) 819–826.
[30] T. Huld, R. Müller, A. Gambardella, A new solar radiation database for estimating PV performance in Europe and Africa, Sol. Energy. 86 (2012) 1803–1815.