Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Home Affairs FDHA
Federal Office of Meteorology and Climatology MeteoSwiss

Technical Report MeteoSwiss No. 251

Predicting photovoltaic power with the COSMO model

Dominik Büeler

ISSN: 2296-0058

Technical Report MeteoSwiss No. 251

Predicting photovoltaic power with the COSMO model

Dominik Büeler

Recommended citation:
Dominik Büeler: 2014, Predicting photovoltaic power with the COSMO model, *Technical Report MeteoSwiss*, 251, 106 pp.

Editor:
Federal Office of Meteorology and Climatology, MeteoSwiss, © 2014

MeteoSwiss
Operation Center 1
CH-8058 Zürich-Flughafen
T +41 58 460 91 11
www.meteoswiss.ch

Abstract

The share of photovoltaic (PV) power in Switzerland is expected to rise during the next decades in response to legal requirements on renewable energy production and energy efficiency. In order to assure constant transition network stability, balance groups are obliged to predict the expected amount of power produced by their PV plants (and automatically fed into the transition network) for every following day in a 15 min-resolution. Due to the strong dependency of PV power production on weather, numerical weather prediction (NWP) models can be used as a basis for day-ahead PV power production forecasting.

In cooperation with the electricity supply company "Elektrizitätswerk der Stadt Zürich" (ewz), which provides power production measurement data in 15 min-resolution over three years for around 270 PV plants spread over Switzerland, different approaches to forecast PV power production based on the NWP model COSMO-2 are tested and verified. Both for ewz and MeteoSwiss, the main objective is to determine and understand the various technical and meteorological aspects influencing a day-ahead PV power production forecast in Switzerland based on COSMO-2.

The case study for a single PV plant in the first part of this study aims to characterize the seasonal sensitivity of the PV power production forecast output to key parameters as the tilt and azimuth angles as well as the temperature and heating coefficients of the PV plant. In a second part, different physical and statistical approaches are applied to compute a power production forecast for all available PV plants in a 1 h-resolution over three years.

Based on the results of this study, the technical uncertainties associated with a PV power production forecast are concluded to dominate over the meteorological ones. On the meteorological side, COSMO-2 irradiance forecast errors seem to be most important. Further contributions might come from the COSMO-2 temperature forecast errors as well as interpolation and topography uncertainties. On the technical side, the deviations of the assumed or fitted from the real metadata parameters are found to be most important: in winter, panel tilt and azimuth angle specifications have the largest impact, whereas the PV module temperature and heating coefficients do not affect the power production forecast at all. In summer, the latter two coefficients become also important due to higher irradiance and temperature levels. Wrong assumptions about these technical coefficients can lead to both increasing and decreasing power production forecast errors, whereas the latter case can occur when wrong assumptions of technical coefficients artificially compensate COSMO global irradiance forecast errors. Snow cover on PV panels is concluded to be the second most important technical uncertainty: also in the flatlands the weather conditions can be such that snow remains on the PV panels for days or weeks, generating "artificial" errors in the PV power production forecast of over 200 %. Furthermore, power production measurement data errors and uncertainties, shading, and uncertainties of the

applied PV module efficiency model contribute to the forecast error.

This study provides an overview of the most important aspects and current shortcomings for building a PV power production forecast model for Switzerland: the largest potential on the meteorological side inheres to an improvement of the irradiance forecast performance of COSMO-2, especially for fog and high fog conditions. Furthermore, proper irradiance forecast data in 15 min-resolution could be made available by the COSMO community. Depending on how the PV sector will develop in the future, which has to be estimated by the energy sector, either fully physical or partly statistical PV power production forecast approaches need to be further developed. In the former case, detailed metadata of each individual PV plant need to be provided by the energy sector, whereas in the latter case this would not be necessary. Furthermore, algorithms detecting errors in a PV power production measurement data set and distincting them from snow-induced patterns have to be developed in order to reduce PV power production forecast errors. Finally, a snow cover parameterization specifically adapted for tilted PV panel systems is an inavoidable feature of a PV power production forecast model for Switzerland.

Contents

1 Introduction

1.1 Motivation

Every of the around 130 Swiss balance groups has to submit a so-called schedule to the Swiss transition network operator (swissgrid) for every following day, the day-ahead. This schedule contains the electricity amount the balance group expects to be consumed and thus drawn from the Swiss transition network by its customers every 15 minutes during the day-ahead, $E_{con}^{schedule}$. The prediction of this value is done with statistical approaches. In order to ensure a balance between electricity consumption and production and thus to enable swissgrid to maintain a constant stability of the transition network, the balance group has to feed the same electricity amount into the transition network as their customers consume. In a second step, therefore, the balance group derives the electricity amount, which has to be fed into the transition network, from the predicted amount of electricity consumption $E_{con}^{schedule}$. It consists of the own electricity production, $E_{prod}^{schedule}$, and the trade of electricity with other suppliers, $E_{trade}^{schedule}$. In summary, the balance group is obliged to plan its electricity production and trade for every 15 minutes of the day-ahead such that it equals the predicted consumption amount given in the schedule:

$$E_{con}^{schedule} = E_{prod}^{schedule} + E_{trade}^{schedule} \tag{1}$$

During the corresponding day, the intra-day, the real electricity amount produced by the balance group every 15 minutes, E_{prod}^{real}, often differs from the planned production in the schedule, $E_{prod}^{schedule}$, by a negative or positive delta due to technically, human-, or weather-induced reasons:

$$E_{prod}^{real} = E_{prod}^{schedule} + \Delta E_{prod} \tag{2}$$

The difference ΔE_{prod} has to be balanced by the balance group, what can be achieved by changing the electricity production of the steerable power plants immediately during the intra-day. The residual differences, which cannot be compensated, have to be balanced by purchasing expensive balancing energy from swissgrid. Therefore, the balance group is forced to minimize the deviation of the real from the planned electricity production, ΔE_{prod}, in order to minimize expensive balancing energy purchases.

New-renewable energy sources (as solar and wind) are strongly weather-dependent. One aspect contributing to the minimization of ΔE_{prod} is thus to use day-ahead forecasts from NWP models in order to predict the expected electricity production from the new-renewable sources. The larger the share of new-renewable sources in the total energy mix of the balance group is, the more important

these predictions are. PV power plants as one of the new-renewable energy sources are gaining importance within the Swiss balance groups. This is why this study investigates the use of the high-resolution NWP model COSMO for the day-ahead prediction of PV power production and thus the minimization of ΔE_{prod}, done exemplarily for the balance group ewz described in the following.

ewz as an active Swiss balance group is the energy service company for the city of Zurich and parts of the Canton of Grisons and belongs to the ten highest-revenue Swiss companies within this field. It employs 1100 people and provides electricity for 220'000 customers. ewz strongly invests in the expansion and promotion of new renewable energies (*Ewz*, 2013).

Figure 1 shows the approximate distribution of the around 270 PV plants managed by ewz in 2012 (the map was downloaded in 2013; however, the qualitative distribution did not change significantly compared to 2012). This small number accounted for less than 1 % of the total ewz electricity production. Therefore, PV power production forecasts were not required at that time because the deviations of the real from the approximately expected PV power production would have been small and could easily be compensated with other production sources. However, during the time of the project, the company was in the phase of defining its future strategy for PV power production, yielding a planned expansion of the PV share of up to 20 % within the following decades. This is why ewz got interested in using COSMO for PV power production forecasts in order to minimize the previously mentioned deviation between planned and real electricity production ΔE_{prod}. [1]

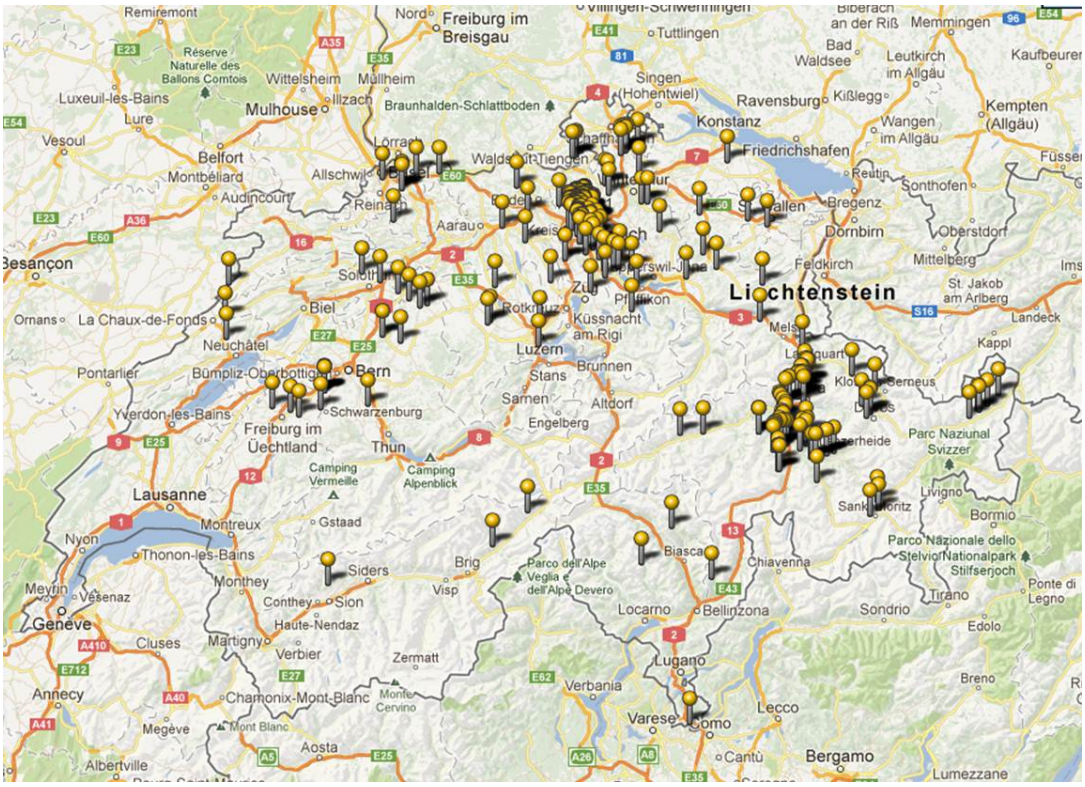


Figure 1: PV plants of ewz in 2013 with a number of around 270 and an installed capacity of around 12 MW in total. The city of Zurich (42 % of all PV plants) and the Canton of Grisons (22 % of all PV plants) are the regions with the highest concentrations. *Ewz* (2013).

An exchange between ewz and and the Business Development division of MeteoSwiss took up the

[1] Personal communication of ewz employees (2013).

aforementioned needs and initialized a project (with a one-year internship in the MeteoSwiss division Coordination Research and Innovation), which is presented in this technical report. The main objective was defined as the determination of the benefit of COSMO forecasts to generate power predictions for PV plants, whereby the term “benefit” refers to a qualitative potential in the sense of associated technical and meteorological aspects influencing the PV power prediction based on COSMO forecasts.

1.2 Theoretical background

1.2.1 PV power production forecast horizons

Depending on the time, for which PV power production is intended to be forecasted into the future, different methods are applied (*Lorenz et al.*, 2010), as summarized in Figure 2. For a PV power production forecast for the next minutes, a combination of irradiance measurement data from surface stations (or satellites) and power production measurements from the PV plants themselves is used as basis for extrapolation approaches. Also PV power production forecasts for the next hours are computed with extrapolation methods. However, satellite-based cloud data serves as background information for this forecast horizon and can be interpolated a few hours into the future by using either satellite-based motion vectors or model-based wind fields. For horizons of more than six hours up to days, NWP models become the main source of information for PV power production forecasts.

This project only deals with the daily forecast horizon, namely the day-ahead forecast as one specific need for balance groups as ewz, by using the NWP model COSMO. Balance groups also require PV power production forecasts for the next minutes and hours (also referred to as “nowcasting”) for the short-term regulation of their electricity production during the intra-day itself. However, this need and its implementation is completely different from day-ahead forecasts and should be treated seperately.

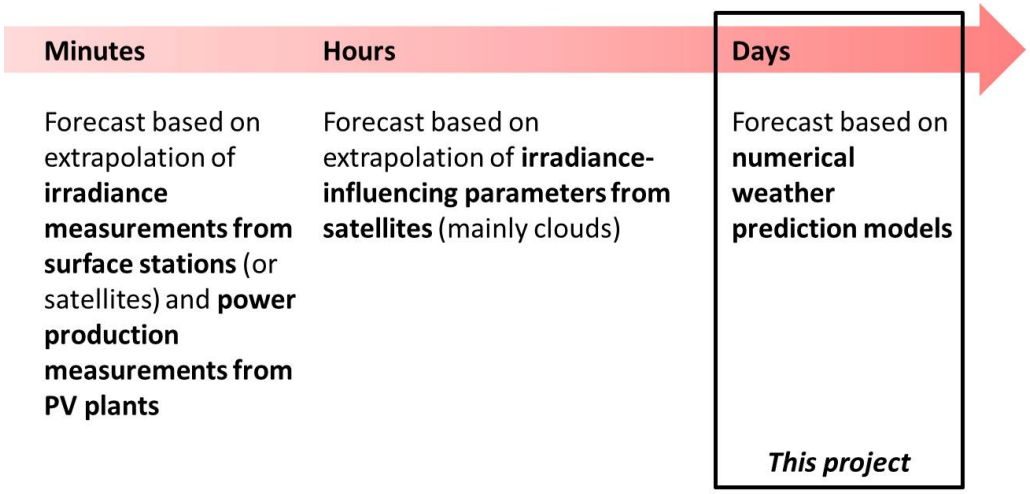


Figure 2: Required horizons for PV power production forecasts and the accordingly applied forecast aproaches. This project deals with the daily horizon only. Sketch according to *Lorenz et al.* (2010).

1.2.2 Factors influencing PV power production

A variety of both technical and meteorological factors have an impact on PV power production. Many of these factors are not accounted for in this study, because they either do not significantly improve PV power production forecasts or the required efforts for their inclusion in this study outweighs any potential benefit. For detailed information about technical and meteorological factors influencing PV power production, the book “Photovoltaics - System Design and Practice” by *Haeberlin* (2012) is recommended.

To give a short overview, the most important rather **technical factors** are listed in the following (details about their implementation in a PV power production forecast model are given in the methods section):

- **PV module efficiency**: determining how much of the incoming solar energy is converted into power is a crucial technical factor. It depends mainly on the incoming solar radiation and the temperature of the PV module. An example for these two dependencies is given in Figure 3, showing a logarithmic increase of the efficiency (of a CIS module) towards high global irradiance values and an efficiency reduction with higher module temperatures (*Beyer et al.*, 2004). The detailed mathematical dependency is given in Equation (22) in the methods section. A further factor influencing the PV module efficiency in a long-term is the age of the module: a commonly used rule of thumb is an efficiency reduction of 10 % after 10 years of operation (*Ewz*, 2013).
- **Installation characteristics of the PV panel**: to calculate the solar energy reaching a PV panel, it is important to know the tilt and azimuth angles as well as the area of the panel. Also important in this context is whether the PV panel is installed free-standing or integrated in roofs or walls, since this has an effect on its thermoregulation and hence module temperature (a free-standing PV panel is ventilated better through wind and thus heats up less) (*Lorenz and Heinemann*, 2012). Finally, the installation location has a strong impact on the power production in the sense that the environment of the PV panel can induce shading (due to topography as shown in Figure 4, objects as other buildings or trees as shown in Figure 5, and the PV panel itself), which leads to a reduction of the power production.
- **Efficiency of further devices**: the PV module efficiency is not the only factor determining the power output. There are additional devices necessary for a PV plant, as the inverter (converting from direct current to alternating current) or the maximum-power-point-(MPP)-tracker (keeping the PV module constantly at its maximum power point, which is the current-voltage condition yielding the maximum possible module power under the current environmental conditions), which both also have an efficiency contributing to the ultimate power production output.

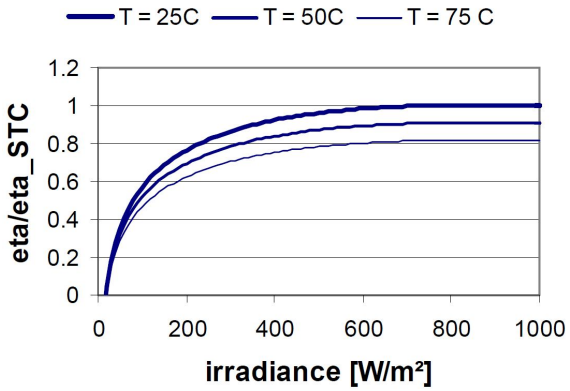


Figure 3: Standardized PV module efficiency (real efficiency divided by efficiency at laboratory standard conditions) for a CIS module as a function of global irradiance. The different curves represent the efficiency at different module temperatures. The figure indicates the logarithmic dependency of the module efficiency on global irradiance and the linear reduction of the same at higher module temperatures. Different PV module types can have slightly different dependency functions and hence shapes of these curves. *Beyer et al.* (2004).

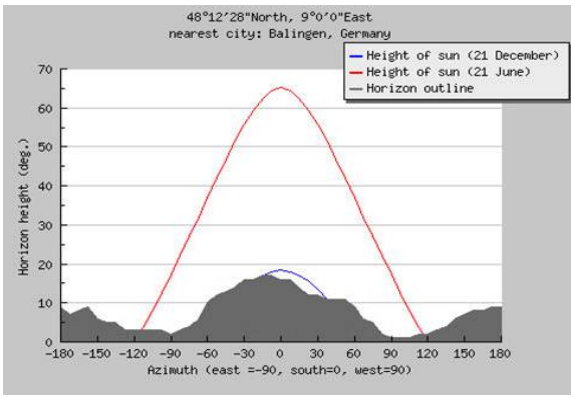


Figure 4: Example of the diurnal course of the sun in winter (blue) and summer (red) at a location in Germany, together with the horizon line at the same location (gray shading). In winter, almost all direct solar radiation is shaded by the topography, whereas in summer, the higher solar elevation allows much more radiation to reach the location. From: http://www.photovoltaik-web.de/ertragsprognose/pvgis/pvgis.html (2014-05-11).

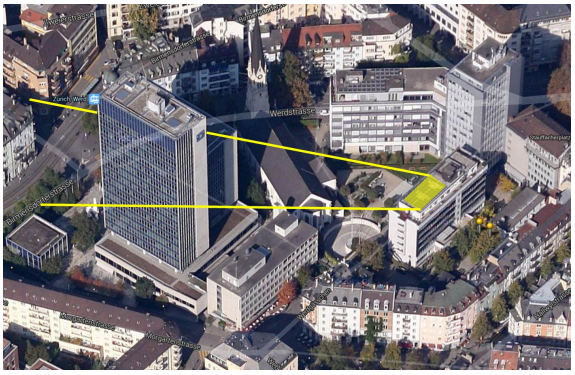


Figure 5: Example of a PV plant (highlighted by the yellow area) in Zurich that can be shaded by a high building close to the plant (indicated by the yellow lines) especially for low solar elevation angles. *Google* (2013).

On the **meteorological side**, the following factors are the most important ones (details about their implementation in a PV power production model are given in the methods section):

- **Solar radiation**: the global irradiance (= incoming global solar radiation) is the "fuel" of a PV module and hence most important. Depending on the weather conditions, the three components direct, diffuse, and ground-reflected irradiance vary in their relative importance. So is direct irradiance most important during clear-sky conditions, diffuse irradiance during cloudy or foggy conditions, and ground-reflected close to grounds with a high albedo as for instance due to snow cover.
- **Air temperature and wind**: mainly these two parameters contribute to the heating or cooling of a PV module and thus its efficiency (*Lorenz and Heinemann*, 2012).
- **Atmospheric turbidity**: it has an influence on the remaining energy of the incoming solar radiation at the surface.
- **Snow**: this parameter influences PV power production in two opposing ways. On one hand, snow can cover a PV panel, what hinders or even stops its power production by reducing or fully cover the PV panel area. Figure 6 indicates, how tilt angle and installation type can determine if and how snow remains lying on a PV panel. On the other hand, snow on the ground can enhance ground-reflection and thus power production.

Figure 6: Example for a partly snow-covered PV panel leading to a reduction of the irradiated area and thus probably of the power production. Zurich, December 2013.

2 Data

2.1 Overview

The data used for this study are the meteorological data provided by MeteoSwiss and the PV power production data as well as metadata provided by ewz. Since the different data sets originally have different temporal and spatial resolutions, most of them need to be converted in order to get to the required resolutions. This means, the meteorological data have to be interpolated or extrapolated to the location of each of the around 270 PV plants and to the required temporal resolution of 1 hour. The final set of data (which was compiled in a database) thus consists of around 270 locations in Switzerland, to which each a set of metadata, of meterological forecast and measurement data, and of power production measurement data is assigned.

All data are available from April 2010 to December 2012 (the required COSMO data are not available for the time before April 2010). The PV power production forecasts are performed retrospectively as hindcasts in order to be verified with the historical power production measurement data of the same period.

Table 1 gives an overview of the different data sets used in this study with additional information about the original temporal and spatial resolutions as well as the source. Details about the individual data sets are given in the subsequent subsections.

Table 1: Overview of the different data sets used in this study. Details are given in the text.

Data	Temporal resolution	Spatial resolution	Source
COSMO-2 forecasts	1 h	2.2 x 2.2 km	MeteoSwiss
Satellite measurements (Meteosat Second Generation)	15 min or 24 h	1.1 x 1.7 km	MeteoSwiss
Surface station measurements (SwissMetNet)	10 min	SwissMetNet stations	MeteoSwiss
PV power production measurements	15 min	PV plants	ewz
PV plant metadata	-	PV plants	ewz

2.2 COSMO-2 forecasts

2.2.1 About COSMO-2

COSMO[2] is a non-hydrostatic limited-area NWP model developed by the **C**onsortium for **S**mall-scale **Mo**deling, which consists of the national meteorological services of Germany, Greece, Italy, Poland, Romania, Russia, and Switzerland. The model is fed with the synoptic boundary conditions of the global model of the European Center for Medium Range Weather Forecast (ECMWF). Currently, MeteoSwiss uses both the regional model COSMO-7 and the local model COSMO-2 for its operational weather forecasts (at the time of the project, COSMO-1 was in development and thus not available yet). With a spatial resolution of 2.2 x 2.2 km (520 x 350 grid points), COSMO-2 is the one with the currently highest resolution, covering a domain of the whole Alpine Arch (and thus more than Switzerland), and thus the one used for this study (if nothing else mentioned, the term "COSMO" refers to "COSMO-2" in this report). In the vertical, the model is divided in 60 levels. The COSMO-2 topography is characterized by a maximum height of 3950 m above sea level and a maximum slope of 15°. The mean model slope within Switzerland amounts to 2.8°. Every 3 h, a COSMO-2 model run is computed, which yields eight runs a day. The computation time step of COSMO-2 is 20 s, whereas for research purposes only hourly outputs are available in general (as it is the case for this study).

2.2.2 Parameters

Table 2: Used COSMO-2 forecast parameters.

Parameter	Description	Unit	Temporal resolution	Temporal type
ASWDIFD_S	Diffuse downward shortwave radiation at surface level on terrain following plane	W m^{-2}	1 h	Average of the previous hour
ASWDIR_S	Direct downward shortwave radiation at surface level on terrain following plane	W m^{-2}	1 h	Average of the previous hour
GLOB	Global downward shortwave radiation at surface level on terrain following plane	W m^{-2}	1 h	Average of the previous hour
T_2M	2m air temperature	°C	1 h	Instantaneous

Table 2 contains the COSMO-2 parameters used for this study. As global irradiance basis, only the sum of the two components (ASWDIFD_S and ASWDIR_S) is used (instead of GLOB directly). The reason is that GLOB is the old COSMO approximation for global irradiance, defined as

$$GLOB = \frac{ASOB\_S}{1 - ALB\_RAD} \tag{3}$$

[2]All technical information about COSMO are either taken from the MeteoSwiss website or are based on internal oral information.

with ASOB_S as the net shortwave radiation at the surface level on a terrain following plane and ALB_RAD $\in [0, 1)$ as the albedo of the surface. This approximation has to be interpreted with caution since ASOB_S is an accumulated and ALB_RAD an instantaneous value. Therefore, any results based on the old COSMO parameter GLOB are not shown in this report.

For the irradiance parameters ASWDIFD_S and ASWDIR_S, the assumption of them being on a horizontal plane has to be made in order to convert them on a tilted and oriented plane (PV panel). This assumption is not fully true, since they are actually computed on a terrain following plane (see Table 2). With this, a bias is introduced. Since the maximum slope of the COSMO-2 topography is only 15°, however, the bias is assumed to be relatively small. Only for (the few) PV plants, which are located in steep mountainous regions, the bias probably has an effect.

2.2.3 Selection of the model run

Even though the PV power production forecasts in this study are computed retrospectively, the conditions for an operational computation have to be considered and assumed, respectively. This means, the COSMO-2 model run of the day before the intra-day has to be selected such that it fulfills the requirements of the Swiss power market. Assuming an operational PV power production forecast for a balance group as ewz, the criteria for the COSMO-2 model run would be in particular:

1. **Time of availability**: the COSMO-2 forecast data have to be available by around 7:00 UTC of the day before the intra-day at the latest, since ewz has to submit the schedule (and thus the PV power production forecast based on the COSMO-2 forecast data) at around 8:00 UTC at this day.
2. **Length of forecast horizon**: the forecast horizon of the COSMO-2 forecast data has to last to the time of the sunset of the intra-day at least in order to cover the whole sunshine duration and thus PV power production potential of the intra-day.
3. **Age of model run**: the COSMO-2 forecast data have to be as new as possible when entering the PV power production forecast model of ewz in order to minimize the forecast uncertainty input by COSMO-2, which is higher for a certain forecasted time when using an older instead of a newer model run.

Figure 7 shows four COSMO-2 model runs (of an assumed Monday), which could potentially be used as a basis for a PV power production forecast for the intra-day (Tuesday). The only COSMO-2 model run meeting all aforementioned requirements is the one available at 03:00 UTC of the day before the intra-day. It is the model run used for aviation applications and thus has a forecast horizon of 45 h (whereas all the other model runs last only 33 h into the future). This horizon makes forecast values available up to 00:00 UTC of the day after the intra-day. The model run available at 06:00 UTC would meet the first requirement (available early enough) and be better considering the third requirement (as new as possible). Due to its shorter forecast horizon, however, it does not fulfill the second criterion (forecast horizon long enough). The model run available at 09:00 UTC would be better considering the third requirement (as new as possible). However, it would not meet the first requirement (availability early enough) due to its late time of availability and also the second requirement (forecast horizon long enough). The model run available at 12:00 UTC would meet the second (forecast horizon long enough) and even be best with regard to the third requirement (as new as possible). However, this model run

would be too late to fulfill the first criterion (available early enough). **Therefore, all meteorological forecast data in this study are based on the COSMO-2 03:00-UTC model run of the day before the intra-day.**

COSMO-2 model run					Lead times							
Mon, 03:00	Mon, 06:00	Mon, 09:00	Mon, 12:00		Tue, 03:00	Tue, 06:00	Tue, 09:00	Tue, 12:00	Tue, 15:00	Tue, 18:00	Tue, 21:00	Wed, 00:00
03:00-run					+ 24 h	+ 27 h	+ 30 h	+ 33 h	+ 36 h	+ 39 h	+ 42 h	**+ 45 h**
	06:00-run				+ 21 h	+ 24 h	+ 27 h	+ 30 h	**+ 33 h**			
		09:00-run			+ 18 h	+ 21 h	+ 24 h	+ 27 h	+ 30 h	**+ 33 h**		
			12:00-run		+ 15 h	+ 18 h	+ 21 h	+ 24 h	+ 27 h	+ 30 h	**+ 33 h**	

Figure 7: Four COSMO-2 model runs (of an assumed Monday), which could potentially be used as a basis for a PV power production forecast for the intra-day (Tuesday). The model runs 03:00-UTC, 06:00-UTC, 09:00-UTC, and 12:00-UTC are shown, including their forecast horizon in hours into the future (= lead time). In red, the time of the intra-day with potential sunshine (and thus required for the PV power production forecast) is indicated, whereas the gray color represents the time after sunset (not required anymore). The blue line indicates the deadline of the latest required availability of the weather forecast data, as defined by the power market. The sketch is based on the information from a project meeting with ewz and technical information about COSMO-2.

2.2.4 Spatial and temporal interpolation

In order to get COSMO-2 forecast data for each of the around 270 PV plant locations, the COSMO-2 grid points have to be interpolated to the desired locations. This is done by an often used procedure for research purposes, which does not necessarily find the closest but most representative COSMO-2 grid point for each PV plant location. Hereby, distance as well as altitude difference between the grid points and the desired location are accounted for.

No temporal interpolation has to be done for the COSMO-2 forecast data, since their temporal resolution is 1 h, which is the resolution intended to be used for the PV power production forecast.

2.3 Satellite measurements

2.3.1 About the satellite

The satellite measurement data are from the Meteosat Second Generation (MSG) satellite, operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT). The used parameters for this study are from a gridded data set, developed and postprocessed at MeteoSwiss.

The gridded irradiance data sets used in this study have a general uncertainty of around 10 % (*Stoeckli*, 2013). This is important to keep in mind, since the different satellite irradiance parameters are used for various verifications (as the one of the COSMO-2 irradiance forecast data) and thus assumed to represent the “reality”.

2.3.2 Parameters

Table 3 shows the satellite measurement parameters used for this study.

Table 3: Used satellite measurement parameters.

Parameter	Description	Unit	Temporal resolution	Temporal type
KI	Heliosat clear-sky index (5 % = completely overcast)	%/100	15 min	Instantaneous
SAA	Solar azimuth angle (0° = north)	°	15 min	Instantaneous
SCAN_TIME	Time of actual scan of pixel in seconds after time specified in time axis	s	24 h	-
SIS	Surface incoming shortwave radiation (global radiation)	$W\,m^{-2}$	15 min	Instantaneous
SISDIF	Surface incoming shortwave radiation (diffuse component on horizontal plane)	$W\,m^{-2}$	15 min	Instantaneous
SISDIR	Surface incoming shortwave radiation (direct component on horizontal plane)	$W\,m^{-2}$	15 min	Instantaneous
SISDIRCF	Surface incoming shortwave radiation (clear-sky - direct component on horizontal plane)	$W\,m^{-2}$	15 min	Instantaneous
SNOWMASK	Snow mask (0 = snow free, 1 = snow covered)	-	24 h	-
SZA	Solar zenith angle	°	15 min	Instantaneous

2.3.3 Spatial and temporal interpolation

The spatial interpolation of the satellite parameters is done by assigning the closest grid point of the gridded satellite data set to each PV plant location.

To provide the satellite parameters with a resolution of 1 h, the 15 min values have to be aggregated. Since the time specified in the time axis of a satellite parameter differs from the real measurement time of the satellite by the parameter SCAN_TIME (explained in Table 3), the latter has to be considered for an accurate temporal aggregation. As sketched in Figure 8, the aggregation procedure looks as follows: the (constant) scan time of a satellite grid point amounts to around 12 min on average. This means, a satellite parameter value specified with 00:00 was actually measured at 00:12. Therefore, to get the value at 01:00, which should be the mean value of the hour between 00:00 and 01:00 (in order to make it comparable to the according irradiance value of COSMO at 01:00, which is also an hourly mean value), the mean over the values at 00:00, 00:15, 00:30, and 00:45 is used, which were actually measured at around 00:12, 00:27, 00:42, and 00:57. A use of the mean of the values at 00:15, 00:30, 00:45, and 01:00 would include the value 01:00, which was actually measured at 01:12 and would thus be outside the investigated hour. Hence, the former procedure seems to be more appropriate than the latter for the purpose of this study.

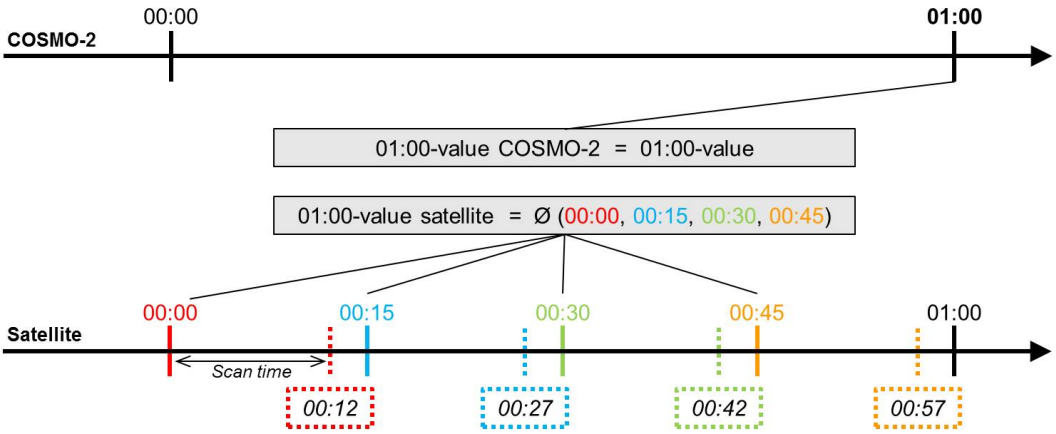


Figure 8: Applied procedure for the temporal aggregation of the satellite data. Details are in the text.

2.4 Surface station measurements

2.4.1 Parameter

Table 4 shows the surface station measurement parameter used for this study.

Table 4: Used surface measurement parameter.

Parameter	Description	Unit	Temporal resolution	Temporal type
tre200s0	2m air temperature	°C	10 min	Instantaneous

2.4.2 Spatial and temporal interpolation

To assign the surface station measurement data to the PV plant locations, the closest SwissMetNet (SMN) surface station, having a measurement time series for the desired period, is assigned to each PV plant. Even though an additional altitude correction between a surface station and the PV plant location would make sense (since a temperature parameter is used), it is not done. The reason is the fact, that the air temperature seems to be of secondary importance for a PV power production forecast (after irradiance) and the effort for the implementation of an algorithm accounting for the altitude correction would be disproportionately large considering the probably small benefit it would yield for the purpose of this study. The use of the gridded temperature data set by (*Frei*, 2014) would be an alternative to minimize the spatial interpolation uncertainties. However, its temporal resolution of 24 h would require temporal interpolation yielding uncertainties that would probably outweigh the aforementioned ones of the spatial interpolation. A gridded temperature data set with a resolution of 10 min or at least 1 h would thus be optimal for further projects within the energy sector.

The surface station measurement parameter does not need to be aggregated to a temporal resolution

of 1 h but to 15 min (the reason is explained in the method section). The original temporal resolution of 10 min yields instantaneous measurement values at 00:00, 00:10, 00:20, 00:30, 00:40, 00:50, and 01:00. Therefore, to get the required values at 00:15 and 00:45, the mean of the values at 00:10 and 00:20 and 00:40 and 00:50, respectively, is calculated.

2.5 PV power production measurements

2.5.1 About the PV plants

The number of ewz PV plants with power production measurement data amounted to 271 at the end of 2012. Hence, this is the maximum number the PV power production forecast is done for. Due to the constant increase of the number of PV plants managed by ewz, not all 271 PV plants existed yet in 2010 and 2011. The number of operative plants providing power production measurement data developed over the three years is shown in Table 5. Therefore, the PV power production forecasts are computed for these sets only and not for all 271 PV plants in every of the three years.

Table 5: Development of the number of operational PV plants from 2010 to 2012.

Year	Number of operational PV plants
2010	244
2011	259
2012	271

2.5.2 Parameter

The power production measurement data are available in a 15 min resolution and thus represent the aggregated power production in kW h over the past 15 minutes.

2.6 PV plant metadata

2.6.1 Parameters

For each individual PV plant, the metadata parameters shown in Table 6 are available. However, not all of them revealed to be necessary for the power production forecast approaches used in this study.

Table 6: Metadata parameters for each PV plant provided by ewz and thus available for this study.

Parameter	Description
Name	Name of the PV plant defined by ewz (example: Alexandra)
Location	Name of the village or town the PV plant is located in
Coordinates	Latitude and longitude of the PV plant
Nominal power	Total installed nominal power of the PV plant in kWp (= kW at standard (peak) conditions)
Operation start date	Month and year, at which the PV plant was put into operation
Building type	Type of the building the PV plant is installed on (example: agricultural building)

3 Methods

The methods section is split into three parts: the first one describes the general approach for PV power production forecasting, with a focus on the conversion of global irradiance on a tilted and azimuthally oriented surface and the modeling of the PV module efficiency. In a second part, the different specific PV power production forecast approaches applied and investigated in this study are introduced. The third part describes the verification methods used to analyze and evaluate the different PV power production forecast approaches. Table 7 gives an overview of all the parameters used for both the PV power production forecast and the verification calculations and thus introduced in this section.

Table 7: Parameters used in this report.

Symbol	Variable	Unit
A	Area of the PV panel	m^2
ψ	Azimuth angle of the PV panel (measured clockwise from north)	°
$TC_{loc \rightarrow lst}$	Conversion factor from local time to local solar time	min
$TC_{std \rightarrow loc}$	Conversion factor from standard time to local time	min
d	Day of the year	-
$I_{dif,h}$	Diffuse irradiance on a horizontal surface	$W\,m^{-2}$
$I_{dif,t}$	Diffuse irradiance on a tilted surface	$W\,m^{-2}$
$I_{dir,b}$	Direct beam irradiance	$W\,m^{-2}$
$I_{dir,h}$	Direct horizontal irradiance	$W\,m^{-2}$
$I_{dir,t}$	Direct irradiance on a tilted surface	$W\,m^{-2}$
a_{1-3}	Factors for PV module efficiency irradiance dependency function	-
k_{1-6}	Factors for PV module efficiency irradiance and air temperature dependency function	-
$I_{glob,h}$	Global irradiance on a horizontal surface	$W\,m^{-2}$

Continued on next page

Table 7 – *Continued from previous page*

Symbol	Variable	Unit
$I_{glob,t}$	Global irradiance on a tilted surface	W m^{-2}
I_{stc}	Global irradiance on a tilted surface at standard conditions	W m^{-2}
ϕ	Latitude	°
LST	Local solar time	h
λ	Longitude	°
λ_{std}	Longitude of the standard meridian of the time zone containing longitude λ	°
T_{noct}	Nominal operating cell temperature (NOCT) / PV module temperature at representative outdoor operation conditions ($I = 800\,\mathrm{W\,m^{-2}}$, $T_{air} = 20\,^{\circ}\mathrm{C}$, wind velocity of 1 m s^{-1}, PV panel being installed free-standing (meaning a good circulation also on the back side))	°C
P_{nom}	Power of the PV plant at standard conditions ($I_{glob,t} = 1000\,\mathrm{W\,m^{-2}}$, $T_{mod} = 25\,^{\circ}\mathrm{C}$, $AM = 1.5$[3]) / nominal power of the PV plant	kW
$\eta = \eta(I_{glob,t}, T_{air})$	PV module efficiency at $I_{glob,t}$ and T_{air}	-
η_{stc}	PV module efficiency at standard conditions ($I_{glob,t} = 1000\,\mathrm{W\,m^{-2}}$, $T_{mod} = 25\,^{\circ}\mathrm{C}$, $AM = 1.5$)	-
γ	PV module heating coefficient (PV module temperature change per W m^{-2} of irradiance change	°C m^2 W^{-1})
T_{mod}	PV module temperature	°C
$T_{mod,stc}$	PV module temperature at standard conditions	°C
E	PV power production (forecasted or measured)	kWh
ρ_X	Sensitivity of the PV power production forecast E to a change in parameter X	kWh per unit of parameter X
θ_i	Solar angle of incidence on the PV panel (angle between the vector normal to the PV panel and the solar irradiance)	°
θ_a	Solar azimuth angle (measured clockwise from north)	°
D	Solar day angle	rad

Continued on next page

[3] AM stands for air mass and defines the direct path length of the solar radiation relative to the direct path length (down to sea level) during a solar zenith angle of 0°. The value of $AM = 1.5$ is a useful standard condition since it is the air mass during a solar zenith angle of 48.2°, representing the approximate year-round average for the mid-latitudes (*Wuerfel*, 2009).

Table 7 – *Continued from previous page*

Symbol	Variable	Unit
δ	Solar declination angle	°
θ_e	Solar elevation angle (angle between the horizontal plane and the solar irradiance beam)	°
ω	Solar hour angle	°
$\theta_z = 90° - \theta_e$	Solar zenit angle (angle between the vector normal to the earth surface and the solar irradiance beam)	°
ST	Standard time	-
α	Temperature coefficient (PV module efficiency change per $°\mathrm{C}$ of module temperature change)	$°C^{-1}$
β	Tilt angle of the PV panel (measured from the horizontal plane)	°
Δt	Time step for the PV power production forecast	h
T_{air}	2m air temperature	°C

3.1 PV power production forecasting in general

3.1.1 General formula for PV power production forecasting

Equation (4) shows the general procedure for the computation of a PV power production forecast:

$$E(t) = \int_{t-\Delta t}^{t} (I_{glob,t}(t) \cdot \eta(t))\,\mathrm{d}t \cdot A \tag{4}$$

$E(t)$ is defined as the predicted aggregated power production of a certain PV plant between the times $t - \Delta t$ and t in kWh. Therefore, it is calculated by the multiplication of the integral of the predicted global irradiance on the (tilted and azimuthally oriented) PV panel $I_{glob,t}(t)$ in $\mathrm{W\,m^{-2}}$ and the predicted efficiency of the PV module $\eta(t)$ (dimensionless) between the times $t - \Delta t$ and t and the constant area of the PV panel A.

Assuming $I_{glob,t}(t)$ to represent the mean global irradiance over the time span between $t - \Delta t$ and t, which is indeed the case for a COSMO irradiance parameter specified with time t, and $\eta(t)$ to represent the according mean PV module efficiency over the same time span, Equation (4) can be written in a discretized form:

$$E(t) = I_{glob,t}(t) \cdot \eta(t) \cdot A \cdot \Delta t \tag{5}$$

The following subsections provide the general procedures for the calculation of the four terms in Equation (5), with a focus on the conversion of the global irradiance on a horizontal to the one on a tilted and azimuthally oriented plane as well as on two PV module efficiency modeling approaches from the literature.

3.1.2 Irradiance on a tilted and azimuthally oriented PV panel

The first key step for accurate PV power production forecasting is to calculate the global irradiance on a tilted and azimuthally oriented PV panel, for which the irradiance components have to be available. As the direct component both the direct irradiance on a horizontal plane and the direct beam irradiance can be used as basis for this conversion. For the diffuse component the diffuse irradiance on a horizontal plane is required. The following conversion calculations are based on *PVeducation* (2014), *Heinemann* (2002), *ITACA* (2014), *Spencer* (1971), and *Lorenz and Heinemann* (2012).

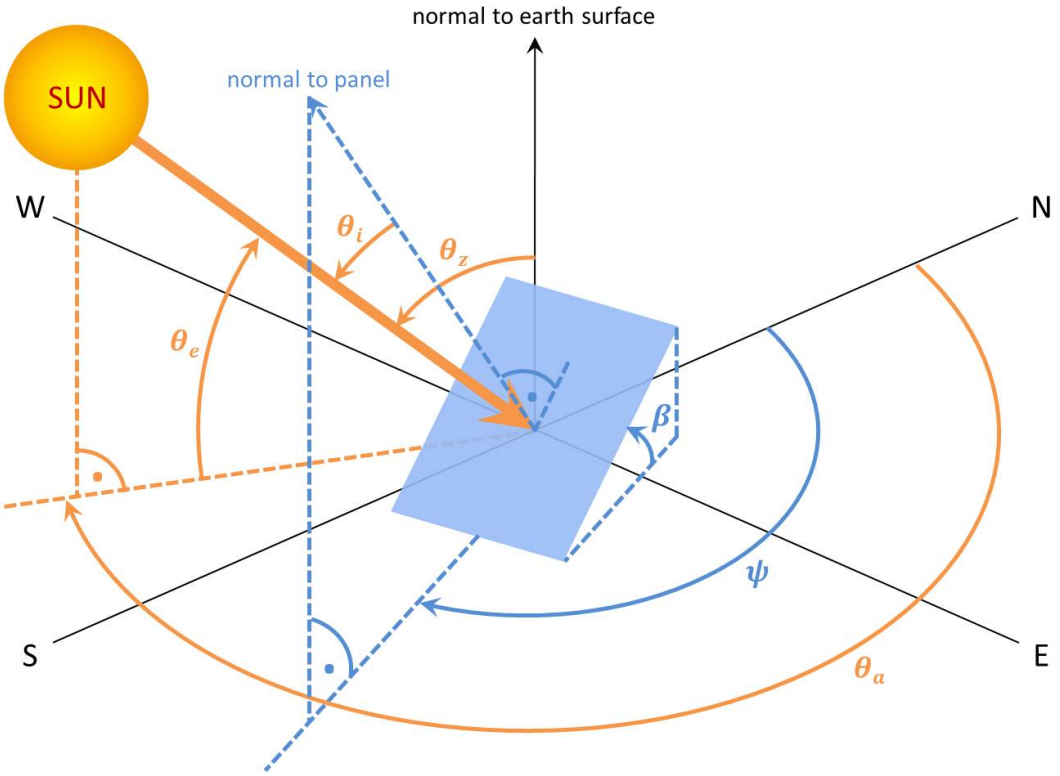


Figure 9: Important angles for the calculation of the angle of incidence θ_i of the sun on a tilted and azimuthally oriented PV panel, defined as the angle between the direct solar irradiance beam and the normal of the PV panel. β and ψ are the tilt and azimuth angles, respectively, of the PV panel. The latter is measured clockwise from north (yielding $\psi = 180°$ for a southward oriented PV panel). θ_a is the solar azimuth angle, also measured clockwise from north. θ_z and θ_e are the solar zenith and elevation angles, respectively, whereby $\theta_z + \theta_e = 90°$.

The conversion of the **direct irradiance** is done as follows:

$$I_{dir,t} = I_{dir,h} \cdot \frac{cos(\theta_i)}{cos(\theta_z)} \tag{6}$$

where $I_{dir,h}$ (in W m^{-2}) is the direct irradiance on a horizontal plane, θ_z (in °) the solar zenith angle required to convert the direct horizontal to the direct beam irradiance, and θ_i (in °) the angle of incidence on the PV panel, which is shown in Figure 9 and whose calculation is explained in the subsequent paragraphs. If the direct beam irradiance is directly available as a basis for the conversion, Equation (6) gets simplified:

$$I_{dir,t} = I_{dir,b} \cdot cos(\theta_i) \tag{7}$$

The conversion of the **diffuse irradiance** is done as follows:

$$I_{dif,t} = \frac{I_{dif,h} \cdot (1 + cos(\beta))}{2} \tag{8}$$

Hereby, $I_{dif,h}$ (in W m^{-2}) is the diffuse irradiance on a horizontal plane and β (in °) the tilt angle of the PV panel (shown in Figure 9).

To obtain the **global irradiance** on a tilted and azimuthally oriented PV panel, $I_{glob,t}$ (in W m^{-2}), the two components from Equations (6) and (7), respectively, and (8) need to be summed up (the third term, the ground-reflected irradiance, is neglected in this study and thus not shown here):

$$I_{glob,t} = I_{dir,t} + I_{dif,t} \tag{9}$$

The conversion of the direct irradiance on a tilted and azimuthally oriented PV panel requires the calculation of the **angle of incidence** θ_i at the investigated time. It mainly depends on the local solar time, the solar elevation angle, and the solar azimuth angle, as described in the following paragraphs.

In a first step, the local solar time is derived from the local time by calculating two correction factors. The first factor $TC_{std \rightarrow loc}$ (in min) accounts for the conversion from standard to local time:

$$TC_{std \rightarrow loc} = 4\,\text{min/deg} \cdot (\lambda - \lambda_{std}) \tag{10}$$

λ is the longitude of the investigated location and λ_{std} the standard meridian of the used time zone (both in °). Since the data of this study are all in UTC, λ_{std} equals 0°.

The second factor is the (empirical) equation of time (also in min) accounting for the conversion from local time to local solar time (*Spencer*, 1971):

$$\begin{aligned} TC_{loc \rightarrow lst} = &(0.000075 + 0.001868 \cdot cos(D) - 0.032077 \cdot sin(D) \\ &- 0.014615 \cdot cos(2 \cdot D) - 0.040849 \cdot sin(2 \cdot D)) \\ &\cdot 4 \cdot \frac{180}{\pi} \end{aligned} \tag{11}$$

with the day angle D (in rad) being a function of the day of the year d (dimensionless) as

$$D = 2 \cdot \pi \cdot \frac{d-1}{365} \tag{12}$$

The local solar time LST (in h) can then be calculated by adding the two time correction factors $TC_{std \to loc}$ and $TC_{loc \to lst}$ to the standard time ST (which is expressed in h since the start of the day):

$$LST = ST + (TC_{std \to loc} + TC_{loc \to std}) \cdot \frac{1}{60} \text{h min}^{-1} \tag{13}$$

Alternatively, the local solar time LST can be expressed as solar hour angle ω (in °). It is negative in the morning, zero at solar noon, and positive in the afternoon, with a change of 15 ° h^{-1}. The conversion looks as follows:

$$\omega = 15° \text{ h}^{-1} \cdot (LST - 12 \text{ h}) \tag{14}$$

The annual solar cycle is represented by the solar declination angle δ (in °), which can be calculated with the following empirical equation (*Spencer*, 1971):

$$\begin{aligned} \delta = (&0.006918 - 0.399912 \cdot cos(D) + 0.070257 \cdot sin(D) \\ &- 0.006758 \cdot cos(2 \cdot D) + 0.000907 \cdot sin(2 \cdot D)) \cdot \frac{180}{\pi} \end{aligned} \tag{15}$$

The previously derived parameters can now be used to calculate the first important angle, the solar elevation angle θ_e (in °; shown in Figure 9) at a certain location with latitude ϕ (in °):

$$\theta_e = sin^{-1}\left(sin(\delta) \cdot sin(\phi) + cos(\delta) \cdot cos(\phi) \cdot cos(\omega)\right) \tag{16}$$

As a second important angle, the solar azimuth angle θ_a (in °; shown in Figure 9) can also be derived from the previously derived parameters and the location's longitude λ (in °) in addition:

$$\theta_a = cos^{-1}\left(\frac{sin(\delta) \cdot cos(\lambda) + cos(\delta) \cdot sin(\lambda) \cdot cos(\omega)}{cos(\theta_e)}\right) \tag{17}$$

Together with the PV panel's tilt and azimuth angles β and ψ, the solar elevation angle θ_e and the solar azimuth angle θ_a (all shown in Figure 9) can finally be combined to get the cosine of the angle of incidence θ_i (in °):

$$cos(\theta_i) = sin(\theta_e) \cdot cos(\beta) + cos(\theta_e) \cdot sin(\beta) \cdot cos(\psi - \theta_a) \tag{18}$$

$cos(\theta_i)$ (dimensionless) can now be used to convert the direct irradiance on a tilted and azimuthally oriented PV panel according to the Equations (6) and (7), respectively, as described in the beginning

of this section.

3.1.3 PV module efficiency

The PV module efficiency η (dimensionless) expresses the proportion of the energy from the global irradiance incident on the PV module that is converted into electricity. It depends mainly on the incoming global irradiance $I_{glob,t}$ itself and the operating temperature of the PV module T_{mod} (in °C). The high complexity of these two dependencies as well as the diversity of PV module technologies and installation types has spawned a significant number of literature studies dealing with different approaches to model these dependencies for PV power forecasting purposes.

Model by Beyer et al. (2004)

For one part of this study, the broadly used model of *Beyer et al.* (2004) (and *Evans and Florschuetz* (1977)) is applied. The model is based on the assumption of the PV module being constantly at its maximum power point. Therefore, the PV module efficiency at the maximum power point, η_{MPP}, is just referred to as η hereafter.

In a first part, the model describes the irradiance dependency of the PV module at the standard condition module temperature of 25 °C as

$$\eta(I_{glob,t}, T_{mod,stc} = 25\,^{\circ}\mathrm{C}) = a_1 + a_2 I_{glob,t} + a_3 \ln(I_{glob,t}) \tag{19}$$

whereas a_{1-3} (dimensionless) are module-specific parameters (*Beyer et al.*, 2004).

The efficiency at irradiance $I_{glob,t}$ and module temperature $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$ changes with changing module temperature according to the following relationship (*Beyer et al.*, 2004):

$$\eta(I_{glob,t}, T_{mod}) = \eta(I_{glob,t}, T_{mod,stc} = 25\,^{\circ}\mathrm{C}) \cdot (1 + \alpha(T_{mod} - 25\,^{\circ}\mathrm{C})) \tag{20}$$

α (in °C^{-1}) is the temperature coefficient representing the efficiency change per °C change of module temperature T_{mod}. It is negative because the efficiency decreases with module temperatures higher than 25 °C and increases when T_{mod} sinks below 25 °C.

The module temperature is a result of the energy from incoming irradiance not converted into electricity but lost as thermal energy. It depends on the ambient air temperature T_{air} (in °C) and the global irradiance $I_{glob,t}$, scaled with a positive heating coefficient γ (in °C m^2 W^{-1}):

$$T_{mod} = T_{air} + \gamma I_{glob,t} \tag{21}$$

Hence, higher air temperature and global irradiance values lead to a stronger heating of the PV module and thus a decreasing efficiency.

The different dependencies can be united in a single equation expressing the efficiency as a function of irradiance and air temperature:

$$\eta(I_{glob,t}, T_{air}) = (a_1 + a_2 I_{glob,t} + a_3 \ln(I_{glob,t})) \cdot (1 + \alpha(T_{air} + \gamma I_{glob,t} - 25\,^{\circ}\mathrm{C})) \tag{22}$$

From the perspective of a power production forecasting based on NWP model data, the described efficiency model depends on five unknown technical parameters, a_{1-3}, α, and γ, and two known meteorological parameters T_{air} and $I_{glob,t}$.

The parameters a_{1-3} can be derived from a parameter fitting based on at least three known efficiency values at 25 °C but different irradiance values $I_{glob,t}$. Usually, the technical data sheet of a PV module (published by the manufacturer) specifies two efficiency values: one of them is the efficiency at (laboratory) standard conditions η_{stc} (that is $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$, $I_{glob,t} = 1000\,\mathrm{W\,m^{-2}}$, and $AM = 1.5$; hence: $\eta_{stc} = \eta(T_{mod,stc} = 25\,^{\circ}\mathrm{C}, I_{glob,t} = 1000\,\mathrm{W\,m^{-2}})$) and another one at any low irradiance values around $I_{glob,t} = 200\,\mathrm{W\,m^{-2}}$. Further efficiency values can be derived by interpolation between the two known values or by derivation from the measured power output of the PV plant.

The temperature coefficient α is usually also given on the technical data sheet of the PV module. It is often around $-0.0045\,^{\circ}\mathrm{C}^{-1}$ on average (*Dubey et al.*, 2013).

Compared to α, the heating factor γ varies for a certain PV module type due to its strong dependency on the installation type of this PV plant. Compared to a roof- or facade-integrated, a free-standing PV panel heats up much less during operation because the backside of the panel is cooled better due to higher air ventilation caused for instance by wind. Therefore, values of γ range from $0.02\,^{\circ}\mathrm{C\,m^2\,W^{-1}}$ for free-standing to $0.06\,^{\circ}\mathrm{C\,m^2\,W^{-1}}$ for sloped-roof-integrated PV panels (*Nordmann and Clavadetscher*, 2003). The dependency on the installation type is the reason why γ is usually not given in the technical data sheet of the PV module and thus has to be estimated.

Figure 10 contains an exemplary outcome of the PV module efficiency model, shown by the effective efficiency η divided by the efficiency at standard conditions η_{stc}. The used parameter set consists of $\eta_{stc} = 0.117$, $\alpha = -0.004\,18\,^{\circ}\mathrm{C}^{-1}$, and $\gamma = 0.056\,^{\circ}\mathrm{C\,m^2\,W^{-1}}$. a_{1-3} are fitted by assuming eight module data points at $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$ and different irradiances $I_{glob,t}$, indicated by the gray-pointed line (whereby two of them are taken from an exemplary technical module sheet - η at $1000\,\mathrm{W\,m^{-2}}$ ($= \eta_{stc}$) and η at $200\,\mathrm{W\,m^{-2}}$ - and the others approximately interpolated). The resulting black solid line shows the resulting efficiency at $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$. The residual three solid colored curves show the efficiency at module temperatures $T_{mod} = 10\,^{\circ}\mathrm{C}$, $T_{mod} = 50\,^{\circ}\mathrm{C}$, and $T_{mod} = 75\,^{\circ}\mathrm{C}$, computed by using the mentioned α.

Expressing the module temperature as a function of air temperature and irradiance, the outcome looks different, as shown in Figure 11. This is due to the module temperature varying along a solid curve of constant air temperature.

Due to the lack of information about the different module types making up the PV plant system of this study, assumptions about the efficiency at standard conditions η_{stc} and the temperature coefficient α have to be made. An analysis of ten random PV plants, for which ewz was able to find specific module information, revealed a diversity of module types and manufacturers. This makes it reasonable to use the mean of η_{stc} and α values from 19 different literature studies (summarized in Table 19 by *Dubey et al.* (2013) in the Appendix), which is $\eta_{stc} = 0.117$ and $\alpha = -0.004\,18\,^{\circ}\mathrm{C}^{-1}$.

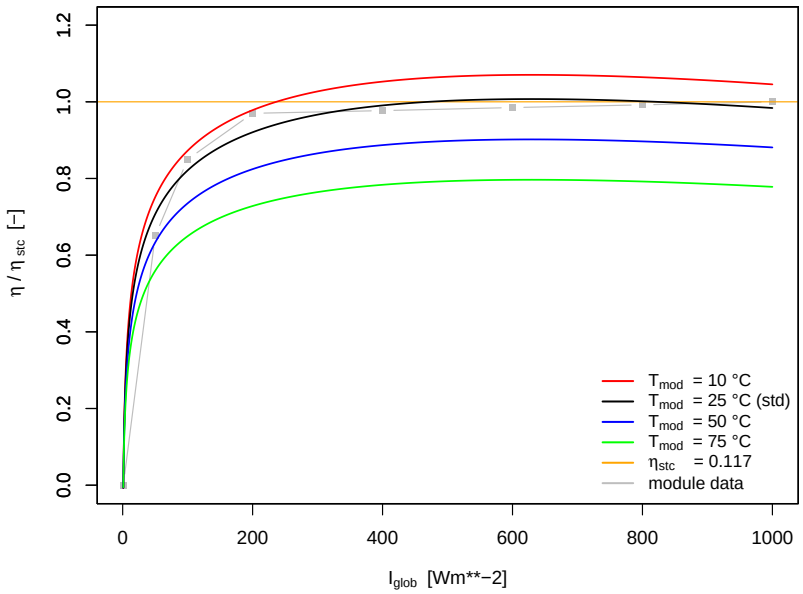


Figure 10: Exemplary PV module efficiency as a function of irradiance and module temperature, shown by the effective efficiency η divided by the efficiency at standard conditions η_{stc}. The used parameter set consists of $\eta_{stc} = 0.117$, $\alpha = -0.004\,18\,^{\circ}\mathrm{C}^{-1}$, and $\gamma = 0.056\,^{\circ}\mathrm{C\,m^2\,W^{-1}}$. a_{1-3} are fitted by assuming eight module data points at $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$ and different irradiances $I_{glob,t}$, indicated by the gray-pointed line (whereby two of them are taken from an exemplary technical module sheet - η at 1000 W m^{-2} ($= \eta_{stc}$) and η at 200 W m^{-2} - and the others approximately interpolated). The resulting black solid line shows the resulting efficiency at $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$. The residual three solid colored curves show the efficiency at module temperatures $T_{mod} = 10\,^{\circ}\mathrm{C}$, $T_{mod} = 50\,^{\circ}\mathrm{C}$, and $T_{mod} = 75\,^{\circ}\mathrm{C}$, computed by using the mentioned α.

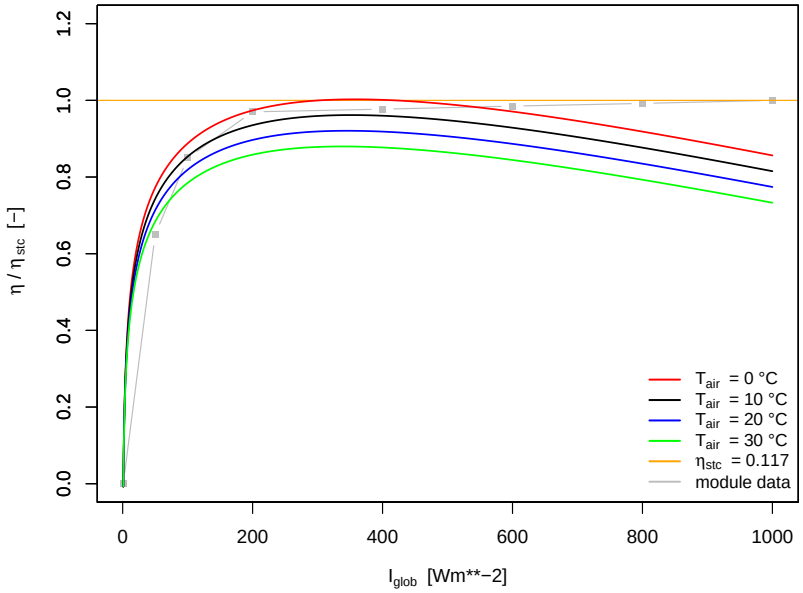


Figure 11: Exemplary PV module efficiency as a function of irradiance and air temperature (instead of module temperature as in Figure 10), shown by the effective efficiency η divided by the efficiency at standard conditions η_{stc}. Further information is given in the caption of Figure 10.

Model by Wagner (2014)

Another PV power production forecast model used in this study is the one by *Wagner* (2014). It uses a different module efficiency approach, which consists of more and of empirically derived coefficients representing the dependency on global irradiance and air temperature. Additionally, unlike the two coefficients α and γ in the model of *Beyer et al.* (2004), the coefficients of this model are linearly separated, which is the reason for its use in this study (explained later).

The PV module efficiency of this second model is defined as follows:

$$\begin{aligned}\eta(I', T') = 1 &+ k_1 \cdot ln(I') \\ &+ k_2 \cdot (ln(I'))^2 \\ &+ T' \cdot \left(k_3 + k_4 \cdot ln\left(I'\right) + k_5 \cdot \left(ln\left(I'\right)\right)^2\right) \\ &+ k_6 \cdot (T')^2\end{aligned} \tag{23}$$

whereby

$$I' = \frac{I_{glob,t}}{I_{stc}} \tag{24}$$

with $I_{stc} = 1000\,\mathrm{W\,m^{-2}}$ being the global irradiance at standard conditions and

$$T' = T_{mod} - T_{mod,stc} \tag{25}$$

with $T_{mod,stc} = 25\,^{\circ}\mathrm{C}$ being the module temperature at standard conditions.

The module temperature T_{mod} is defined differently compared to the previous model:

$$T_{mod} = T_{air} + (T_{noct} - 20\,^{\circ}\mathrm{C}) \cdot \frac{I_{glob,t}}{800\,\mathrm{W\,m^{-2}}} \tag{26}$$

The new parameter T_{noct} is the nominal operating cell temperature (NOCT), which is defined as the module temperature under the "realistic" (or better: more realistic than the laboratory standard conditions) outdoor conditions of $I = 800\,\mathrm{W\,m^{-2}}$, $T_{air} = 20\,^{\circ}\mathrm{C}$, a wind velocity of $1\,\mathrm{m\,s^{-1}}$, and a PV panel being installed free-standing (meaning a good circulation also on the back side). For CdTe and pcSi modules T_{noct} is 45 °C.

For the k-factors, *Wagner* (2014) empirically derived a set of values between -1 and 1 both for CdTe and pcSi modules, which is listed in Table 8. A drawback of this model is thus the fact that it is harder to assign any physical meaning to the k-factors, as it can be done for the temperature coefficient α and the heating coefficient γ of the model by *Beyer et al.* (2004).

Table 8: Literature values for k_{1-6}.

k-value	CdTe module	pcSi module
k_1	$-6.04464 \cdot 10^{-2}$	$-2.98282 \cdot 10^{-2}$
k_2	$-2.86750 \cdot 10^{-2}$	$-2.92604 \cdot 10^{-2}$
k_3	$-1.86757 \cdot 10^{-3}$	$-2.29474 \cdot 10^{-3}$
k_4	$+3.18431 \cdot 10^{-4}$	$+1.60938 \cdot 10^{-4}$
k_5	$-8.40270 \cdot 10^{-5}$	$+5.50864 \cdot 10^{-5}$
k_6	$-9.73482 \cdot 10^{-6}$	$-7.02740 \cdot 10^{-6}$

3.1.4 PV panel area

The lack of the PV panel area A of each individual PV plant (see data section) requires to derive it. To do so, the following approach can be applied (as done by *Lorenz and Heinemann* (2012)):

$$A = \frac{P_{nom}}{\eta_{stc} \cdot 1000\,\mathrm{W\,m^{-2}}} \tag{27}$$

P_{nom} is the nominal power of the PV plant, which means the power the whole PV plant (= all installed PV modules together) would yield under standard conditions. It is one of the only technical parameters that are known for each individual PV plant of this study. The module efficiency at standard conditions η_{stc} is not known and thus has to be estimated.

Equation (27) is also applied in practice (as recognized in a detailed construction plan about a specific PV plant, which ewz could provide for this study): if a PV plant to be installed needs to fulfill certain certification requirements or desires of the owner, the installer receives the value of the nominal power at standard conditions P_{nom}, which should be installed. By knowing the nominal efficiency at standard conditions η_{stc} of the single PV modules planned to be installed, the installer can then decide how many PV modules (of a certain area) have to be installed in order to cover the desired total area A and / or to reach the required P_{nom}.

3.1.5 Time step

In general, the time step for the PV power production forecast is set by the Swiss power market. As mentioned in the introduction, it amounts to 15 min. The temporal resolution of the Swiss power market, however, conflicts with the one of COSMO, having a time step of 1 h only (at least for this study). Therefore, the PV power production forecast is done for a 1 h-time-step mainly due to the following two reasons:

- The main objective of this project is to analyze the various errors and uncertainties associated with a PV power production forecast. This rather qualitative objective can already be reached when looking at the 1 h-time-step.
- The analysis of the PV power production forecast errors on a 15 min basis should be content of a further, more energy-sector-oriented project. It would require a stronger effort on the interpolation

of COSMO forecast data in a temporal resolution other than 15 min to the one of 15 min. These issues cannot be covered here, since they would go beyond the scope of this study.

3.2 Power production forecast sensitivity case study for a single PV plant

Before the computation of power production forecasts for the whole available set of PV plants based on the previously described approaches, a sensitivity case study is done for a single PV plant. The idea of this first step is to understand how sensitive the output of the PV power production forecast is to a change in different key parameters. Furthermore, it helps to understand the general PV power production characteristics on one hand and the results of the power production forecast for the whole set of PV plants on the other hand.

3.2.1 Case study PV plant

The PV plant used for the sensitivity case study has to meet the following two criteria:

- There have to be available as many metadata about the selected PV plant as possible, in order to minimize the uncertainties in the power production forecast due to assumptions about the technical parameters.
- The power production measurement data of the selected PV plant have to be as proper as possible (in terms of accuracy and uncertainties), in order to minimize the proportion of the total power production forecast error induced by wrong power production measurement data.

The combination of the two criteria yielded a relatively large PV plant in the city of Zurich, suitable for the sensitivity study. Out of the construction plans ewz could provide for this PV plant, the set of metadata shown in Table 9 is known.

3.2.2 Procedure

Power production forecast formula

The power production forecast for the sensitivity case study PV plant is computed based on the general Equation (5).

As the first term, the global irradiance on a horizontal plane is converted to the one on the tilted and azimuthally oriented PV panel. The conversion is done componentwise by calculating direct and diffuse irradiance separately.

For the conversion of the direct irradiance on the tilted and azimuthally oriented PV panel, the previously described Equation (6) is used by inserting the according direct model output parameter from COSMO, the direct irradiance on a horizontal plane $I_{dir,h}^{COSMO}$:

$$I_{dir,t}^{COSMO} = I_{dir,h}^{COSMO} \cdot \frac{cos(\theta_i)}{cos(\theta_z)} \tag{28}$$

Table 9: Known technical parameters of the sensitivity case study PV plant.

Parameter	Value	Source
Area per module	1.574 · 0.804 m	Construction plan
Area of the whole plant	1063 m^2	Derived from area per module and number of modules
Area of the whole plant	1073 m^2	Derived according to Equation (27)
Azimuth angle (ψ)	180°	Estimated by sketch on construction plan and from Google Maps
Heating coefficient (γ)	-0.0027 °C m^2 W^{-1}	Value for free-standing panels according to literature
Installation type	Free-standing, installed on a pedestal	Construction plan
Module efficiency at standard conditions (η_{stc})	13.7 %	Construction plan
Module type	CNPV-175M (mcSI) module by the company S.E.T	Construction plan
Nominal power per module	175 W	Construction plan
Nominal power of the whole plant (P_{nom})	147 kW	Construction plan
Number of installed modules	840	Construction plan
Shading angle	16.8°	Construction plan
Temperature coefficient (α)	-0.0045 °C^{-1}	CNPV-175M module sheet (downloaded from the Internet)
Tilt angle (β)	18°	Construction plan

For the diffuse irradiance on the tilted and azimuthally oriented PV panel, the COSMO diffuse irradiance on a horizontal plane, $I_{dif,h}^{COSMO}$, can be inserted into Equation (8):

$$I_{dif,t}^{COSMO} = \frac{I_{dif,h}^{COSMO} \cdot (1 + cos(\beta))}{2} \tag{29}$$

To obtain the global irradiance on the tilted and azimuthally oriented PV panel, the two components from Equations (28) and (29) are summed up according to Equation (9):

$$I_{glob,t}^{COSMO} = I_{dir,t}^{COSMO} + I_{dif,t}^{COSMO} \tag{30}$$

For the module efficiency, the previously introduced model of *Beyer et al.* (2004) is used. Therefore, inserting both the COSMO global irradiance on the tilted and azimuthally oriented PV panel $I_{glob,t}^{COSMO}$ and the COSMO ambient 2 m air temperature T_{air}^{COSMO}, Equation (22) becomes:

$$\eta(I_{glob,t}^{COSMO}, T_{air}^{COSMO}) = \left(a_1 + a_2 I_{glob,t}^{COSMO} + a_3 \ln(I_{glob,t}^{COSMO})\right) \cdot \left(1 + \alpha(T_{air}^{COSMO} + \gamma I_{glob,t}^{COSMO} - 25\,^{\circ}\mathrm{C})\right) \tag{31}$$

Using the terms $I_{glob,t}^{COSMO}$ and $\eta(I_{glob,t}^{COSMO}, T_{air}^{COSMO})$, the final forecast formula for the sensitivity case study PV plant / location becomes:

$$\begin{aligned} E(t) = {} & I_{glob,t}^{COSMO} \\ & \cdot \left(a_1 + a_2 I_{glob,t}^{COSMO} + a_3 \ln(I_{glob,t}^{COSMO})\right) \cdot \left(1 + \alpha(T_{air}^{COSMO} + \gamma I_{glob,t}^{COSMO} - 25\,^{\circ}\mathrm{C})\right) \\ & \cdot \frac{P_{nom}}{\eta_{stc} \cdot 1000\,\mathrm{W\,m^{-2}}} \\ & \cdot \Delta t \end{aligned} \tag{32}$$

The temperature coefficient α, an estimate of the heating coefficient γ, the nominal power P_{nom}, and the module efficiency at standard conditions η_{stc} in Equation (32) can now be extracted from the metadata set in Table 9. The three shape parameters a_{1-3}, however, need to be estimated. This is done by an approximate parameter fitting explained in the general part at the beginning of this section and shown in Figure 10.

Sensitivity calculations

In a first step, the sensitivity ρ of the power production forecast E for the case study PV plant to a change in different parameters X_i during a clear-sky day both in winter and summer is analyzed:

$$\rho_{X_i} = \frac{\delta E}{\delta X_i}, with\ E = f(X_i)\, i \in [1...n] \tag{33}$$

The two clear-sky days are extracted with an algorithm finding points in time, where the satellite measurement data indicate cloud-free conditions and the power production measurement data of the case study PV plant yield a low variability. The idea of selecting these two days is to draw conclusions on the seasonal characteristics of the PV power production forecast sensitivity. Then, for the selected two days, the diurnal power production course of the case study PV plant is forecasted after slightly changing one of the following four parameters X_i (whereas all the other parameters are left constant):

- Tilt angle β
- Azimuth angle ψ
- Temperature coefficient α
- Heating coefficient γ

In a second step, the sensitivity study for the same four parameters is done, however, not only for the two case study days but for the total available time period from 2010 - 2012.

The sensitivity study is done based on both meteorological forecast and measurement data. This means, in one run, the power production forecasts are computed using COSMO forecast data for

global irradiance and air temperature, as described by Equation (32), whereas in another run, satellite (for global irradiance) and surface station measurement data (for air temperature) are used as basis for the "forecast". The idea of the latter, the use of meteorological measurement data, is to eliminate or minimize, respectively, the meteorological uncertainties in the sensitivity study. The surface station used for this step is the SwissMetNet station in Zurich Affoltern, which is approximately 3 km away from the case study PV plant.

3.3 Power production forecast for all PV plants

After the sensitivity case study for one single PV plant, the power production forecast for the whole set of PV plants is computed. Since the set of known technical parameters for each of the 270 PV plants (shown in Table 6) is much smaller than the one of the previously described sensitivity case study PV plant (shown in Table 9), Swiss-wide assumptions have to be made to compute a power production forecast for the whole set of PV plants. In particular, the following parameters are unknown and have to be estimated:

- PV panel tilt angle β
- PV panel azimuth angle ψ
- PV plant area A
- PV module efficiency at standard conditions η_{stc}
- Temperature coefficient α
- Heating coefficient γ

This results in different sets of assumptions and estimations yielding four different PV power production forecast methods, which are described in the following subsections. Three of them are really forecasts, which means they are based on COSMO forecast data, whereas the fourth method is based on satellite and surface station measurements and thus acts as reference method assuming a "perfect" COSMO forecast.

3.3.1 Global assumptions for all forecast methods

Before the specific description of the four PV power production forecast methods, Table 10 summarizes the assumptions, which are common for all methods, and the associated uncertainties.

Table 10: Global assumptions for all PV power production forecast methods.

Topic	Assumption	Uncertainties
Efficiency	The PV module is constantly at its maximum power point.	A PV module contains a MPP tracker, which is constantly steering the product of current and voltage to its conditional optimum. The efficiency of this tracker, however, is not necessarily 1 all the time.
Efficiency	The inverter constantly works at its maximum power and thus has an efficiency of 1.	It is often the case that the power of the inverter is lower than the one of the PV module. Hence, the inverter can act as a limiting factor.
Efficiency	The efficiency of the PV module is constant during its whole life / operation time.	The efficiency decreases with increasing age of the PV module. Manufacturers usually specify a number of 10 % efficiency decrease in 10 years operation time.
COSMO	The available COSMO irradiance parameters are computed on a horizontal plane.	The COSMO irradiance parameters on a horizontal plane follow the model topography, meaning that they are perpendicular to the surface which does not need to be horizontal. However, the maximum slope of the model topography amounts to around 15°, which occurs in very mountainous Alpine regions only. Hence, this bias only gets relevant in these regions where usually only few PV panels are installed.
Irradiance	The ground-reflected irradiance as a third component of global irradiance is zero.	It is appropriate to neglect ground-reflected irradiance for purposes as PV power forecasting in general. However, a high surface albedo due to snow cover can significantly enhance the ground-reflected component and thus the power generation of a tilted PV panel.
Irradiance	The atmosphere is isotropic.	The isotropic assumption approximately holds for overcast situations. However, clear-sky conditions show a strong anisotropic behaviour, since the radiance in the circumsolar region is significantly larger than in the rest of the sky.

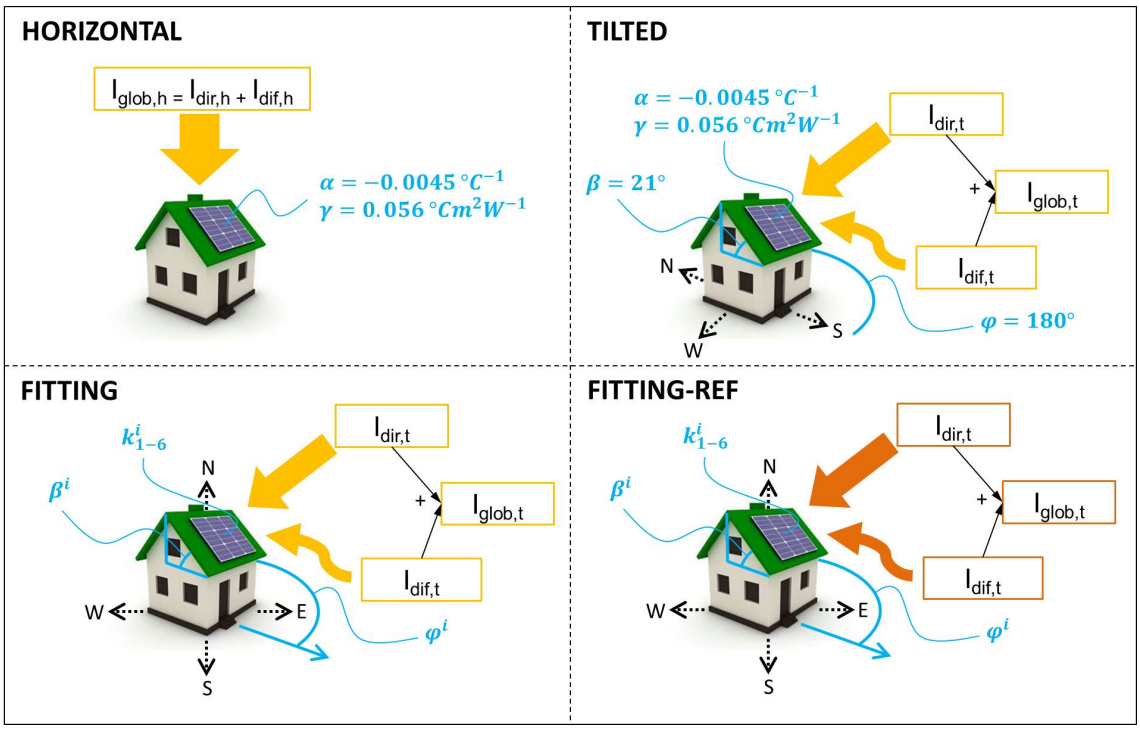


Figure 12: Applied PV power production forecast methods. Details are in the text.

3.3.2 Forecast method HORIZONTAL

This method, sketched at the top left of Figure 12, is characterized by the assumption of all PV plants being horizontally oriented. Herewith, it allows to use the **global irradiance** on a horizontal plane as a direct model output parameter of COSMO, $I_{glob,h}^{COSMO}$, whereby the conversion of the irradiance on a tilted plane is not necessary.

As in the sensitivity case study, for the **module efficiency** the model of *Beyer et al.* (2004) is used. Inserting both the COSMO global irradiance on a horizontal plane $I_{glob,h}^{COSMO}$ and the COSMO ambient 2 m air temperature T_{air}^{COSMO}, Equation (22) becomes:

$$\begin{aligned}\eta(I_{glob,h}^{COSMO}, T_{air}^{COSMO}) = &\left(a_1 + a_2 I_{glob,h}^{COSMO} + a_3 \ln(I_{glob,h}^{COSMO})\right) \\ &\cdot \left(1 + \alpha(T_{air}^{COSMO} + \gamma I_{glob,h}^{COSMO} - 25\,^{\circ}\mathrm{C})\right)\end{aligned} \tag{34}$$

With the described assumptions, the final HORIZONTAL **forecast formula** becomes

$$\begin{aligned}E(t) = &I_{glob,h}^{COSMO} \\ &\cdot \left(a_1 + a_2 I_{glob,h}^{COSMO} + a_3 \ln(I_{glob,h}^{COSMO})\right) \cdot \left(1 + \alpha(T_{air}^{COSMO} + \gamma I_{glob,h}^{COSMO} - 25\,^{\circ}\mathrm{C})\right) \\ &\cdot \frac{P_{nom}}{\eta_{stc} \cdot 1000\,\mathrm{W\,m^{-2}}} \\ &\cdot \Delta t\end{aligned} \tag{35}$$

The residual **metadata parameter values** are set as follows: for the shape parameters a_{1-3}, the same values as in the sensitivity case study are used. The temperature coefficient α is assumed to be $-0.0045\,°C^{-1}$ and the heating coefficient $\gamma = 0.056\,°C\,m\,W^{-1}$ for all PV plants. For the nominal power P_{nom}, the known specific value for each PV plant is used, whereas for the module efficiency at standard conditions η_{stc} a value of 0.117 is inserted.

3.3.3 Forecast method TILTED

For this method, sketched at the top right of Figure 12, the **global irradiance** on a horizontal plane is converted to the one on a tilted and azimuthally oriented PV panel. Due to the lack of information about tilt and azimuth angles of the individual PV plants, a tilt angle β of 21° and an azimuth angle ψ of 180° (which is southward) is assumed for all PV plants. These values are based on the estimation of ewz. The conversion is done componentwise by calculating direct and diffuse irradiance separately, following the same procedure as in the sensitivity case study (Equations (28), (29), and (30)).

Also for the **module efficiency**, the same assumptions as in the sensitivity study are made, which is shown in Equation (31).

Using the terms $I_{glob,t}^{COSMO}$ and $\eta(I_{glob,t}^{COSMO}, T_{air}^{COSMO})$, the final TILTED **forecast formula** becomes the same as for the sensitivity case study:

$$\begin{aligned} E(t) = I_{glob,t}^{COSMO} \\ &\cdot \left(a_1 + a_2 I_{glob,t}^{COSMO} + a_3 \ln(I_{glob,t}^{COSMO})\right) \cdot \left(1 + \alpha(T_{air}^{COSMO} + \gamma I_{glob,t}^{COSMO} - 25\,°C)\right) \\ &\cdot \frac{P_{nom}}{\eta_{stc} \cdot 1000\,W\,m^{-2}} \\ &\cdot \Delta t \end{aligned} \tag{36}$$

3.3.4 Forecast method FITTING

Compared to the methods HORIZONTAL and TILTED, FITTING, sketched at the bottom left of Figure 12, is a mixture of a physical and a statistical method. It is physical in the sense that it uses similar models as in the TILTED method to convert global irradiance on a tilted and azimuthally oriented PV panel and to include the irradiance and temperature dependency of the module efficiency. However, the unknown PV plant parameters as the tilt and azimuth angles of the panel and the irradiance and temperature dependency factors of the module are not chosen through assumptions anymore but by deriving through a non-linear parameter fitting. This parameter fitting can be done for each single PV plant in order to get plant-specific parameters, which is indicated by the i-superscripts on the bottom left of Figure 12. The approach requires irradiance and temperature measurement data from satellite and surface stations on the meteorological side and power production measurement data on the technical side. With the use of this data, the method becomes partially statistical.

The calculation of the **global irradiance** on the tilted and azimuthally oriented PV panel is done in the same way as for the sensitivity case study or the method TILTED: the direct and diffuse components

from Equations (28) and (29) are summed up according to Equation (9) to get the desired global irradiance parameter shown in Equation (30). The only difference from the method TILTED is that FITTING uses PV-plant-specific tilt and azimuth angle values, β^i and ψ^i, instead of making the same general assumptions for all PV plants (which is a tilt angle β of 21° and an azimuth angle ψ of 180°).

For the **module efficiency**, however, FITTING uses the model of *Wagner* (2014) instead of the one from *Beyer et al.* (2004) applied in HORIZONTAL and TILTED. First, the idea was to use also the one from *Beyer et al.* (2004). However, the model turned out to be not very suitable for a fitting approach since the used fitting algorithm experienced difficulties in converging the temperature and heating coefficients (α and γ) towards physically reasonable values. The problem of this model might be the fact that its parameters are multiplied with each other and thus not linearly separated. With the model of *Wagner* (2014), the fitting algorithm performed much better in converging to reasonable values, even though the number of parameters is even higher in this model. The reason for this is probably the fact, that the parameters to be fitted are linearly separated, which is not the case in the model of *Beyer et al.* (2004). With the new model of *Wagner* (2014), shown in Equation (23), the module efficiency based on the COSMO forecast data thus looks as follows (when replacing the two variables I' and T' according to the Equations (24) and (25)):

$$\begin{aligned}\eta(I',T') = 1 &+ k_1 \cdot ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right) \\ &+ k_2 \cdot \left(ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right)\right)^2 \\ &+ \left(T_{mod}^{COSMO} - T_{mod,stc}\right) \cdot \left(k_3 + k_4 \cdot ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right) + k_5 \cdot \left(ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right)\right)^2\right) \\ &+ k_6 \cdot \left(T_{mod}^{COSMO} - T_{mod,stc}\right)^2\end{aligned} \tag{37}$$

According to *Wagner* (2014), the final FITTING **forecast formula** becomes:

$$\begin{aligned}E(t) = &\frac{I_{glob,t}^{COSMO}}{I_{stc}} \\ &\cdot \left(1 + k_1 \cdot ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right)\right. \\ &+ k_2 \cdot \left(ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right)\right)^2 \\ &+ \left(T_{mod}^{COSMO} - T_{mod,stc}\right) \cdot \left(k_3 + k_4 \cdot ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right) + k_5 \cdot \left(ln\left(\frac{I_{glob,t}^{COSMO}}{I_{stc}}\right)\right)^2\right) \\ &\left. + k_6 \cdot \left(T_{mod}^{COSMO} - T_{mod,stc}\right)^2\right) \\ &\cdot P_{nom}^i \\ &\cdot \Delta t\end{aligned} \tag{38}$$

The FITTING method consists of three main steps: the selection of the data for the parameter fitting as the first one, the fitting of the PV-plant-specific parameters as the second one, and the PV power production forecast based on the fitted parameters as the third one.

Step 1: data selection

The parameter fitting is done based on five measured parameters listed in Table 11.

Table 11: Measurement parameters used for the fitting.

Parameter	Abbr.	Unit	Data source	Selection criterion for PV plant location	Temporal resolution
Surface incoming shortwave radiation (direct beam component)	SISDIR	$W\,m^{-2}$	Gridded satellite data	Data from the gridpoint closest to the PV plant	15 min
Surface incoming shortwave radiation (diffuse component)	SISDIF	$W\,m^{-2}$	Gridded satellite data	Data from the gridpoint closest to the PV plant	15 min
Heliosat clear-sky index (1 = clear sky, 0.05 = completely overcast)	KI	-	Gridded satellite data	Data from the gridpoint closest to the PV plant	15 min
PV power production	-	kWh	PV plant measurement by ewz	-	15 min
Air temperature 2m above surface	tre200s0	°C	SMN surface station data	Data from the SMN station closest to the PV plant	10 min

For each parameter, the time series for 2012 is used, which gives around 35'000 data points per parameter. However, only around 5 % of these data points can ultimately be used for the fitting. Therefore, a function is used to choose the proper points in time specifically for each PV plant by applying a set of criteria. This procedure is shown in Table 12, giving the criterion and its computational implementation.

The idea behind step two of Table 12, the application of a so-called running variance on the power production measurement data, is to have a second, more accurate criterion for the selection of cloud-free points in time. This is, the development of the power measurement time series is a better criterion to represent the real-time cloud development right above the PV plant than remote measurements from the satellite. In particular, the second step checks how variable the power production is on an hourly resolution.

The algorithm of the running variance function implements the following four steps, as sketched in Figure 13:

- Load the desired power production time series in 15-minutes resolution
- Detrend the time series
- Compute the variance (of the detrended data) for a 1h-window centered around a data point (this means the window with the two data points before and after the data point) for every data point

Table 12: Criteria and computational implementation of the data selection for the fitting.

Appl. step	Criterion	Computational implementation
1	Daytime	All indices, at which the direct beam radiation (SISDIR) does not contain missing values
2	Absolutely cloud-free sky	All indices, at which the running variance over a moving 1h-window of the power production measurement is below the arbitrary threshold of 0.05
3	Absolutely cloud-free and clear sky	All indices, at which the heliosat clear-sky index (KI) is higher than the arbitrary threshold of 0.99
4	Sun elevation angle not too low	All indices, at which the sum of direct beam and diffuse radiation (SISDIR and SISDIF) is above the arbitrary threshold of 100 $W\,m^{-2}$

of the time series

- Choose all data points, which have a running variance below the arbitrarily set threshold of 0.05. The selected data points should be points in time that do not experience (significant) cloud cover from 30 minutes before and 30 minutes after.

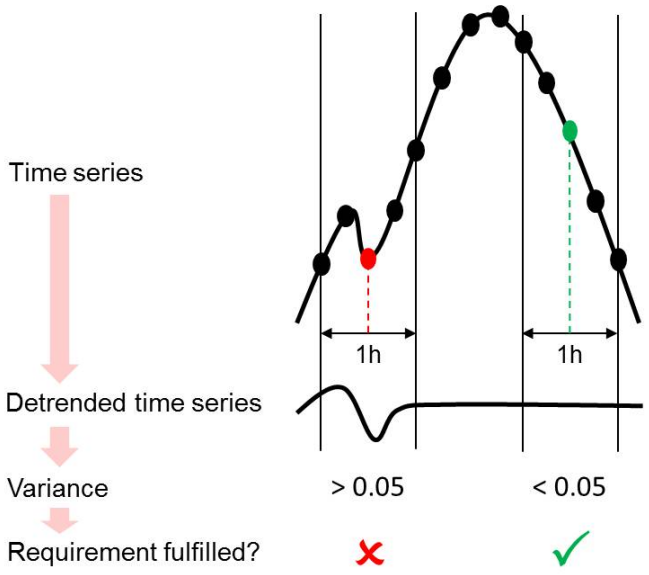


Figure 13: Steps performed by the running variance function. Details are in the text.

Figures 14 and 15 show the selected data points for an exemplary PV plant during two different summer days, a rather good and a rather bad one, as selected by the algorithm explained in Table 11. Details about the procedure are given in the figure caption.

Step 2: parameter fitting

The basic concept of this step is to use the power production forecast function shown in Equation (38) as input for a fitting algorithm, which uses all the known forecast function parameters (as among

others the previously explained satellite and SMN parameters) to fit the unknown parameters (as panel tilt angle, panel azimuth angle, and the irradiance and temperature factors) to the also known power measurement data by applying the non-linear least square approach.

As fitting algorithm, the R function $nls()$ is used. Among others, it requires a list of the function parameters that have to be fitted including the information about what start values as well as upper and lower bounds the fitting algorithm should use for the parameters to be fitted. For the start values for tilt and azimuth angles, the estimated average values by ewz are used, whereas for the k-factors the values for a CdTe PV module, derived by *Wagner* (2014), are set as start values. The lower and upper bounds are set such that the fitted parameters lie within a realistic range. However, this range has to be set as large as possible in order to assure the fitting algorithm can find the real parameters "by itself". Table 13 shows the chosen start values as well as the lower and upper bounds for the eight parameters to be fitted. For some PV plants, the fitting algorithm gets stuck at either the lower or upper bound of tilt and / or azimuth angles. In these cases, the fitting is done again but with the nearly fully possible range of tilt and azimuth angles, which is shown in brackets.

Table 13: Start values, lower bound, and upper bound chosen for the fitting procedure.

Parameter	Start value	Lower bound	Upper bound
Tilt [°]	20	1 (0.5)	89 (89.5)
Azimuth [°]	180	70 (0.5)	290 (359.5)
k_1 [-]	$-6.04464 \cdot 10^{-2}$	-1	1
k_2 [-]	$-2.86750 \cdot 10^{-2}$	-1	1
k_3 [-]	$-1.86757 \cdot 10^{-3}$	-1	1
k_4 [-]	$+3.18431 \cdot 10^{-4}$	-1	1
k_5 [-]	$-8.40270 \cdot 10^{-5}$	-1	1
k_6 [-]	$-9.73482 \cdot 10^{-6}$	-1	1

Step 3: power production forecast

After the derivation of the specific metadata parameters for each PV plant i, the power production forecast is calculated by using the COSMO forecast instead of the measured parameters, but with the same formula shown in Equation (38).

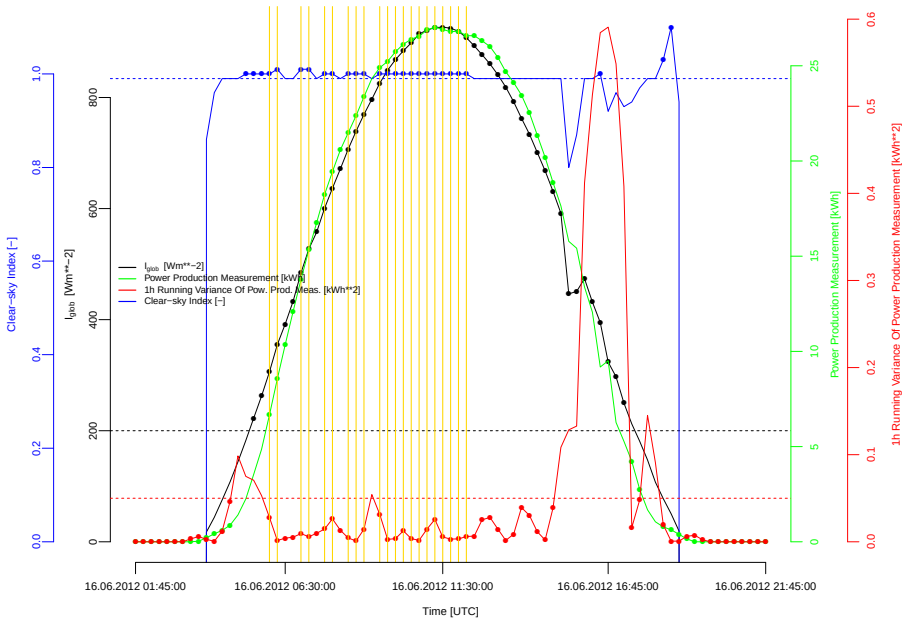


Figure 14: Results of the data points selection algorithm for the FITTING method at an exemplary good day. The black curve is the global irradiance at the location of the PV plant and the green one the produced power of the plant. In red, the running variance of the detrended power measurement is shown. The blue curve shows the development of the heliosat clear-sky index at the location of the PV plant. The dots on the curves indicate, which data points fulfill their own criterion (which means the data points above or below, respectively, the according thresholds indicated by the dashed lines): the black dots are the points in time with global irradiance greater than 200 W m^{-2}, the red and green dots, respectively, the points in time where the running variance of the power production is below 0.05, and the blue dots the points in time with a heliosat clear-sky index greater than 0.99. Ultimately, the vertical yellow lines indicate the data points, which fulfill every criterion (= dots on every of the four lines) and are thus chosen for the fitting.

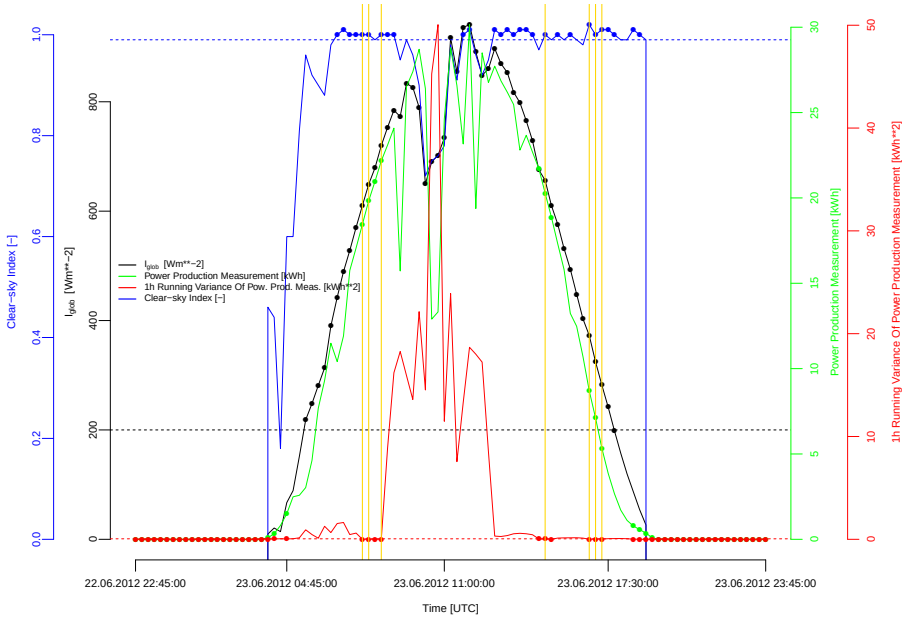


Figure 15: Results of the data points selection algorithm for the FITTING method at an exemplary bad day (details in the caption of Figure 14).

3.3.5 Forecast method FITTING-REF

FITTING-REF represents the power production forecast reference method. It is the same as the FITTING method with regard to the derivation of the metadata parameters through fitting, but is based on meteorological measurement (satellite (SAT) and surface stations (SMN)) instead of forecast data (COSMO). The idea of this method is the elimination of the meteorological uncertainties in the power production forecast by using meteorological measurement data and thus assuming "perfect" COSMO forecasts. With the use of satellite and surface measurement data for irradiance and temperature, the final FITTING-REF **forecast formula** becomes:

$$
\begin{aligned}
E(t) = & \frac{I^{SAT}_{glob,t}}{I_{stc}} \\
& \cdot \left(1 + k_1 \cdot ln\left(\frac{I^{SAT}_{glob,t}}{I_{stc}} \right) \right. \\
& + k_2 \cdot \left(ln\left(\frac{I^{SAT}_{glob,t}}{I_{stc}} \right) \right)^2 \\
& + \left(T^{SMN}_{mod} - T_{mod,stc} \right) \cdot \left(k_3 + k_4 \cdot ln\left(\frac{I^{SAT}_{glob,t}}{I_{stc}} \right) + k_5 \cdot \left(ln\left(\frac{I^{SAT}_{glob,t}}{I_{stc}} \right) \right)^2 \right) \\
& \left. + k_6 \cdot \left(T^{SMN}_{mod} - T_{mod,stc} \right)^2 \right) \\
& \cdot P^i_{nom} \\
& \cdot \Delta t
\end{aligned}
\tag{39}
$$

3.4 Verification

3.4.1 Statistical values used

A single error ϵ of both COSMO and PV power production forecasts at location i and time t is calculated as follows:

$$
\epsilon^i(t) = x^i_{forecast} - x^i_{measurement} \tag{40}
$$

$x^i_{forecast}$ can be either a COSMO or a PV power production forecast and $x^i_{measurement}$ is the measured reference value. For the latter, satellite data is used in case of a COSMO forecast verification and the measured PV power production in case of a PV power production forecast verification.

For the verification of any forecast, mainly the **root mean squared error (RMSE)** is used:

$$
RMSE = \frac{1}{\sqrt{N}} \cdot \sqrt{\sum_{i=1}^{N} (\epsilon^i)^2} \tag{41}
$$

The **relative RMSE** is defined as the RMSE divided by the mean measurement over the same space and time the RMSE is calculated for:

$$RMSE_{rel} = \frac{RMSE}{\overline{x}_{measurement}} \tag{42}$$

This relative RMSE is just one possible definition and is only useful for limited verification perspectives. This means, it can be used to set a certain error in relation to the magnitude of the according observed value *at the same point in space and time*. However, the relative RMSEs are not suitable to compare among each other, since the mean measurement at different points in space and time can be quite different, which could lead to significantly different relative RMSE values even though the according absolute RMSEs were quite similar. Also does a higer relative RMSE value not automatically mean a higher absolute RMSE as well, which is an important fact when interpreting errors from a power market perspective. To give an example: a higher relative RMSE of a PV power production forecast in winter compared to summer does not necessarily mean a higher financial loss for an electricity supply company as ewz in winter. The reason is that the measured PV power production is naturally lower in winter compared to summer, revealing a higher relative RMSE through the division, even though the absolute RMSE might be smaller in winter compared to summer.

Other parameters used for the verification are the **mean error (ME)** and the **mean absolute error (MAE)**:

$$ME = \overline{\epsilon^i} = \frac{1}{N} \cdot \sum_{i=1}^{N} \epsilon^i \tag{43}$$

$$MAE = |\overline{\epsilon^i}| = \frac{1}{N} \cdot \sum_{i=1}^{N} |\epsilon^i| \tag{44}$$

3.4.2 Aggregation

Depending on whether a verification (of both COSMO and PV power production forecasts) is done from a power-market or a plant-specific, rather physical perspective, it is helpful to use different levels of spatial and temporal aggregation to calculate forecast errors (as the RMSE) for a certain set of PV plants (as the one of this study consisting of 270 plants) and a certain time period (as a month). The following three levels are used in this study:

- **No aggregation in space and time**: the forecast errors of every single PV plant at every 1h-time-step are used as a basis for the calculation of the RMSE for the analyzed time period. This yields two types of RMSEs: one is the RMSE for any single (specific) PV plant per 1h-time-step and another one is the RMSE per PV plant (refering to the whole set of PV plants) and 1h-time-step.
- **Spatial aggregation**: the forecast errors of every single PV plant are first summed up at each 1h-time-step to get the spatially aggregated error for all PV plants together at each time step. These spatially aggregated errors are then used as a basis for the RMSE for all PV plants together and per 1h-time-step. From a power market perspective, this is the most important value: an energy

service company as ewz is interested in the forecast error for the total covered region (rather than the one of a single PV plant) for each time step since this is the deviation that has to be paid ultimately.

- **Spatial and temporal aggregation**: the forecast errors of every single PV plant are first summed up at each 1h-time-step to get the spatially aggregated error for all PV plants together at each time step. These spatially aggregated errors are then summed up again over the whole time period. This yields one single value, which is the forecast error for all PV plants together over the whole time period.

3.4.3 Interpretation of PV power measurement data

Importance of a correct interpretation

The power production measurement data in 15-minutes resolution serves as basis for the verification of the computed PV power production forecasts. Since the verification is made for every single PV plant, the interpretation of the outcoming power production forecast errors requires comprehension of quality and characteristics of the power production measurement data from the according PV plants. There are several technical and meteorological factors, which can affect the power production and lead to abnormal power production measurement data and hence power production forecast error patterns. Some of them having occured during the study are listed in the following.

Occuring abnormal patterns

The following types of patterns were found in the PV power production measurement data available for this study:

1. One or several constant low values that are repeated alternately with zero over a time period of several days
2. One or several constant low values over a time period of several days
3. A value of zero over a time period of several days
4. Very low values over a time period of several days
5. Significantly different (lower) values during a certain time window compared to the rest of the day

The reasons for these patterns can be diverse. Based both on discussions with ewz and other experts and on own speculation, the following aspects could be potential reasons: one of the most important reasons is snow covering PV panels. Depending on the panel tilt angle, the thickness of the snow layer, the module and air temperatures, and the weather conditions in general, a snow layer can stay on a PV panel for days, weeks, or even months, preventing radiation to enter the PV module and thus to limit or even shut down power production. This is mainly reflected by production data pattern as the third or fourth. Also could snow and the associated load even damage a PV module, which could be reflected by patterns as the first or second one. Other reasons for the first two patterns, which sometimes also occur during the night, could be some energetic disturbances from the environment of the PV plant, as external light sources, or any electrical reactions within the PV module. The reason for the fifth pattern can be shading by topography or by any other objects nearby.

Approaches to detect error patterns

The challenge in using these PV power production measurement data is the fact that some of the previously mentioned patterns are naturally caused, as for instance the low production due to snow cover or due to shading. Other irregularities include PV module defects or false transmission of production data. An analysis of the former patterns is crucial since they indicate forecast errors caused by model deficiencies (e.g. missing snow cover or shading parameterization).

An attempt to detect irregular power production measurement patterns could look as follows: for each PV plant the power production measurement data is compared to the global irradiance from the satellite data at the according location over several months of the desired time period. Out of this comparison, the monthly correlation coefficient between power production and global irradiance is calculated for each PV plant. Choosing the PV plants with low correlation coefficients (below a certain threshold) allows to make a first selection of PV plants with potentially critical measurement data (this is because high global irradiance normally coincides with high power production - if this is not the case, there is probably a problem with the PV plant (defects, shading etc.)). In the next step, a function is used to detect days during those a PV plant yields power production measurement values greater than zero but with a variance of zero over the whole day. By applying this test, power production measurement data yielding constantly one specific value (as the previously described second pattern) can be detected. By setting a threshold of for instance five days, ultimately all the PV plants with a constant low power production measurement value over a time period equal to or longer than the set time period can be extracted.

The extracted set with potentially wrong measurement data can additionally be filtered by looking at the satellite parameter snow mask, which allows to state whether during a certain period of days or weeks snow could have been lying on a PV panel (if snow mask equals 1) or not (if snow mask equals 0).

Dealing with error patterns in this study

The aforementioned approaches were applied on all PV power production measurement data of this study but the fitting was only successful for a subset of the PV plants. Due to the diversity of potentially wrong patterns within the available data, significantly more effort would have to be taken in order to detect all power production measurement errors automatically with an algorithm (which would be enough content for a potential further study). Therefore, the PV power production measurement data available for this study are all used without filtering. This means that parts of the PV power production forecast errors in this study are due to wrong power production measurement data.

4 Results and discussion

The following results and discussion section is structured in three parts. In a first short part, the meteorological uncertainties influencing a PV power production forecast are analyzed with some results of the COSMO-2 verification with respect to global irradiance. The second part contains the results of the sensitivity case study described in the methods section. It aims at helping to understand the characteristics of the applied PV power production forecast model and hence the results of the PV power production forecast for all available PV plants (as one of the ultimate objectives of this study), which are discussed in the third and last part.

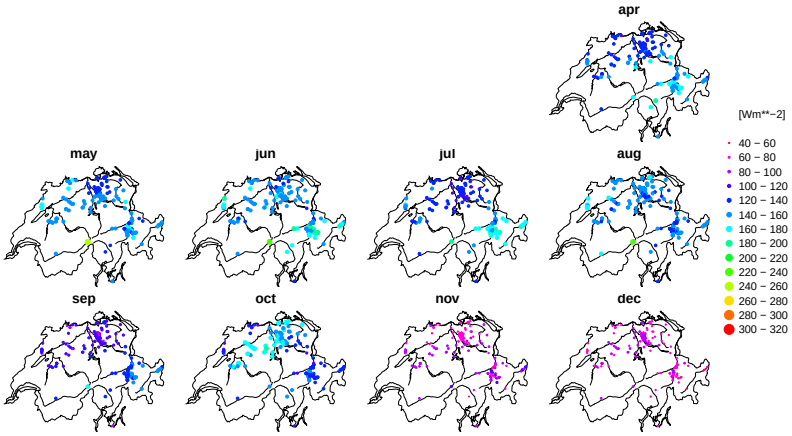
4.1 COSMO verification

The COSMO performance in forecasting global irradiance strongly depends on local weather conditions. In absolute terms, the performance is worse in summer than in winter, whereas in relative terms it is vice versa. In winter, fog or high fog conditions and concurrent forecast biases of COSMO with respect to forecasting global irradiance cause the drop in performance. Northeastern Switzerland, including the region of Zurich, seems to be one of the core regions facing this problem.

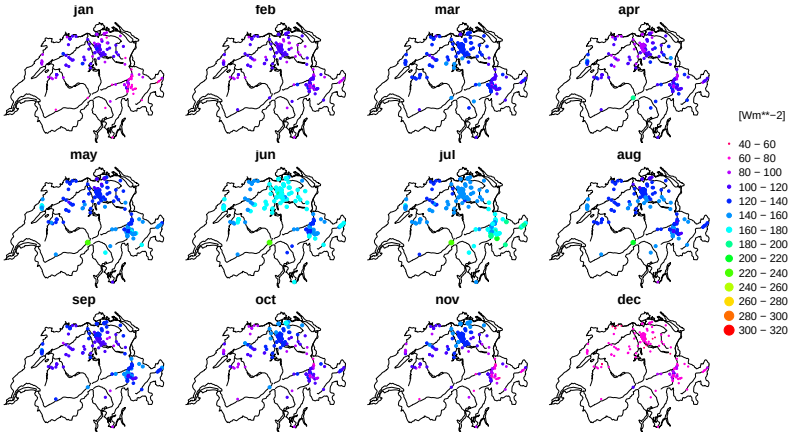
* * *

Figures 16a, 16b, and 16c show the hourly absolute RMSE of the COSMO global irradiance at each PV plant location over the three years based on the satellite as reference (details about the uncertainties of the satellite data can be found in *Stoeckli* (2013)). The COSMO global irradiance is represented by the sum of the direct and diffuse components ASWDIR_S and ASWDIFD_S, whereas SIS is the used satellite parameter (details in the data section). The values range from 40 to 320 $W\,m^{-2}$. The winter months reveal a generally smaller RMSE than the summer months, which is a consequence of the higher shortwave irradiance during summer. Inter-annual differences in the RMSE pattern between same months can be quite large, indicating the COSMO performance to depend on weather conditions to a large degree. July as one such example shows RMSE values in northeastern Switzerland almost 50 % higher in 2012 than in 2010. Furthermore, the RMSE values in the different regions of the Swiss Plateau are quite homogeneous, whereas in the Alps they are more heterogeneous. In Upper Valais, for instance, the RMSE partly reaches double the value of Under Valais. On one hand, this might be due to the mountainous topography as a crucial difficulty for a NWP model as COSMO also in terms of solar irradiance. On the other hand, the applied verification method - the use of the satellite data grid point being closest to the PV plant location to be verified - can yield biases due to height differences: a certain PV plant location can be in a valley, its assigned closest satellite grid point, however, several hundred

(a) 2010



(b) 2011



(c) 2012

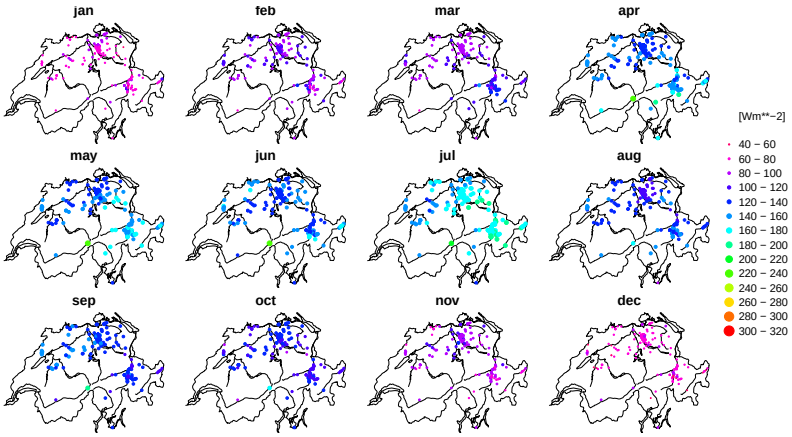


Figure 16: Hourly absolute RMSE (in W m^{-2}) of the COSMO global irradiance at each PV plant location over the three years based on the satellite as reference. The COSMO global irradiance is represented by the sum of the direct and diffuse components ASWDIR_S and ASWDIFD_S, whereas SIS is the used satellite parameter. A higher RMSE is indicated both by a changing color and a larger size of the circle. Since each map shows the RMSE at around 270 PV plant locations, circles are often overlapping, especially in the region of Zurich and the canton of Grisons.

meters higher on a mountain chain. Since these two locations can have quite different characteristics in terms of global irradiance (shading, fog etc.), the corresponding COSMO forecast error can get biased, both in a positive and a negative direction.

Figures 18a, 18b, and 18c show the relative RMSE, which is calculated by dividing the absolute RMSE by the satellite-measured mean global irradiance (SIS) over the according month. Compared to the absolute, the seasonality pattern reverses in the relative RMSE: the winter are higher than the summer values and amount to more than 100 % in some months. Hence, relative to the amount of solar irradiance, COSMO generally seems to perform worse in winter than in summer, even though it can strongly vary in an inter-annual perspective.

Figures 19a, 19b, and 19c show the histograms of the absolute errors of the global irradiance forecast of COSMO for the individual months. One histogram thus contains the absolute errors of every PV plant (around 270) at every hour of the respective month. Most of the months reveal a positive mean error (which is the red line on the right-hand side of zero). Exceptions are July 2010, April, May, and September 2011, and February and March 2012. These are the months with low relative RMSE values, as shown in Figures 18a, 18b, and 18c. Since around one third of the analyzed PV plant locations are in the region of Zurich, one of the main conclusions could be that COSMO tends to overestimate global irradiance and thus has a slightly positive bias for this region.

Fog or high fog is likely to be one of the origins for regional differences in the COSMO performance, what can be indicated by two examples: in November 2011, especially the relative RMSE is significantly higher in northeastern Switzerland than in the Alps (Figures 16b and 18b). Persistent high pressure conditions during this month supplied the higher elevated regions with record amounts of sunshine duration, whereas in the Swiss Plateau (especially the eastern part) this was not the case due to persistent fog (*MeteoSwiss*, 2011). Figure 17 showing the deviation of sunshine durations from the norm period supports this hypothesis. Another example is October 2010, which reveals a similar pattern (Figures 16a and 18a). Also during this month, the Swiss lowland got much less sunshine than the Alpine regions due to persistent fog (*MeteoSwiss*, 2010a). Hence, forecasting fog or high fog (still) seems to be an important weakness of the COSMO performance influencing solar irradiance.

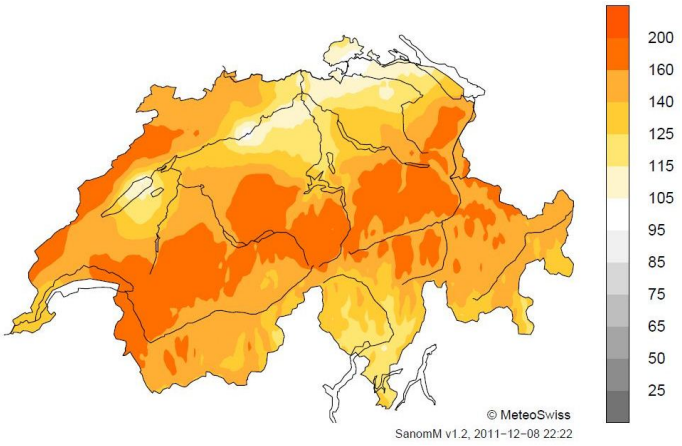
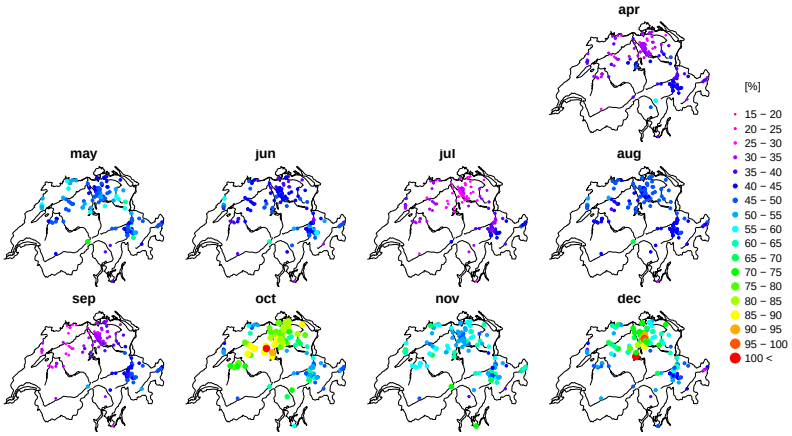
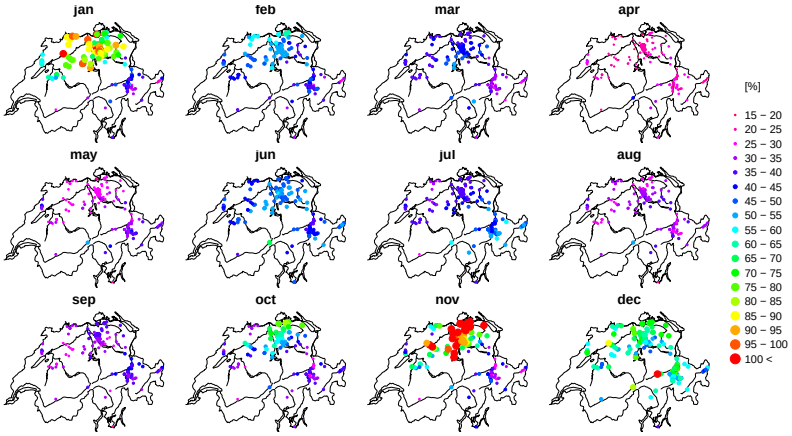


Figure 17: November 2011 sunshine duration in % of the norm, with the reference period of 1961 - 1990. *MeteoSwiss* (2011).

(a) 2010



(b) 2011



(c) 2012

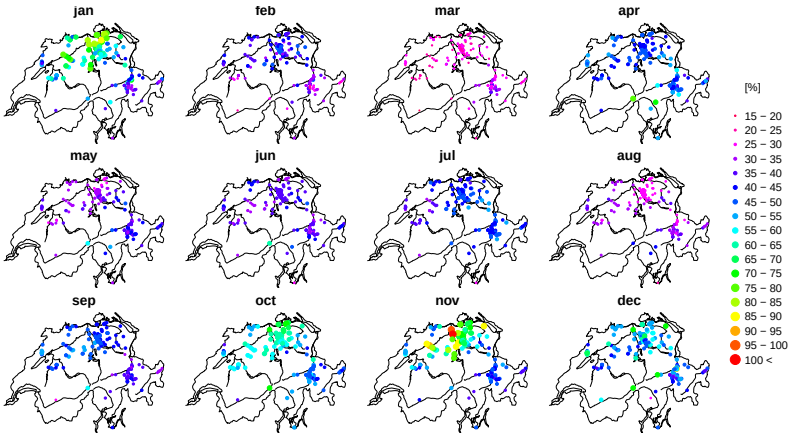
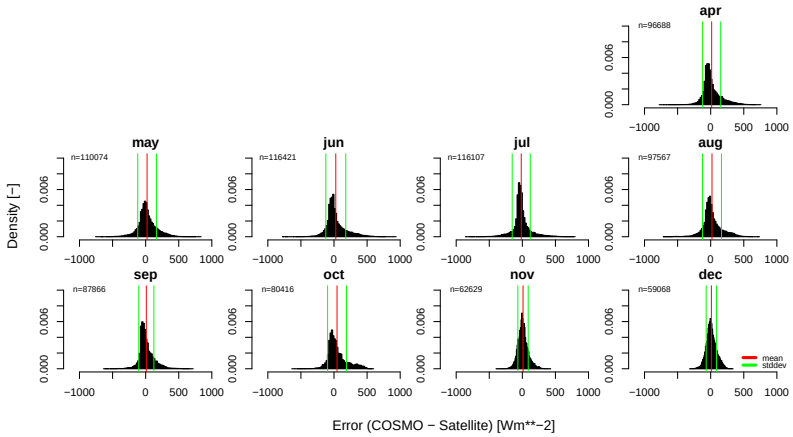
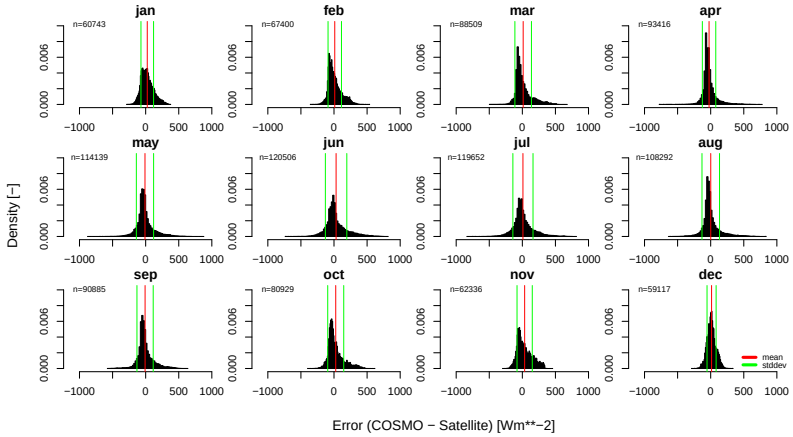


Figure 18: Hourly relative RMSE (in %) of the COSMO global irradiance at each PV plant location over the three years based on the satellite as reference, which is calculated by dividing the absolute RMSE by the satellite-measured mean global irradiance over the according month. The COSMO global irradiance is represented by the sum of the direct and diffuse components ASWDIR_S and ASWDIFD_S, whereas SIS is the used satellite parameter. A higher relative RMSE is indicated both by a changing color and a larger size of the circle. Since each map shows the relative RMSE at around 270 PV plant locations, circles are often overlapping, especially in the region of Zurich and the canton of Grisons.

(a) 2010



(b) 2011



(c) 2012

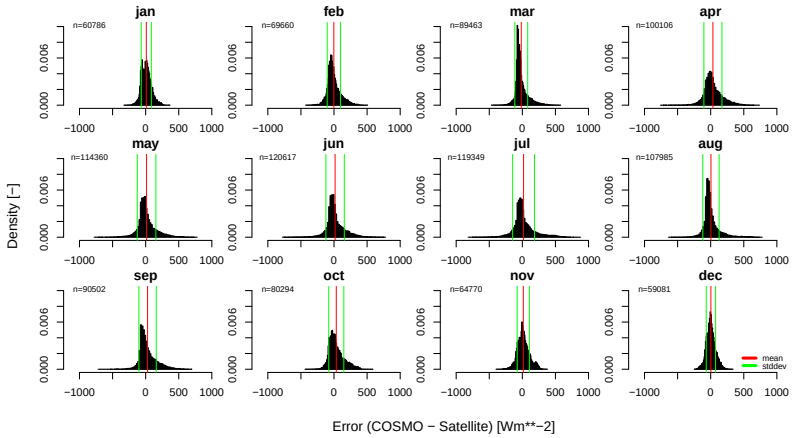


Figure 19: Histograms of the absolute errors (in W m^{-2}) of the global irradiance forecast of COSMO for the individual months. One histogram contains the absolute errors of every PV plant (around 270) at every hour of the according month. The red lines represent the mean error or bias and the green lines range the mean minus and plus one standard deviation. The unit of the y-axis is the relative density.

4.2 Power production forecast sensitivity case study for a single PV plant

The differences in the solar zenith angle, global irradiance, and air temperature between winter and summer yield different sensitivities of the power production forecast to technical parameters in the two seasons: in winter, panel tilt and azimuth angle specifications have the largest impact, whereas the PV module temperature and heating coefficients do not affect the power production forecast at all. In summer, the latter two coefficients become also important due to higher irradiance and temperature levels. Wrong assumptions about technical coefficients can lead to both increasing and decreasing power production forecast errors. The latter can occur when wrong assumptions of technical coefficients artificially compensate COSMO global irradiance forecast errors.

* * *

As described in the methods section, a sensitivity case study is done for a single PV plant prior to the computation of power production forecasts for the whole available set of PV plants. This serves to understand, how sensitive the output of the PV power production forecast is to a change in different key parameters. Furthermore, it helps to understand the general characteristics of PV power production and its response to environmental drivers and technical constraints. The first part of the case study deals with the sensitivity of the power production forecast for a single day, whereas in the second part the same but over the whole time period of the three years is analyzed. The analyzed parameters are the tilt angle β, the azimuth angle ψ, the temperature coefficient α, and the heating coefficient γ.

4.2.1 Sensitivity from a diurnal cycle perspective

This first part contains the results of the sensitivity case study from a diurnal perspective. The two selected days, January 15, 2012, and June 16, 2012, should represent typical clear-sky conditions in winter and summer. The comparison of the global irradiance forecast of COSMO with the measured one by the satellite for the two case study days is shown in Figure 20: there is a clear difference between the satellite and COSMO in both cases, which is an important fact for the following sensitivity studies.

The figures of the subsequent four paragraphs (including the results of the four analyzed parameters) contain the following information: the top half shows the results for winter, the bottom half the ones for summer. The first row of both the winter and summer blocks shows the forecasted power production (in kWh) based on the different parameter perturbations (colored lines) as well as the measured power production (black line) over the hours of the day (in UTC). On the second row, the development of the difference between the power production forecast and measurement is displayed. The two numbers (after the arrow) in the legends of these second rows mean the following: the first number is the accumulated absolute difference (sum of every hourly error of the day) and the second number (in brackets) the accumulated measured power production (sum of every measured hourly power production of the day). All subfigures in the left column represent the results based on satellite (for global irradiance) and surface station (for air temperature) measurements, whereas the ones in the right column are calculated with the use of COSMO forecasts.

(a) January 15, 2012 **(b)** June 16, 2012

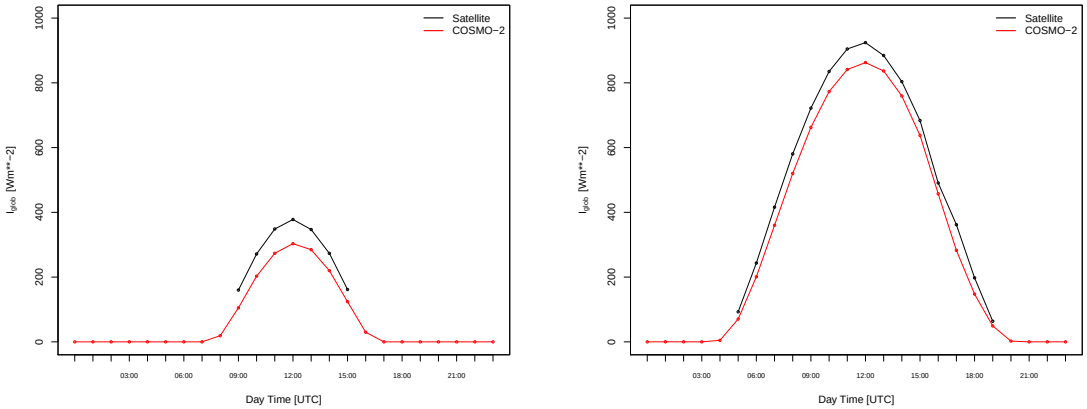


Figure 20: Global irradiance (in W m^{-2}) forecast of COSMO (red line) and measurement of satellite (black line) for the two clear-sky winter and summer days. COSMO underestimates the global irradiance on both days.

Tilt angle (β) sensitivity: $\rho_\beta = \frac{\delta E}{\delta \beta}$

First of all, Figure 21 indicates the consequence of the underestimation of global irradiance by COSMO for the power production forecast: the forecast based on the assumptions closest to reality, which is a tilt angle of 18°, are significantly below the measured power production. Using the satellite and surface measurements, however, the power production forecast based on the same assumption gets closest to the measurement. This can also be interpreted as a verification of the PV power production forecast model, showing that the model captures the most relevant processes and parameters associated with the power production.

As shown in Figure 21, the tilt angle sensitivity is highest in winter, with hourly values of up to 1 kWh per ° change in β around solar noon. Since the peak power production of the PV plant at solar noon amounts to around 80 kWh per hour, this is a relatively high value. In summer, the sensitivity is much lower.

The reason for the difference is displayed in Figure 22: optimal (perpendicular) solar irradiance for power production is given when the tilt angle of the PV panel is equal to the solar zenith angle θ_z. The high solar zenith angle in winter, reaching 70° at solar noon only, is clearly higher than the investigated tilt angles. The closer the tilt gets to the solar zenith angle, the more power production output the model yields, as shown in Figure 21 by the red curve yielding the highest and the blue curve the lowest power production. In summer, however, the minimum solar zenith angle of 25° at solar noon lies between the tilt angles of 18° and 30° and thus causes the highest power productions for these two cases, followed by the angles 40° and 10°, and finally 0° and 50° (as it can be seen Figure 21). According to the principle explained for winter, the early morning and late afternoon hours in summer should also yield the highest power production for the highest tilt angles. However, with the highest power production in the horizontal case (tilt angle of 0°; blue curve in Figure 21) the opposite occurs. The reason for this is probably the self-shading of a tilted and southward-oriented PV panel in summer, as it occurs in the early morning and late afternoon when the solar azimuth angle is slightly northeast and northwest, respectively, causing the direct solar irradiance to fall on back of the panel. The less the PV panel is tilted, the weaker its self-shading is. Therefore, the blue curve in Figure 21 is highest during these hours, but becomes almost lowest in the middle of the day, when the aforementioned effect of the solar

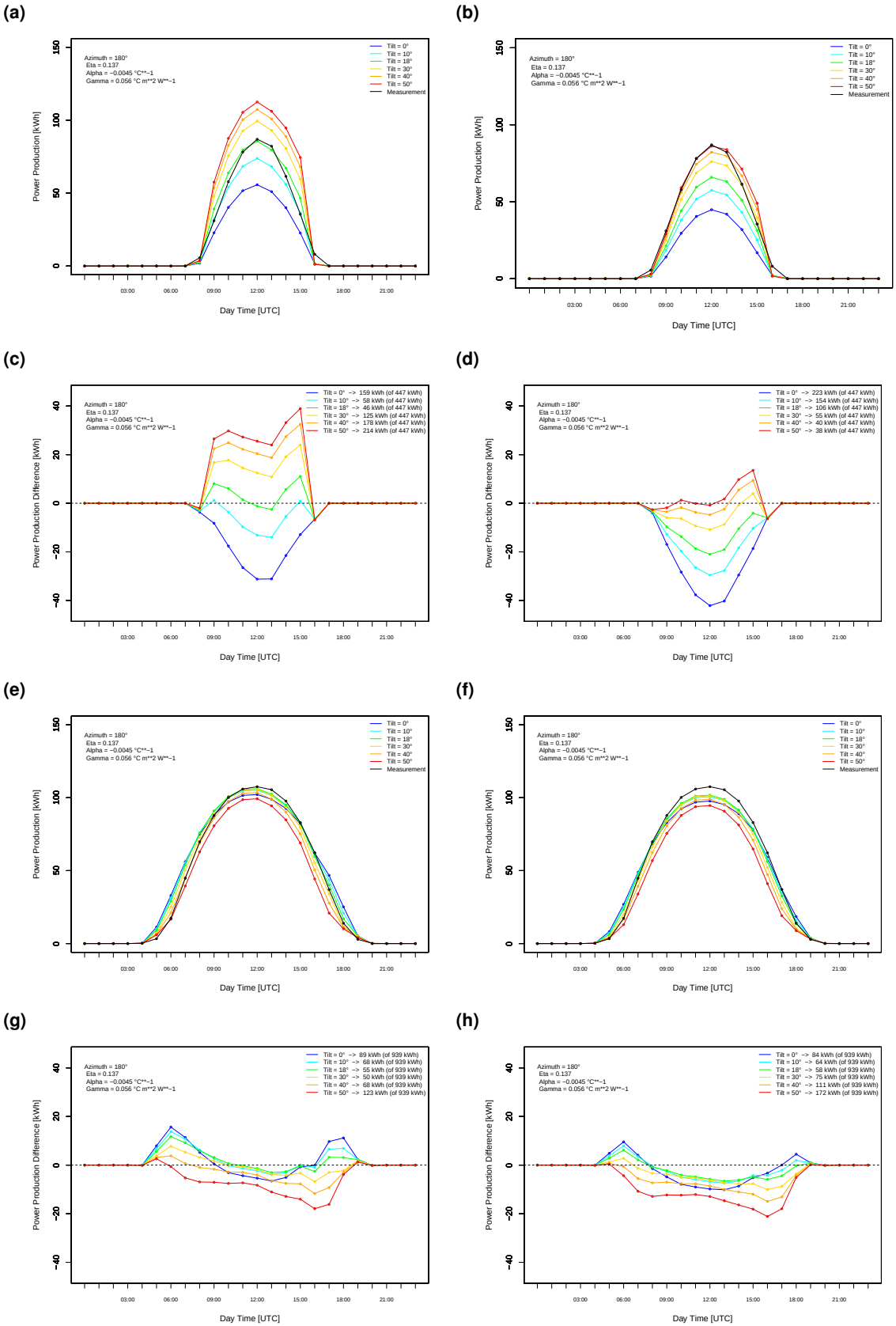


Figure 21: Tilt angle sensitivity $\rho_\beta = \frac{\delta E}{\delta \beta}$. Winter: **a) - d)**. Summer: **e) - h)**. Power production forecasts (colored lines) and measurement (black line): **a), b), e), and f)**. Error of the forecasted power production: **c), d), g), and h)**. Left column: satellite (for global irradiance) and surface station (for air temperature) measurements. Right column: COSMO forecasts.

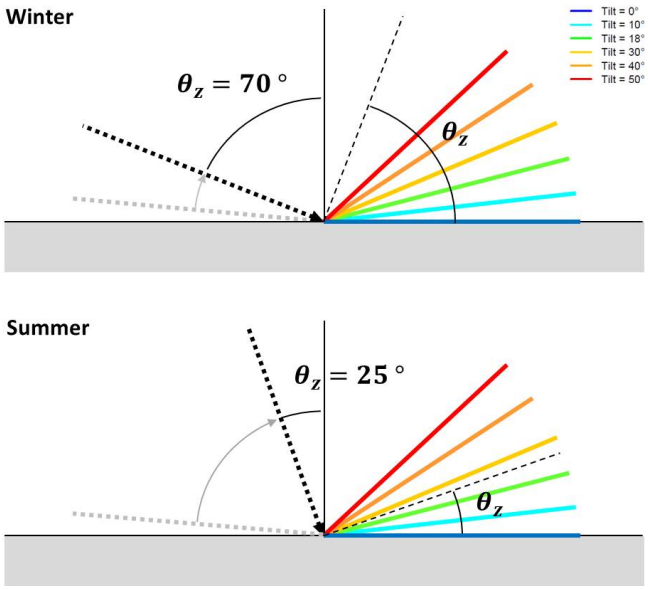


Figure 22: Typical winter and summer solar zenith angles (θ_z; dashed lines) in relation to different tilt angles (β). The solar zenith angle at sunrise and sunset is colored in gray, the one at solar noon in black. The grey arrow thus indicates the diurnal course of the sun.

zenith angle comes into play.

The errors on the left-hand side of Figure 21 further show the performance of the forecasts with the different tilt angles in absolute terms over the day: in winter, the forecast with the (real) tilt angle of 18° generates a daily aggregated error of 46 kWh and the one with 50° one of 214 kWh. With a daily aggregated power production of 447 kWh, the two errors make up 10 % and 47 %, respectively, which is a quite big difference. In summer, the worst forecast with the tilt angle of 50° yields an aggregated error of 123 kWh, which is 13 % of the daily production of 939 kWh.

The power production forecast based on COSMO for the winter day assuming high tilt angles (40° and 50°) shows that errors of COSMO forecasts can be “artificially” compensated by wrong assumptions about PV plants (as in this case the tilt angle, which is higher in the model than 18° in reality). Hence, even though COSMO might reveal a large forecast error in a relative sense in certain situations, the power production forecast model does not necessarily need to perform badly in these situations. The same is valid the other way round. Therefore, assuming the same tilt angle for every PV plant in Switzerland, leads to power production forecast results that have to be interpreted with caution. The following example should illustrate this: a power production forecast is computed for each PV plant of Switzerland for a sunny winter day, assuming the same tilt angle of 20° for all plants. In reality, PV plant A has a tilt angle of 10° and PV plant B one of 20°. Under these conditions, the situations shown in Tables 14 and 15 can occur for the two PV plants.

The reason for the positive power production forecast error of PV plant A resulting from a COSMO irradiance forecast error of zero is the tilt angle assumption of 20°, being too high for plant A and hence generating too much power production output because of the low solar elevation angle in winter. If the COSMO forecast underestimates the irradiance at location A, however, no power production forecast error occurs, because the tilt angle overestimation in the model compensates the COSMO irradiance underestimation to reveal a net power production forecast error of zero.

Table 14: Hypothetical situations for PV plant A (real tilt angle: 10°; assumed tilt angle: 20°).

COSMO irradiance forecast error	PV power production forecast error
Zero	Positive
Negative	Zero

Table 15: Hypothetical situations for PV plant B (real tilt angle: 20°; assumed tilt angle: 20°).

COSMO irradiance forecast error	PV power production forecast error
Zero	Zero
Negative	Negative

In terms of power production yield, the comparison of the winter and summer forecasts based on COSMO in Figure 21 finally reveals the following interesting aspect: with a tilt angle of 50° in winter, the same peak power production at solar noon of around 110 kWh per hour as in summer could be reached. Aggregated over the whole day, of course, the summer production can not be reached. Nevertheless, this might be an important fact to keep in mind when constructing new PV plants in Switzerland.

Azimuth angle (ψ) sensitivity: $\rho_\psi = \frac{\delta E}{\delta \psi}$

Figure 23 shows the sensitivity of the power production forecast to the change of the azimuth angle. For the angles of 180° ±20°, the sensitivity is only moderate all over the day. Using more extreme angles, southeast to east ($\leq$135°) and southwest to west ($\geq$225°), increases the sensitivity. The largest differences to the power production measurement occur in the morning for westward-oriented and in the afternoon for eastward-oriented PV panels. The reason is again the self-shading of the PV panel, which occurs due to the low solar elevation angles in winter inducing the direct irradiance falling on the PV panel from behind (even though the panel is tilted by 18° only). In summer, the sensitivity is moderate and approximately equal for the whole range of azimuth angles between eastward and westward. This is because compared to winter solar elevation angles get much higher in the morning and afternoon, whereby self-shading is not a dominant problem anymore since due to the low tilt angle of 18° the direct irradiance can fall onto the panel from above in the early morning (in the late afternoon) already (still). Both in winter and summer, the smallest sensitivity in general occurs at solar noon, which is due to the highest possible solar elevation angle at this time.

The relatively small sensitivity for azimuth angles close to the real 180° in winter can also be recognized by looking at the error curves in Figure 23: the daily aggregated error for all the three azimuth angles between 160° and 200° are basically the same (45 - 47 kWh, which is 10 % of the daily production), whereas they get much higher for the extreme azimuth angles (up to 202 kWh for 90°, which is 45 % of the daily production). In summer, the aggregated error increases more linearly from the azimuth angle of 180° to the extreme ones. The lowest aggregated error of 52 kWh accounts for 5 %, the highest one of 209 kWh for 22 %.

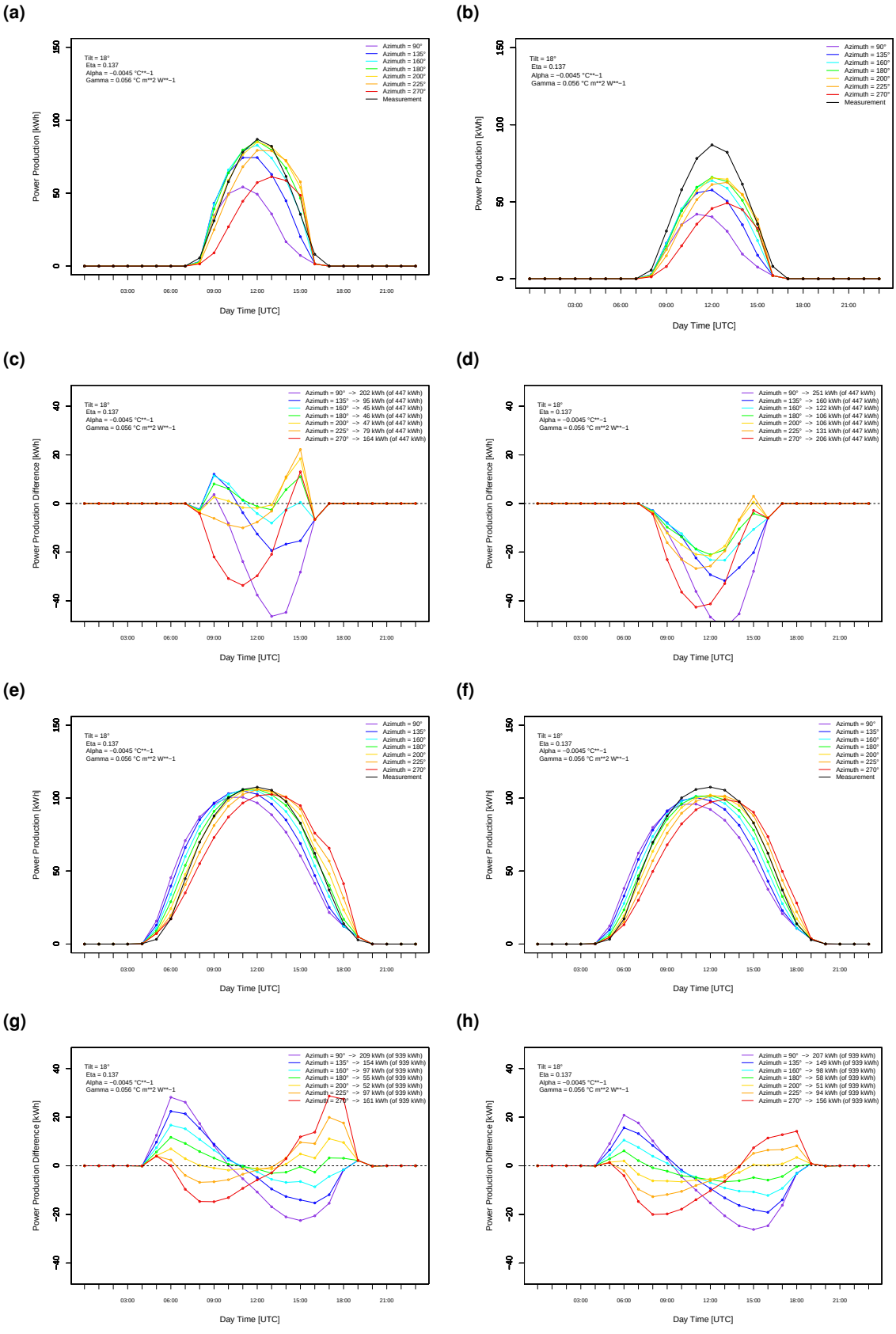


Figure 23: Azimuth angle sensitivity $\rho_\psi = \frac{\delta E}{\delta \psi}$. Explanations are in the caption of Figure 21.

Temperature coefficient (α) sensitivity: $\rho_\alpha = \frac{\delta E}{\delta \alpha}$

As Figure 24 shows, the temperature coefficient sensitivity is almost zero in winter but relatively high in summer. The reason for this can be found in the following two equations, introduced in the methods section ((20) and (21)):

$$\eta(I, T_{mod}) = \eta(I, T_{mod} = 25\,^\circ\mathrm{C}) \cdot (1 + \alpha(T_{mod} - 25\,^\circ\mathrm{C})) \tag{45}$$

with the module temperature approximation

$$T_{mod} = T_{air} + \gamma I \tag{46}$$

Assuming the (measured) conditions of the analyzed winter day - air temperatures (T_{air}) of 0 °C and global irradiance values (I) of up to 600 W m^{-2} around solar noon - the resulting module temperature (T_{mod}) is around 34 °C. With this value, the term $(T_{mod} - 25\,^\circ\mathrm{C})$ that is multiplied with any α gets 9. Hence, the module efficiency change factor $(1 + \alpha(T_{mod} - 25\,^\circ\mathrm{C})$ stays close below 1, whereby the power production forecasts reveal only slight differences using different α.

With air temperature values of 27 °C and global irradiance values of up to 990 W m^{-2} as the solar noon conditions of the summer day, the resulting module temperatures amount to 82 °C. Inserting this value into the module efficiency change term $(1 + \alpha(T_{mod} - 25\,^\circ\mathrm{C})$ yields factors between 0.57 (for the most negative α) and 0.91 (for the least negative α), which means a reduction of the module efficiency of 43 to 9 %. Accordingly, the power production forecasts based on different temperature coefficients result in significantly different values as most pronounced at solar noon. The highest power production is forecasted when using the highest (or least negative) α.

The comparison of the results based on satellite and surface measurements with the ones based on COSMO further show that the temperature coefficient sensitivity depends on irradiance and temperature even within a certain season. So is the sensitivity in the summer and winter forecasts based on the measurements slightly higher than in the ones based on COSMO, because COSMO predicts lower global irradiance values than the satellite measures on both days (Figure 20).

Considering the error curves of Figure 24, the daily aggregated error of the power production forecast based on $\alpha = -0.0045\,^\circ\mathrm{C}^{-1}$ of 55 kWh is the smallest. This result confirms that it makes sense to use this α, which is also given in the technical sheet for the PV module type installed on the case study PV plant.

Heating coefficient (γ) sensitivity: $\rho_\gamma = \frac{\delta E}{\delta \gamma}$

As it can be seen in Figure 25, the heating coefficient sensitivity qualitatively looks quite similar to the temperature coefficient sensitivity: it is quite low in winter but significantly higher in summer. The reason is, as for the temperature coefficient sensitivity, the difference in the irradiance intensity between winter and summer but also within the day itself. Since the heating coefficient only flows into Equation (46), its sensitivity only depends on the irradiance but not additionally on the air temperature as it is the case for the temperature coefficient sensitivity.

Considering the daily aggregated power production forecast error using different γ, the best winter forecast is reached when using $\gamma = 0.065\,^\circ\mathrm{C\,m^2\,W^{-1}}$ (yields a daily error of 45 kWh) and the best

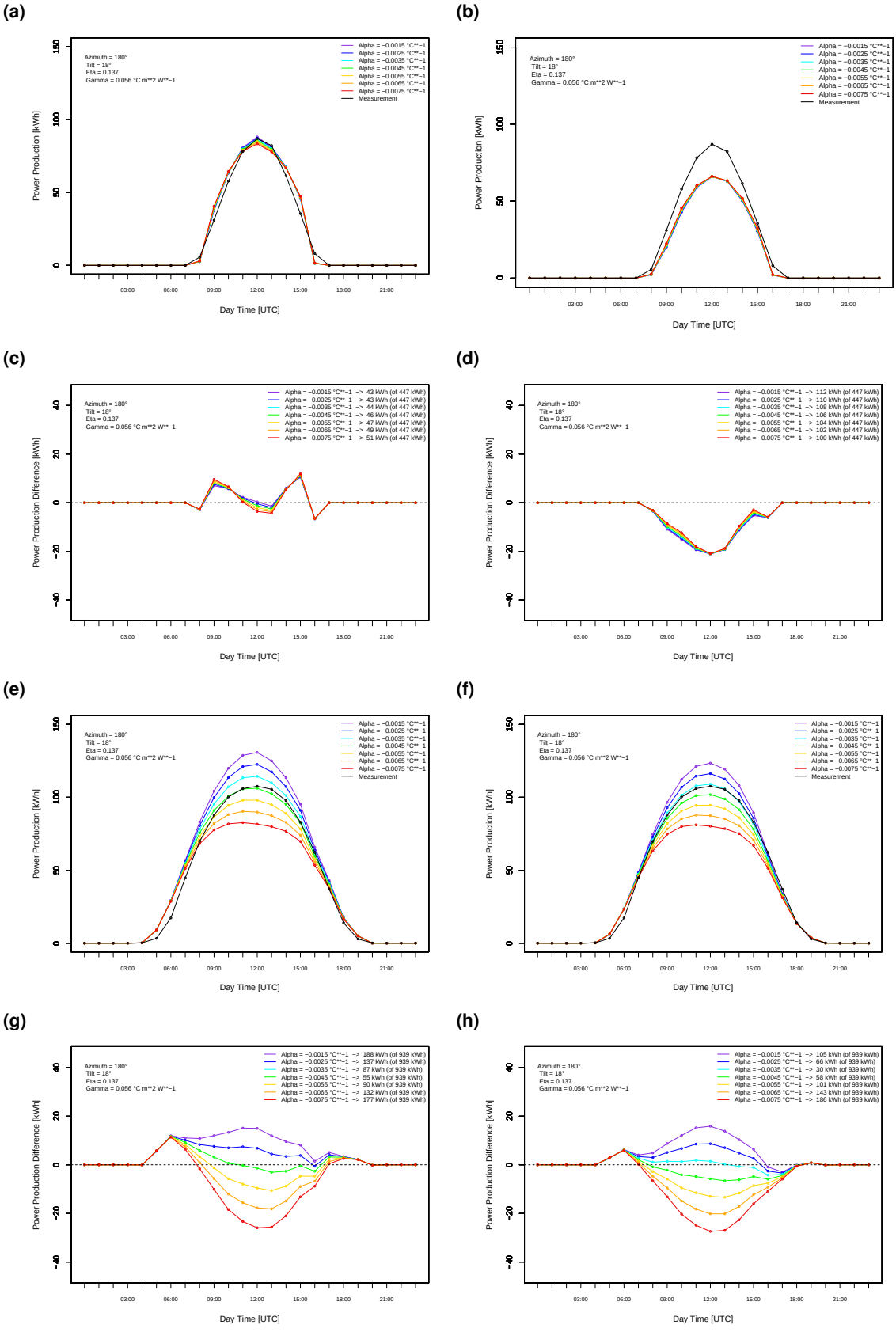


Figure 24: Temperature coefficient sensitivity $\rho_\alpha = \frac{\delta E}{\delta \alpha}$. Explanations are in the caption of Figure 21.

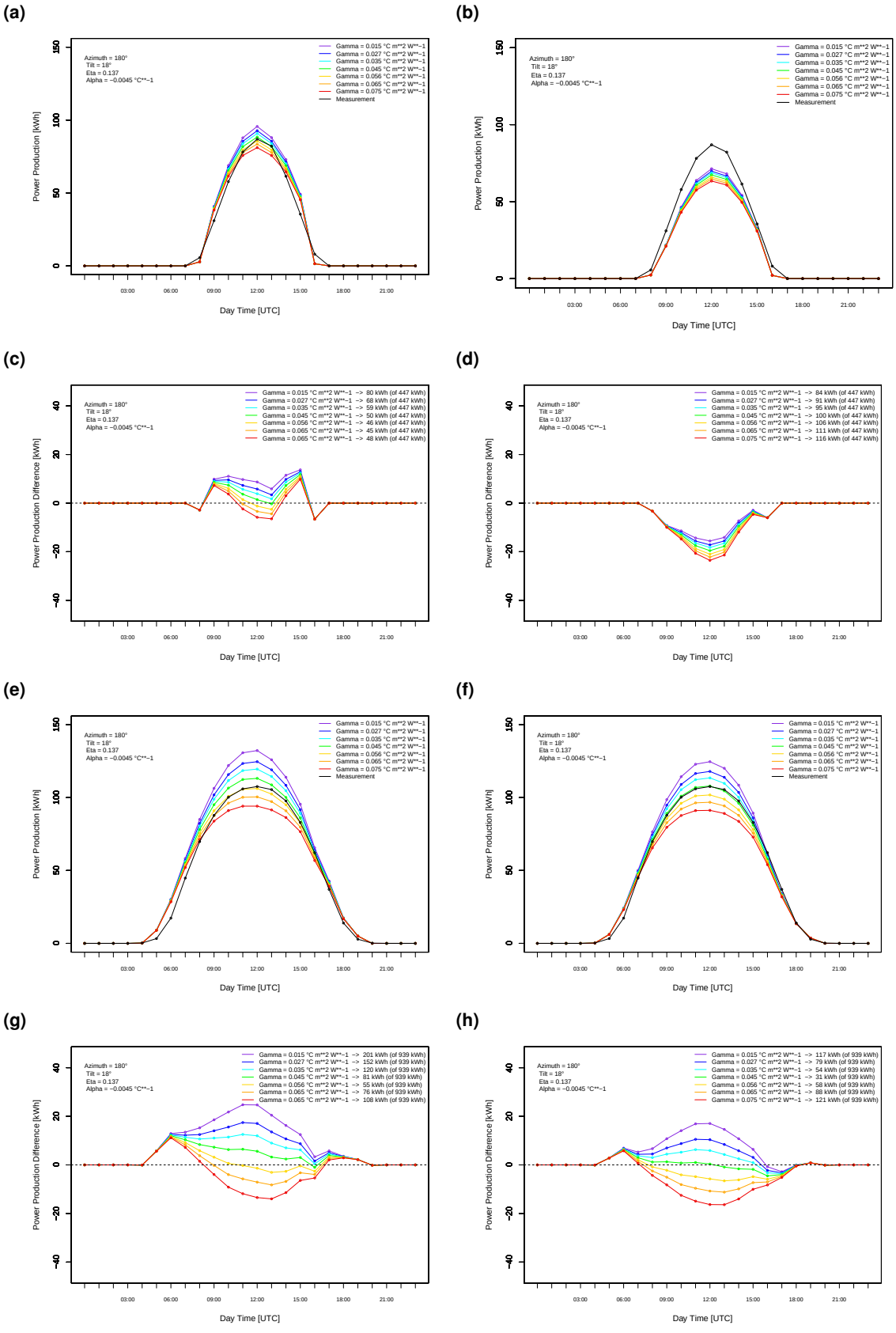


Figure 25: Heating coefficient sensitivity $\rho_\gamma = \frac{\delta E}{\delta \gamma}$. Explanations are in the caption of Figure 21.

summer forecast when using $\gamma = 0.056\,°\mathrm{C}\,\mathrm{m}^2\,\mathrm{W}^{-1}$ (yields a daily error of 55 kWh).

Beside the sensitivity, the results in Figure 25 also show the weakness of the (global) assumptions about γ made in the literature: *Lorenz et al.* (2010) suggest to use $\gamma = 0.027\,°\mathrm{C}\,\mathrm{m}^2\,\mathrm{W}^{-1}$ if a PV panel is free-standing and $\gamma = 0.056\,°\mathrm{C}\,\mathrm{m}^2\,\mathrm{W}^{-1}$ if it is roof-integrated. The case study PV plant consisting of free-standing PV modules, however, does not correspond to this assumption. It rather seems that even though the panels are free-standing, they heat up stronger and thus reveal higher module temperatures at typical summer conditions. That is why the power production forecast model needs a higher γ than $0.027\,°\mathrm{C}\,\mathrm{m}^2\,\mathrm{W}^{-1}$ in order to catch all the influence factors heating up the modules and reducing their efficiency.

Therefore, the heating coefficient γ is suggested to be the largest weakness of the used PV efficiency model by *Beyer et al.* (2004). This could be due to the fact that the heating of a PV module cannot be explained by one single scaling factor only, since it depends on too many factors as the following (some of them are described in literature, some are speculative):

- Stronger wind $\rightarrow$ better ventilation of the PV panel (especially if free-standing) $\rightarrow$ stronger cooling of the PV modules
- Lower albedo of the roof below PV panel $\rightarrow$ stronger energy absorption from solar irradiance $\rightarrow$ stronger thermal radiation from the roof up to the PV panel $\rightarrow$ stronger heating of the PV modules
- Higher tilt angle of the PV panel $\rightarrow$ better ventilation of the PV panel (if free-standing) especially with windy conditions $\rightarrow$ stronger cooling of the PV modules

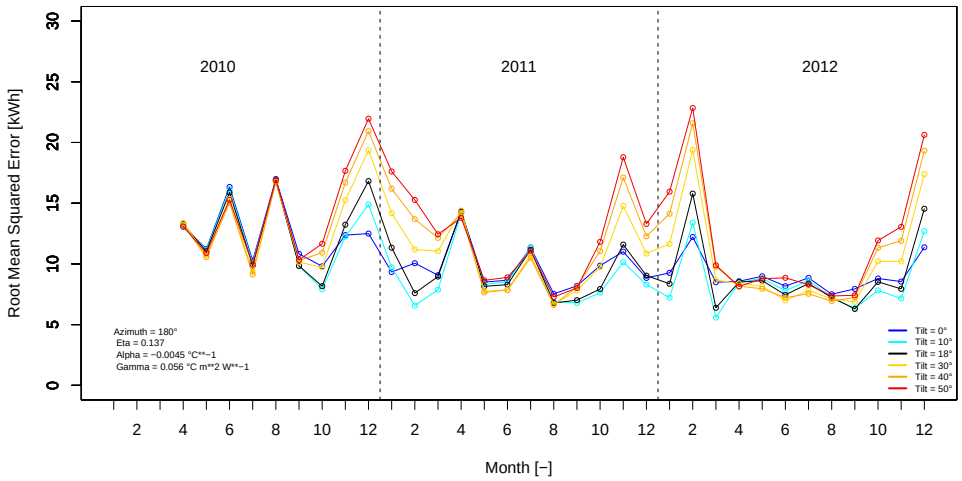
4.2.2 Sensitivity from a monthly perspective

In the following, the same sensitivity tests are applied for the same case-study PV plant. However, they are computed and verified for the whole three-years period from 2010 to 2012 and not on a daily basis anymore. The figures show the errors RMSE, ME, and MAE per 1h-time-step, which are computed for each month. Compared to the last section, only the power production forecasts based on satellite (irradiance) and surface station (temperature) measurements are shown. It is important to mention that the colors are the same for each parameter value as in the previous section with one exception: for each sensitivity study, the error curve based on the best-estimate parameters (this is 18° for the tilt angle, 180° for the azimuth angle, $-0.0045\,°\mathrm{C}^{-1}$ for the temperature coefficient, and $0.056\,°\mathrm{C}\,\mathrm{m}^2\,\mathrm{W}^{-1}$ for the heating coefficient) is colored in black. Hence, the black curve does not represent the power production measurement anymore (which would be the zero-line in the following figures).

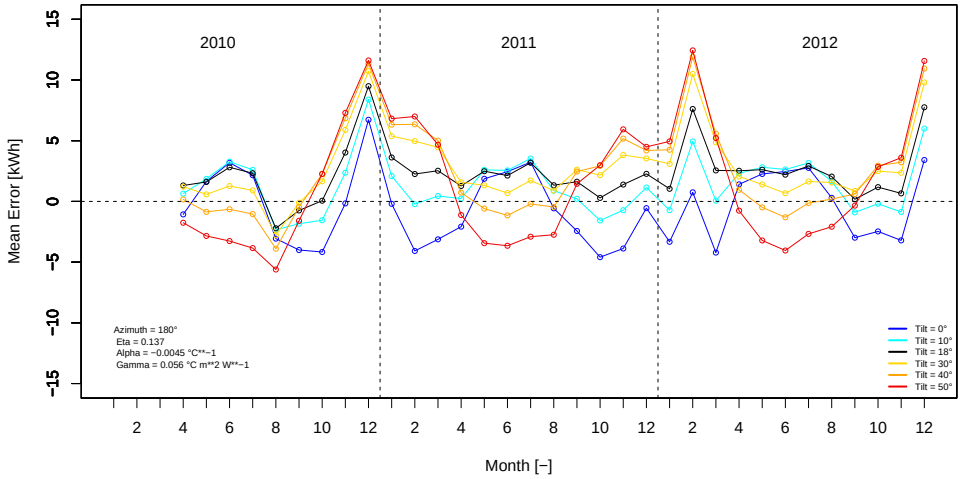
Tilt angle (β) sensitivity: $\rho_\beta = \frac{\delta E}{\delta \beta}$

Figures 26a, 26b, and 26c show the monthly errors of the different tilt angle assumptions. The RMSE development shows the high sensitivity to the tilt angle in winter compared to summer, what can be recognized in the diurnal verification already. The ME in Figure 26b further reflects the effect of the seasonal characteristics of the solar zenith angle described in the previous section: in winter, the assumption of a tilt angle of 50° overestimates the power production most strongly, whereas in summer the order reverses and the same tilt angle leads to the strongest underestimation. Finally, the errors

(a) Root mean squared error



(b) Mean error



(c) Mean absolute error

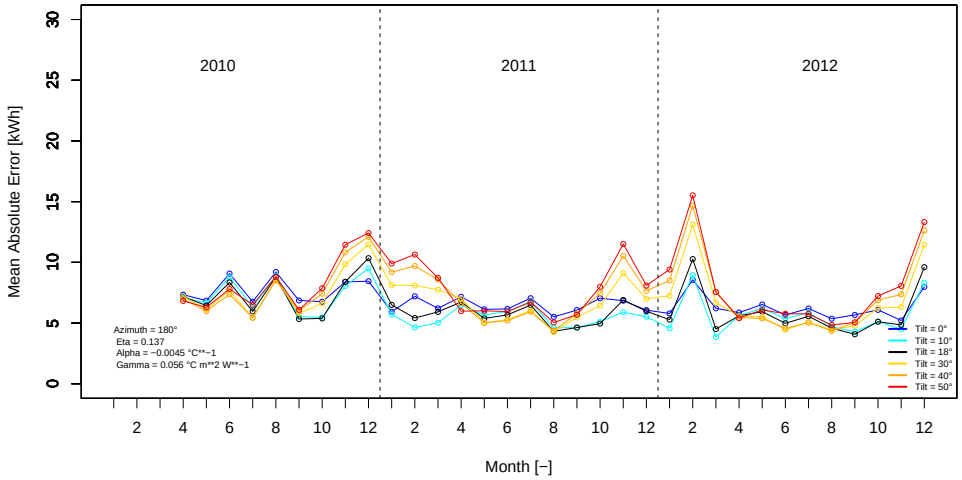


Figure 26: RMSE, ME, and MAE per 1h-time-step for each month (in kWh) for the case-study PV plant, assuming different tilt angles β. The power production forecast is based on satellite (irradiance) and surface station (temperature) measurements only (and not on COSMO forecasts). The colors are the same for each parameter value as in the previous section (with the diurnal sensitivity) with one exception: the error curve based on the best-estimate parameters (this is 18° for the tilt angle, 180° for the azimuth angle, $-0.0045\,^{\circ}\mathrm{C}^{-1}$ for the temperature coefficient, and $0.056\,^{\circ}\mathrm{C\,m^2\,W^{-1}}$ for the heating coefficient) is colored in black. Hence, the black curve does not represent the power production measurement anymore (which would be the zero-line).

indicate that using the known tilt angle of 18° for the power production forecast of this PV plant does not yield the best performance: in winter, a lower tilt angle would lead to significantly better results and in summer, a higher tilt angle would slightly improve the forecast. This shows that even though the applied power production forecast model seems to perform well on clear-sky days when using measured irradiance and temperature data as well as all the best-possible set of PV plant metadata parameters (according to the diurnal verifications in the previous section), significant uncertainties come up when calculating the forecast over a whole year including all non-perfect weather conditions.

Azimuth angle (ψ) sensitivity: $\rho_\psi = \frac{\delta E}{\delta \psi}$

In summer, an increasing deviation from the known azimuth angle of 180° also increases the power production forecast error, as it can be seen in Figures 27a, 27b, and 27c. In certain winter months, however, the forecasts based on the two extreme azimuth angles (90° and 270°) abruptly switch to the ones with the best performance. This finding shows the compensation effect mentioned in the previous section: totally wrong assumptions within the power production forecast model can improve the forecast performance in an artificial way by increasing or reducing the irradiance on a PV panel such that any other error within the forecast gets compensated.

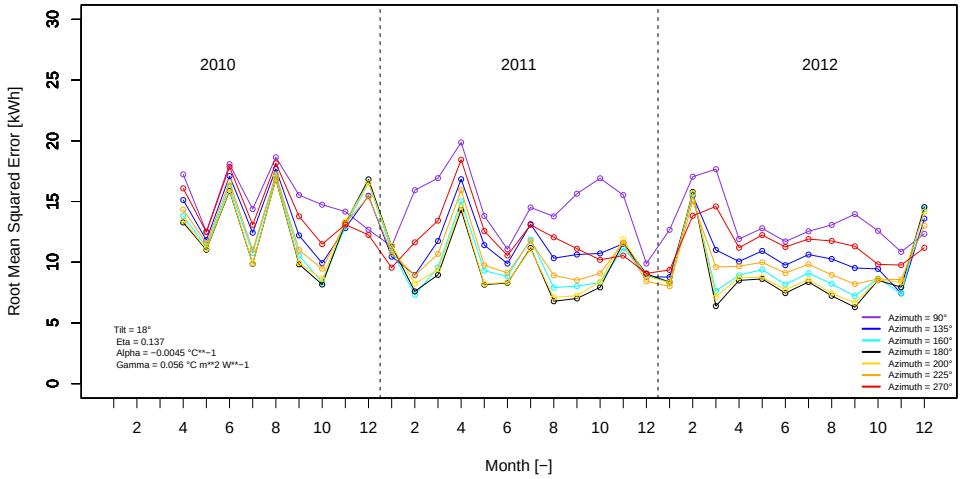
Temperature coefficient (α) sensitivity: $\rho_\alpha = \frac{\delta E}{\delta \alpha}$

For the temperature coefficient sensitivity test, the errors in Figures 28a, 28b, and 28c again reflect the seasonal sensitivity difference: changing α in summer significantly changes the power production forecast error, whereas a change in winter almost has no effect. What can be concluded, however, is that using the value of $-0.0045\,°C^{-1}$ as the known α parameter for this PV plant indeed yields the best power production forecast performance over the whole period.

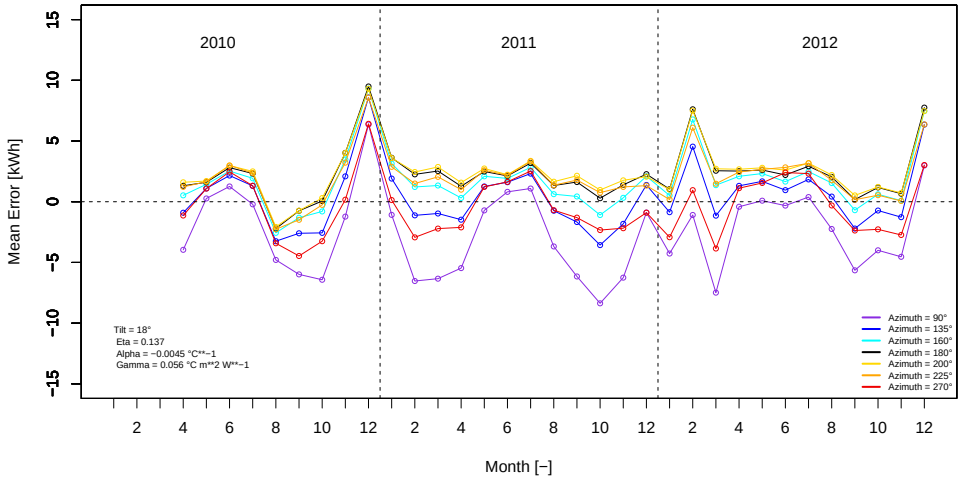
Heating coefficient (γ) sensitivity: $\rho_\gamma = \frac{\delta E}{\delta \gamma}$

Figures 29a, 29b, and 29c show the monthly errors of the different heating coefficient assumptions. As it is observed in the error figures for the temperature coefficient sensitivity test already, the sensitivity of the power production forecast to a change in the heating coefficient is quite high in summer but low in winter. Furthermore, Figure 29a indicates that a γ of $0.056\,°C\,m^2\,W^{-1}$ or slightly higher yields the best results for the analyzed PV plant.

(a) Root mean squared error



(b) Mean error



(c) Mean absolute error

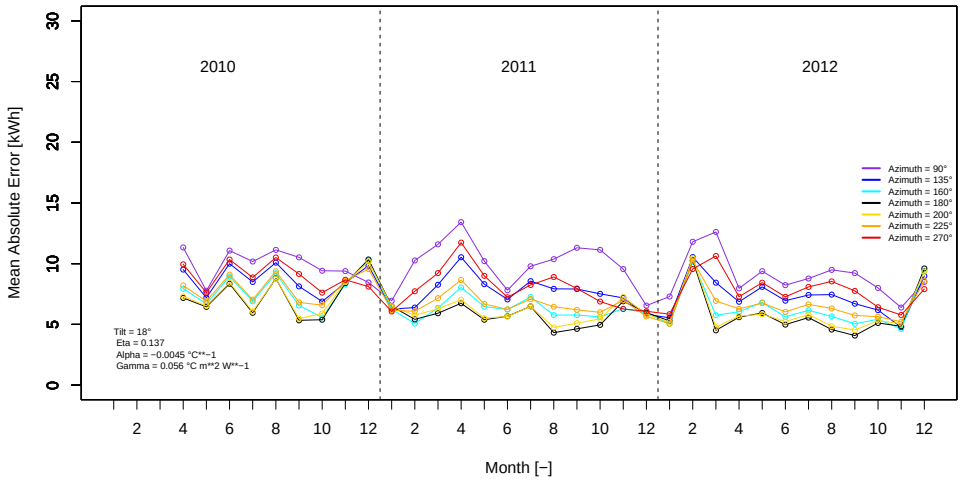
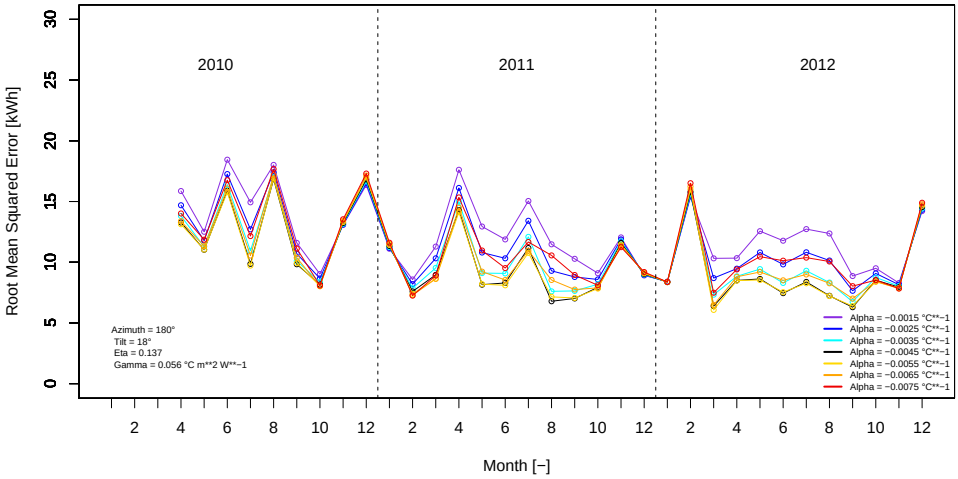
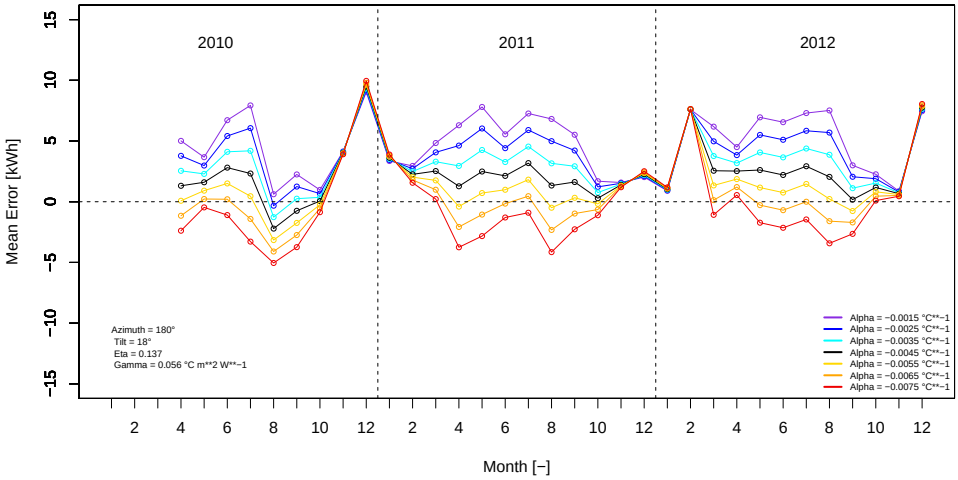


Figure 27: RMSE, ME, and MAE per 1h-time-step for each month (in kWh) for the case-study PV plant, assuming different azimuth angles ψ. Explanations are in the caption of Figure 26.

(a) Root mean squared error



(b) Mean error



(c) Mean absolute error

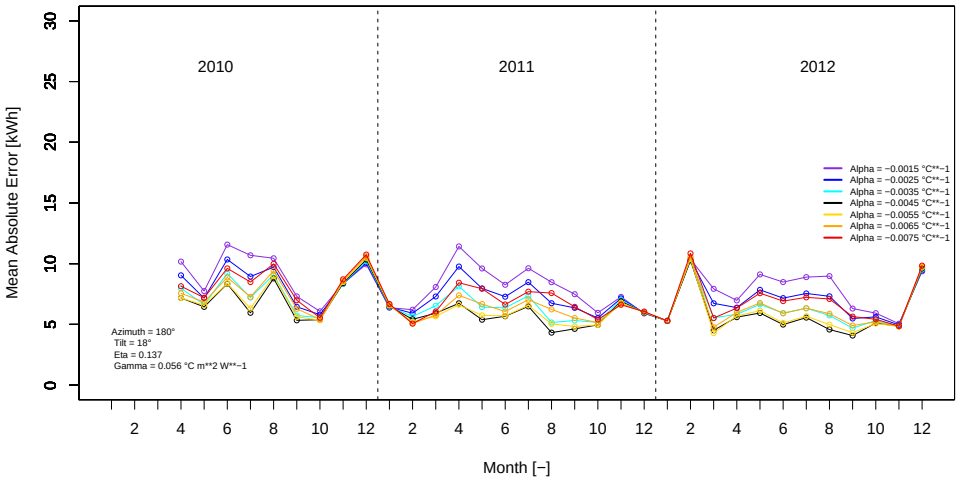
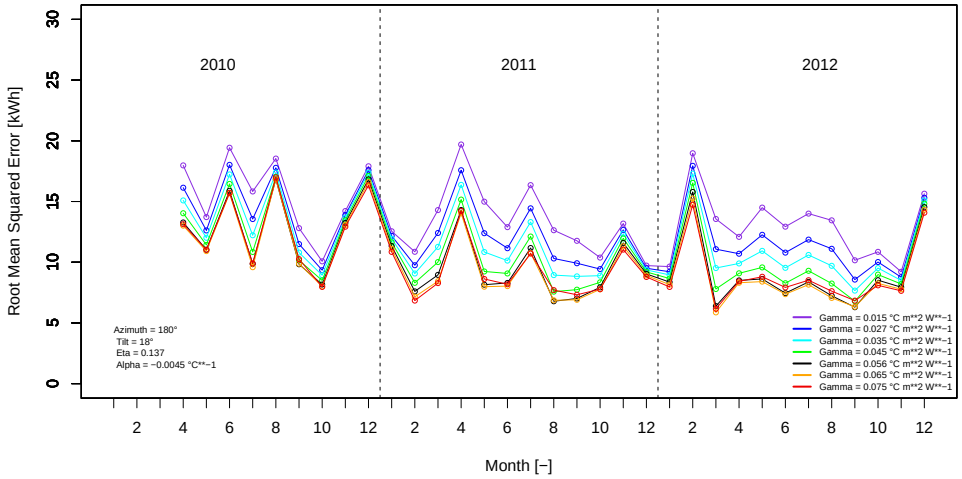
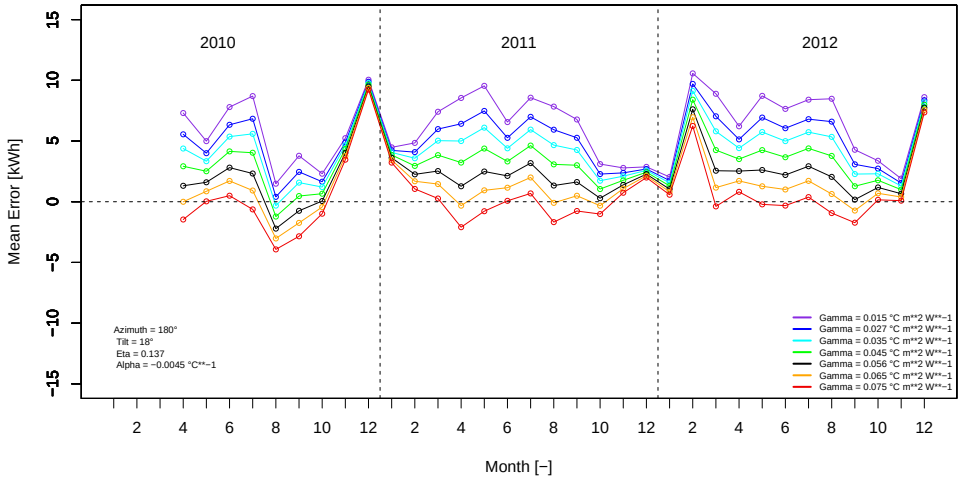


Figure 28: RMSE, ME, and MAE per 1h-time-step for each month (in kWh) for the case-study PV plant, assuming different temperature coefficients α. Explanations are in the caption of Figure 26.

(a) Root mean squared error



(b) Mean error



(c) Mean absolute error

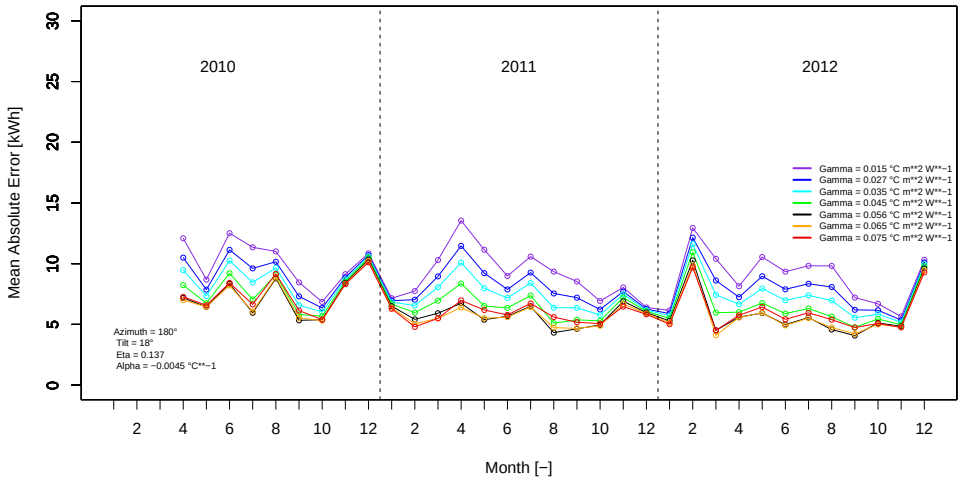


Figure 29: RMSE, ME, and MAE per 1h-time-step for each month (in kWh) for the case-study PV plant, assuming different heating coefficients γ. Explanations are in the caption of Figure 26.

4.2.3 Summary of the sensitivities

Table 16 summarizes the sensitivities for the four analyzed parameters from a seasonal perspective, based on the findings of the previous two sections. The number of stars represents the magnitude of the difference of the daily aggregated power production forecast error based on the worst and and the one based on the best parameter value (as shown by the numbers in the legends of the previous figures displaying the diurnal sensitivities to the different parameters). This means, the larger this difference is, the higher the sensitivity of the according parameter over the analyzed value range is. One star represents a difference range between 0 and 50 kWh, two stars one between 50 and 100 kWh, and three stars one between 100 and 150 kWh.

Table 16: Summary of the seasonal parameter sensitivities.

Parameter	Winter sensitivity	Summer sensitivity
Tilt angle β	***	**
Azimuth angle ψ	***	***
Temperature coefficient α	*	***
Heating coefficient γ	*	***

4.3 Power production forecast for all PV plants

The forecast for a large set of PV plants is dominated by technical uncertainties: even with the use of the actual (non-forecasted) satellite and surface measurement data the power production forecast errors remain large. Using same technical coefficients for all PV plants is necessary due to the lack of information about plant-specific technical details. Snow covering PV panels additionally enhances power production forecast errors in winter if it is not accounted for in the model. A spatial and temporal aggregation of the power production forecast values preceding the calculation of the statistical parameters (as the RMSE) improves the performance of the power production forecast model, as a consequence of the smoothing of the small-scale variabilities.

* * *

This last section contains the results of the power production forecast for all available PV plants, which would be the ultimate goal of a balance group as ewz. As described in the methods section, the following approaches (sets of assumptions) are applied to compute the PV power production forecast:

- HORIZONTAL
- TILTED
- FITTING
- FITTING-REF

The figures in this section show the different errors for each of the upper PV power production forecast methods. The colors of the error curves are blue for the HORIZONTAL, red for the TILTED, yellow for the FITTING, and black for the FITTING-REF method. The HORIZONTAL method is split into the three following methods, which are all based on slightly different assumptions:

- HORIZONTAL (Comp, Eta Tdep Idep): this is the HORIZONTAL method as it is described in the methods section.
- HORIZONTAL (Glob, Eta Tdep Idep): instead of using the sum of the two COSMO irradiance components on a horizontal plane (ASWDIR_S and ASWDIFD_S), the global irradiance parameter GLOB is used as irradiance input.
- HORIZONTAL (Glob, Eta Tdep): additionally to the use of GLOB instead of ASWDIR_S and ASWDIFD_S, the used PV module efficiency model is simplified by assuming the module efficiency to be independent on global irradiance (which means, the logarithmic dependency on global irradiance is not accounted for).

Since the differences between the three different HORIZONTAL methods are negligible compared to the differences between all methods (HORIZONTAL, TILTED, FITTING, and FITTING-REF), they are not accounted for in this section even though shown in the error figures. This means, only the curve colored in cyan, which is the one of the method HORIZONTAL (Comp, Eta Tdep Idep), should be considered in the following figures.

It is important to mention again the total number of PV plants changing over the three years, as described in the data section. Hence, intercomparisons in absolute terms between the three years have to be interpreted with caution.

4.3.1 Verification for a three-years period

Verification per PV plant and 1h-time-step (no aggregation)

Figure 30 shows the measured Swiss-wide mean hourly power production per PV plant for every month of the three years. It ranges from 1 to 15 kWh, whereby the power production in the summer months is two to three times higher than in the winter months. What Figure 30 also indicates is that the PV plants being added to the total set every year do not seem to change the per-plant power production significantly. Figure 30 is shown in order to understand the magnitude of the different errors discussed in the following.

For every month, Figures 31a, 31b, 31c, and 31d show the power production forecast errors on the non-aggregated PV plant and time step level. This means, they indicate the errors inhering to the power production forecast of a single PV plant at every 1h-time-step of the month. The results of four methods are displayed: the blue curve represents the HORIZONTAL, the red one the TILTED, and the yellow one the FITTING method. In black, again the FITTING method is shown but based on satellite (for irradiance) and surface (for air temperature) measurements instead of COSMO forecasts. Hence, the black curve represents the reference forecast (FITTING-REF). The error types shown are the RMSE, the ME, the MAE, and the relative RMSE. In the following, the term “error” generally stands for all of the three absolute errors (RMSE, ME, MAE). Interpretations of specific error types are mentioned accordingly.

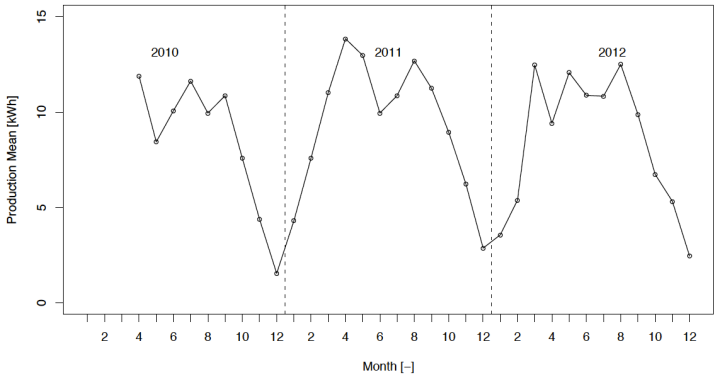


Figure 30: Monthly power production mean (in kWh) per PV plant and 1h-time-step.

The RMSE of all methods varies roughly between 4 and 12 kWh, the ME between -2 and 6 kWh, and the MAE between 2 and 7 kWh. There is a high inter-annual variability in the sense that a certain month can reveal significantly different error values over the three years. This overall variability pattern indicates the high dependency of the PV power production and thus its forecasting performance on the weather conditions of a certain month itself, which can be seen in the high variability of the mean power production in Figure 30 as well.

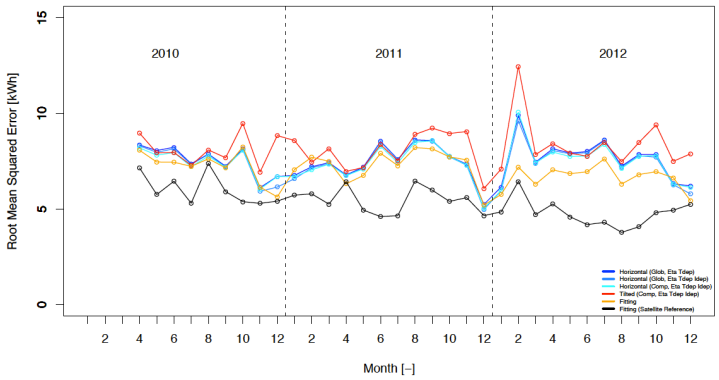
Qualitatively, the three methods based on the COSMO forecast show a quite similar error development over the three years, especially in summer, even though they are based on different assumptions. This indicates the COSMO forecast and its performance, respectively, to have a main influence on the power production forecast performance because it serves as the (same) basis for all three methods.

The ME in Figure 31b indicates a positive bias and thus a general overestimation of the power production by the HORIZONTAL and TILTED methods. With a few exceptions, the positive bias tends to be larger in summer than in winter. With the FITTING method, the positive bias seems to get eliminated, since its ME is distributed around zero over the three years. A similar pattern inheres to the reference forecast.

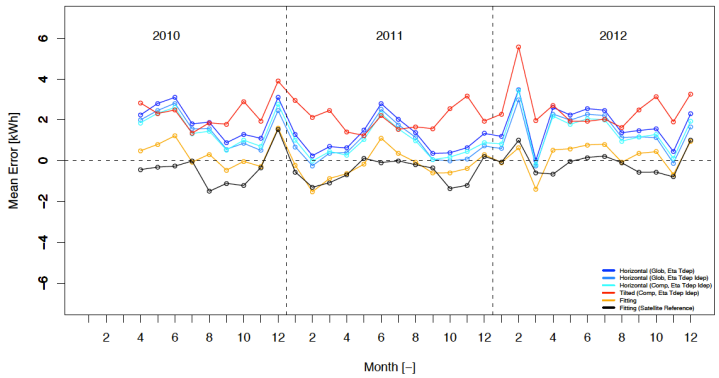
Comparing winter and summer months, the errors do not show a clear seasonality pattern on the first glance. However, considering the much larger amounts of energy from incoming shortwave global radiation and thus the power generation during summer, shown in Figure 30, the winter errors having the same magnitude as the summer errors means a significantly worse performance of the power production forecast model for winter. This is indicated in Figure 31d showing the relative RMSE, which is calculated by dividing the absolute RMSE of Figure 31a by the mean power production of Figure 30. The reasons for the seasonal performance difference could be mainly the following:

1. As it is summarized in Table 16, wrong assumptions about tilt and azimuth angles in winter can generate a significant error in the power production forecast. Both for the HORIZONTAL and TILTED methods, the same assumptions for tilt and azimuth angles are made for all PV plants. For both methods, there is certainly a number of PV plants fulfilling these assumptions, but on the other side, there are many PV plants having different tilt and azimuth angles than prescribed by the two methods. These PV plants are the ones raising the average power production forecast error in winter. Even within the FITTING method using individual tilt and azimuth angles for each PV plant, there are still deviations from the real angles of a few degrees for most of the PV plants, which is enough to generate significant power production forecast errors in winter.

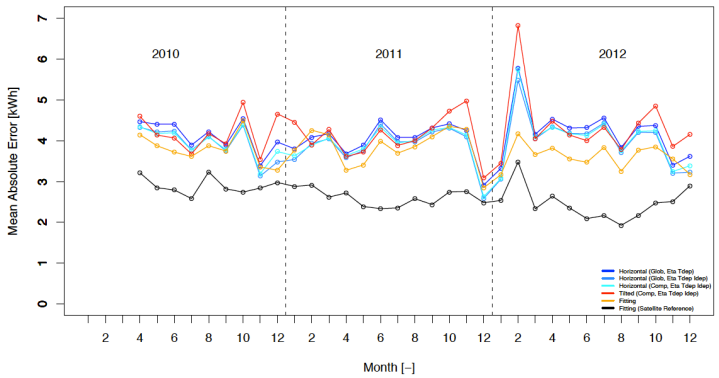
(a) Root mean squared error



(b) Mean error



(c) Mean absolute error



(d) Relative root mean squared error

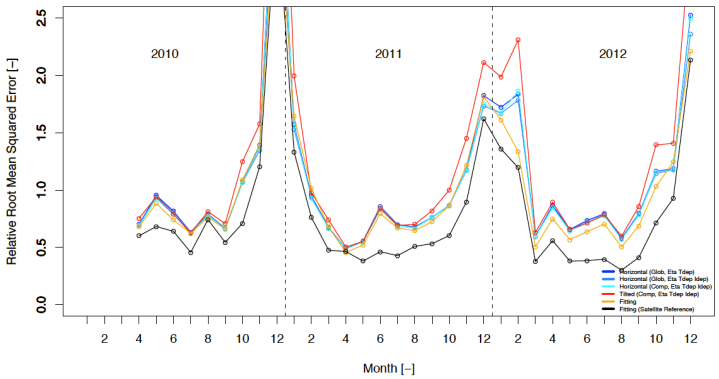


Figure 31: RMSE, ME, MAE (in kWh), and relative RMSE (in %) per PV plant and 1h-time-step for each month, based on the four methods explained in the text.

2. Snow covering the PV plants prevents them from generating power, which can lead to measured power production values of certain PV plants towards 0 kWh. Since snow cover is not accounted for in the power production forecast model, the latter generates a normal power production output (based on the available global irradiance), which leads to a high error when comparing with the low measured power production values.

The two winter months revealing the largest relative errors (according to Figure 31d) are December 2010 and February 2012. To understand their magnitude, the performance of COSMO during the two months is analyzed. As it can be seen in Figures 16a, 16b, and 16c, COSMO indeed performs a bit worse in December 2010 compared to the other Decembers but (beside November) still best compared to the other months of 2010. In February 2012, COSMO models irradiance even slightly better than in February 2011. And also from a seasonal perspective, February 2012 does not stick out in terms of COSMO performance. This proves that the COSMO performance cannot be (mainly) responsible for the large power production forecast errors in these two months, being partly larger than many errors in summer. The climate bulletins of MeteoSwiss for the two months, however, might provide an explanation for this: according to *MeteoSwiss* (2010b), December 2010 was a month abnormally rich of snow in the Swiss Plateau. With 57 cm, Berne registered a record amount of new snow for December. Also Zurich got a monthly amount of new snow of 55 cm, which was close to a record. Considering the fact that a big share of all PV plants is located in the region of Zurich, probably many PV panels were covered with a thick snow layer over many days of the month, preventing them from producing any power. This is quite likely to be the main reason for the low power production values (shown in Figure 30) and thus the (relatively) large error of the power production forecast. The reason for the large error in February 2012 might be a combination of temperature, snow, and sunshine duration (*MeteoSwiss*, 2012): the month was characterized by a heavy cold wave in the first half of the month, which made it to the ten coldest Februaries since measurement start (see Figure 32). Smaller lakes of the Swiss Plateau froze in the middle of the month. Especially northeastern Switzerland registered extremely cold temperatures. Additionally, Februrary 2012 was very dry and got a lot of sunshine (see Figure 32). Even though there were only a few small snow fall events during the cold wave, the cold temperatures helped to maintain a permanent snow layer also in the Swiss Plateau during the whole first half of the month, which was, however, only of small thickness. Connecting these conditions to the power production forecast error reveals the following explanation: the thin but permanent snow layer on many PV panels reduced (but not necessarily stopped) their power production, which is not accounted for in the power production forecast model. However, COSMO probably captures the large amounts of sunshine, which leads to an abnormally high output of the power production forecast model and thus a high error. In conclusion, for the power production forecast error of February 2012 the abnormally high global irradiance probably dominates over the slightly reduced power production due to snow cover, whereas in December 2010 the strongly reduced power production due to snow cover might be mainly responsible for the error.

After having interpreted all power production forecast methods as a whole, the different methods in Figures 31a, 31b, and 31c are compared among each other. In summer, the HORIZONTAL and TILTED methods perform similarly. This coincides with the results from the case study revealing only a moderate sensitivity of the tilte angle to the forecasted power production in summer. However, significant differences between the HORIZONTAL and TILTED methods occur in winter, where the latter performs worse. On first glance, this is surprising since the assumption of all PV plants being 21° tilted and southward oriented (rule of thumb used by ewz) seems to be more realistic than the one of all PV

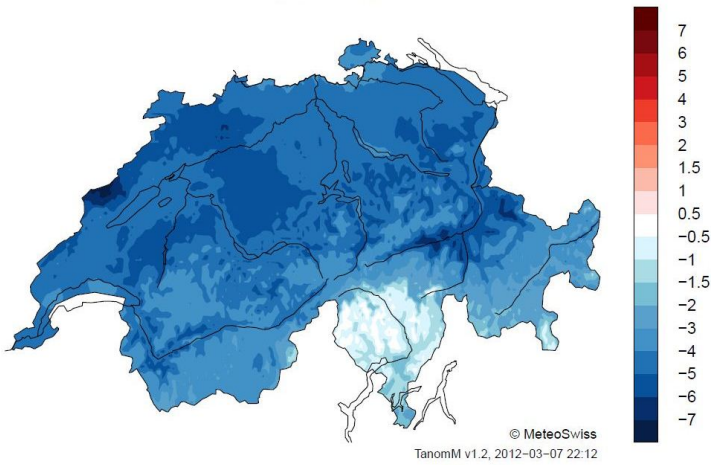


Figure 32: Februar 2012 temperature deviation in °C from the norm, with the reference period of 1961 - 1990. *MeteoSwiss* (2012).

plants being horizontally installed. However, the combination of two previously mentioned effects might be responsible for the error differences:

1. For all PV plants being tilted less than 21°, the assumptions made in the TILTED method lead to an overestimation of the forecast model by improving the conditions for power production considering the low solar elevation angle in winter. In other words, this is the representation of the high sensitivity of the tilt angle to the power production forecast in winter as found in the case study.

2. For all PV plants, no matter what tilt angle they have, the aforementioned problem of the winter error rise due to snow cover is more pronounced when using the TILTED instead of the HORIZONTAL method. The following hypothetical example should explain this: a certain PV plant (with any tilt angle) has a measured power production value of 0 kWh due to a thick snow layer covering it. Because of the lack of snow cover in the power production forecast model, the error of the forecasted power production output for this plant amounts to 10 kWh using the HORIZONTAL method. The use of the TILTED method, however, generates an even higher error of 15 kWh, which results from the missing snow cover parameterization (as for the HORIZONTAL method) but additionally also from the improved conditions for power production with regard to the low solar elevation angle in winter due to the prescribed tilt angle of 21° instead of 0° in the HORIZONTAL method.

The FITTING method as the third one reveals the best overall performance. With up to 3 kWh, the improvement of the RMSE is highest in 2012, whereas in 2010 and 2011 it performs similarly to the HORIZONTAL method (Figure 31a). The reasons for the larger improvement in 2012 might be the following: first of all, the additional PV plants in 2012 compared to the previous years are likely to increase the uncertainties of the assumptions of tilt and azimuth angles, which are made for the HORIZONTAL and TILTED but avoided by the FITTING method. Hence, the difference between the performances of the methods increases. Second, the fact that the PV plant parameters are fitted based on data of 2012 only but used for the power production forecast of every year might slightly improve the forecast based on the FITTING method in 2012 relative to 2010 and 2011. In summer, the improvements of the FITTING method are generally larger than in winter. This could be due to the fact that the FITTING

method also uses plant-specific module efficiency coefficients (instead of overall-assumptions made in the HORIZONTAL and TILTED methods), to which the power production forecast is very sensitive in summer according to the findings of the case study. On the other hand, there are single months as for instance November 2012 or February 2011, where the power production forecast based on the FITTING method does not reveal the best results. This shows the fact that overall assumptions (as made for the HORIZONTAL and TILTED methods) might even be better for certain monthly weather conditions than the PV-plant-specific assumptions derived through parameter fitting.

The black curve in Figures 31a, 31b, 31c, and 31d represents the power production reference forecast. It has to be compared mainly to the yellow curve of the FITTING method, since it is based on exactly the same PV plant metadata (tilt angle, azimuth angle, module efficiency) as the FITTING method but computed by using measurements instead of COSMO forecasts. The reference method performs significantly better in almost all months. Nevertheless are the errors still quite high considering the fact that the uncertainties of irradiance and temperature forecasts are eliminated (mainly but not fully, since there are still errors both from the satellite uncertainty itself and the interpolation or extrapolation of the gridded satellite data set and the temperature measurements from the surface stations to the PV plant locations). This result indicates the impact of technical uncertainties as orientation and module efficiency coefficients of the PV plants to dominate over the meteorological COSMO forecast uncertainties. Considering the average errors of the black and the yellow curves in Figures 31a, 31b, 31c, the technical uncertainties make up at least two thirds and the meteorological uncertainties the residual third. The meteorological uncertainties can be analyzed by subtracting the RMSE of the reference method from the one of the FITTING method. The higher this difference is, the worse COSMO performs in the according month. This can approximately be recognized when analyzing each month of the year having the largest RMSE difference: October for 2010, June for 2011, and July for 2012. In Figures 16a, 16b, and 16c, showing the RMSE of the COSMO global irradiance forecasts, the same three months reveal relatively high RMSE values in the Zurich region (and mostly also in the Grisons) where most of the PV plants are. Therefore, the COSMO performance in the region of Zurich seems to be quite crucial for the performance of a power production forecast for the used PV plant set. More difficult to interpret are the technical uncertainties constituting to the errors of the reference forecast itself. The following aspects might be responsible for it:

- Measurement uncertainties of the satellite
- Extrapolation / interpolation uncertainties from the gridded satellite data and the surface measurement stations (both spatially and temporally)
- Deviations of the fitted PV-plant-specific from the real metadata parameters
- Snow cover on PV panels leading to low power production measurement and thus high forecast errors
- Errors in power production measurement data
- Uncertainties / incompleteness of used module efficiency model
- Uncertainties in the conversion of solar irradiance on tilted and oriented PV panel

Verification for all PV plants together and per time step (spatial aggregation)

Figure 33 shows the measured Swiss-wide mean hourly power production of all PV plants together for

every month of the three years. It ranges from 0.4 to 3.5 MWh. Qualitatively it is the same as the per-PV-plant power production shown in Figure 30, which is due to mathematical reasons. It is important to mention that the scale used in this paragraph is MWh and not kWh anymore.

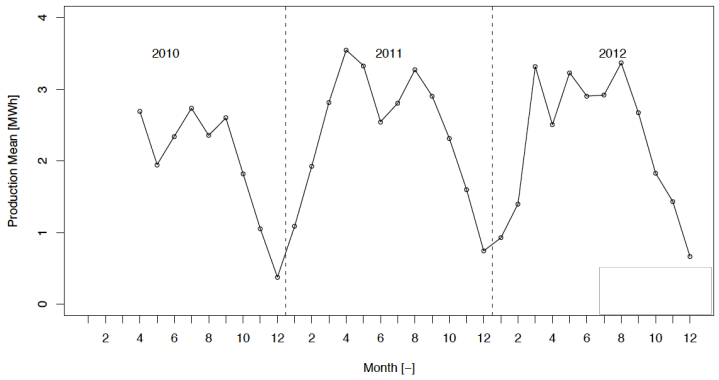


Figure 33: Monthly power production mean for all PV plants together and per 1h-time-step.

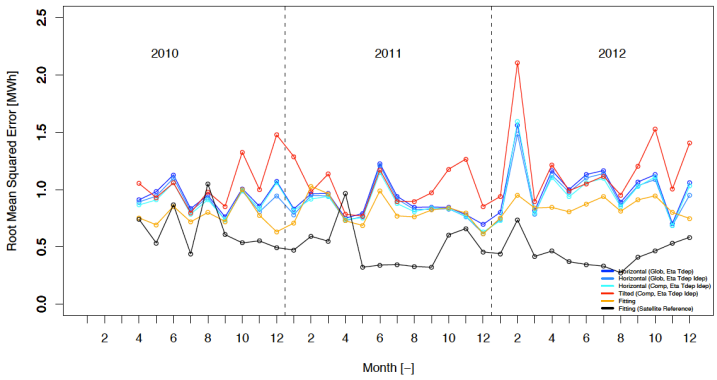
For every month, Figures 34a, 34b, 34c, and 34d show the power production forecast errors on the spatially aggregated PV plant and time step level. This means, they indicate the errors inhering to the power production forecast of all PV plants together at every 1h-time-step of the month. The curves represent the same methods as in the previous paragraphs.

The RMSE of all methods varies between 0.3 and 2.1 MWh, the ME between -0.5 and 1.5 MWh, and the MAE between 0.3 and 1.5 MWh. In general, the qualitative picture of the error looks very similar to the one of the previously described, non-aggregated verification. The relative RMSE, shown in Figure 34d, goes down to around 25 % and thus is significantly lower than the non-aggregated one in Figure 31d. This reduction is caused by the smoothing of small-scale variability due to the spatial aggregation.

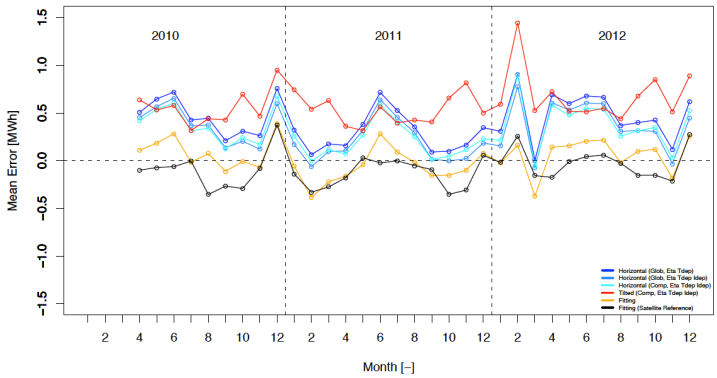
There are two months, August 2010 and April 2011, in which all power production forecast methods perform better than the reference forecast when considering the aggregated RMSE in Figure 34a. The reason for this result might be compensation effects as described in the case study section: for instance, the use of certain module efficiency parameters (as α and γ) might be wrong for certain weather conditions, which would lead to an overestimation of forecasted power production. Using measured irradiance or temperature values as input, the power production forecast would yield values too high. Assuming, COSMO would underestimate irradiance or temperature in this situation, the power production overestimation due to the module efficiency coefficients would be compensated by the underestimation of COSMO. Hence, using COSMO forecasts instead of measurements could end up in even lower errors. This example demonstrates that the performance of the used irradiance and temperature forecasts is not necessarily the limiting factor for the performance of the power production forecast, but rather the model and its set of PV plant metadata parameters used for the power production forecast.

The Swiss-wide aggregated MAE per time step of Figure 34c is specifically interesting for electricity supply companies as ewz. This is because the latter are mainly interested in the Swiss-widely summed absolute (meaning both negative and positive) deviations of their forecasts from the measured power production, since this is the electricity amount that has to be balanced somehow. Assuming a certain constant balancing energy price, which would have to be paid per MWh power production forecast error, Figure 34c indicates that the mean monthly payment per 1h-time-step would vary by up to 50 %.

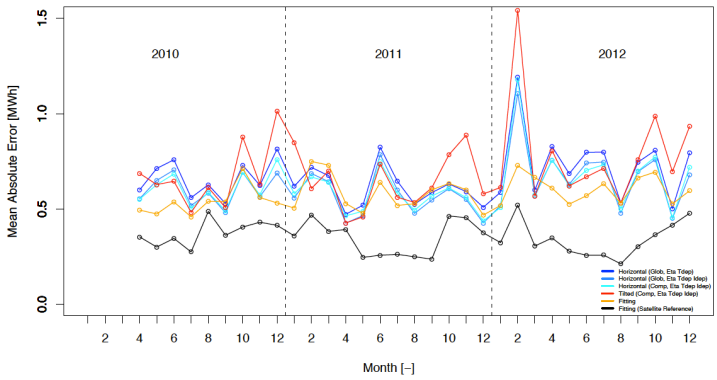
(a) Root mean squared error



(b) Mean error



(c) Mean absolute error



(d) Relative root mean squared error

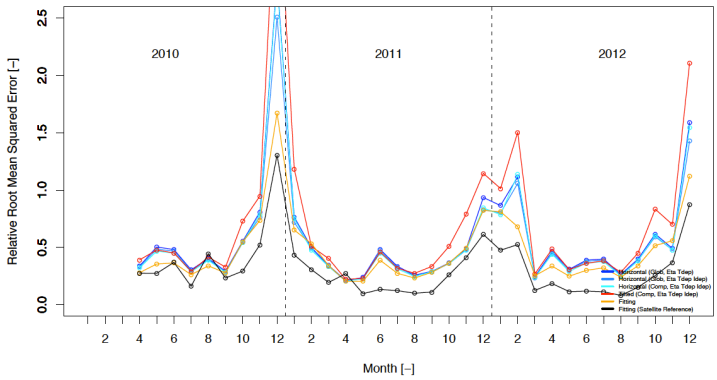


Figure 34: RMSE, ME, MAE (in MWh), and relative RMSE (in %) for all PV plants together and per 1h-time-step for each month, based on the four methods explained in the text.

Verification for all PV plants together and per month (spatial and temporal aggregation)

Figure 35 shows the absolute error of the Swiss-wide and monthly aggregated power production forecast. It indicates the magnitude of deviations from the measured power production of a whole month the different power production forecast methods would yield. Almost all methods reveal the largest monthly errors in the summer month. One exception is the TILTED method, revealing a constantly high positive bias, independent of the season. What is also confirmed by Figure 35 is that the FITTING method performs best over the three years, since its monthly error is spread around zero over the three years.

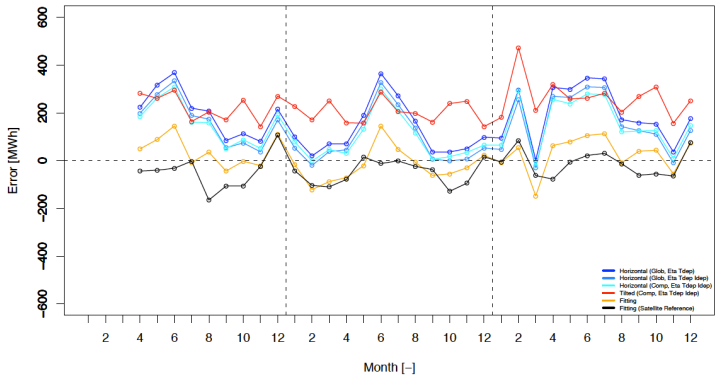


Figure 35: Total monthly error (in MWh) for all PV plants together, based on the four methods explained in the text.

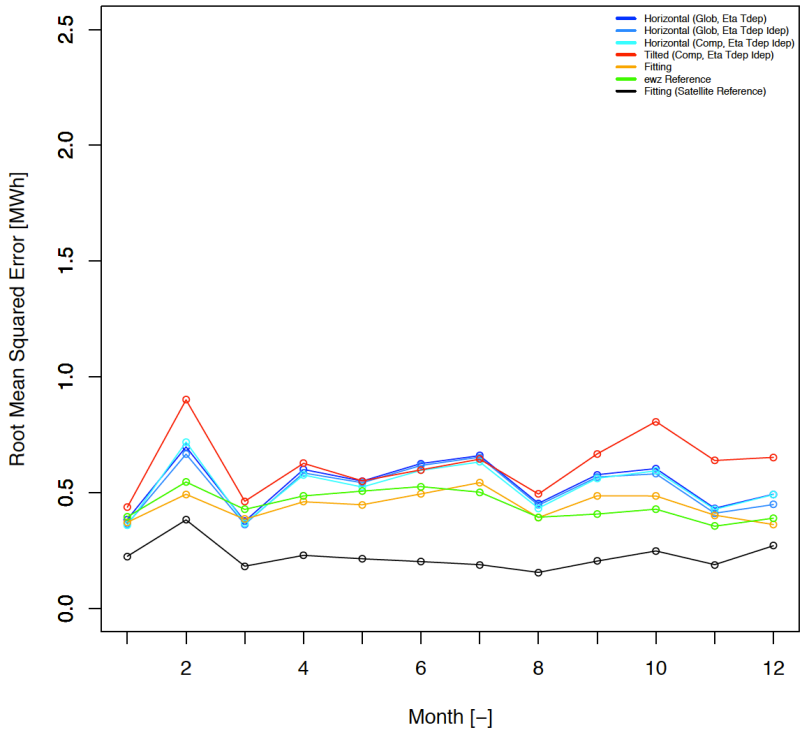
4.3.2 Comparison with ewz reference method for a one-year period

Figure 36a compares the power production forecasts of all previously discussed methods with the one of the ewz reference method for the PV plants of the city of Zurich and the year 2012 (for the previous years, no ewz reference method data ara available). This comparison, however, has to be interpreted with caution because of crucial differences between the HORIZONTAL, TILTED, and FITTING methods and the reference method of ewz (referred to as EWZ hereafter), which are listed in Table 17.

Table 17: Differences of the power production forecast methods of this study to the reference method of ewz

	HORIZONTAL / TILTED / FITTING methods	ewz reference method
Used model	COSMO-2	COSMO-7
Used model run	03-UTC run of the day before the day to be forecasted	12-UTC run of the day before the day to be forecasted
Used COSMO grid points	Most representative grid point for each PV plant	Grid point of Zurich, Affoltern for each PV plant
Set of PV plants	All PV plants of Switzerland	PV plants of the city of Zurich

(a) Root mean squared error



(b) Relative root mean squared error

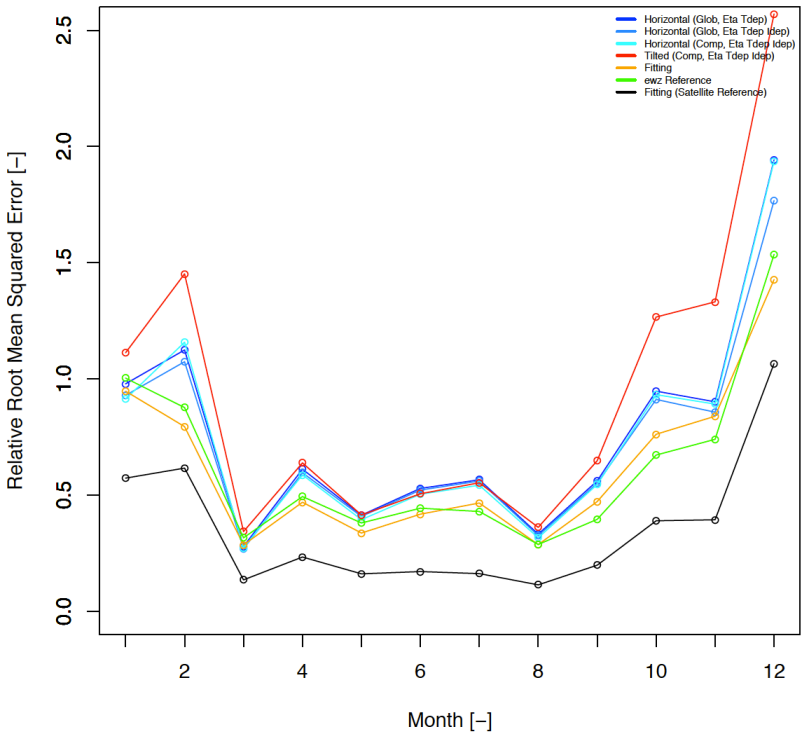


Figure 36: RMSE (in MWh) and relative RMSE (in %) for all PV plants of the city of Zurich and per 1h-time-step for each month of 2012, based on the four methods explained in the text. Additionally, the error curve of the ewz reference method is shown in green.

The RMSE magnitude (absolute and relative) of the EWZ method, shown in green in Figures 36a and 36b, is similar to the one of the FITTING method. In the first half of the year, the EWZ performs better than the FITTING method, whereas it is the other way round in the second half. Hence, the EWZ performs better than the HORIZONTAL and TILTED but worse than the satellite reference method in most of the months. What can be further observed is that the absolute RMSE development of the EWZ method is smoother than of the other ones and it does not show the same qualitative development. This could be due to the facts that the EWZ method is based on one single COSMO grid point only and on COSMO-7 instead of COSMO-2, which might have a different performance pattern for the year 2012.

5 Overview of uncertainties

Meteorological uncertainties

- COSMO irradiance and temperature forecast errors
- COSMO forecast data interpolation uncertainties
- COSMO topography uncertainties

Technical uncertainties

- Deviations of the assumed or fitted from the real metadata parameters (tilt and azimuth angles, irradiance- and temperature-dependency parameters)
- Snow cover on PV panels
- Power production measurement data errors and uncertainties
- Shading by objects, topography, and PV panel itself
- PV module efficiency model uncertainties
- Uncertainties in the conversion of the solar global irradiance on a tilted and oriented PV panel
- Uncertainties in the satellite irradiance data

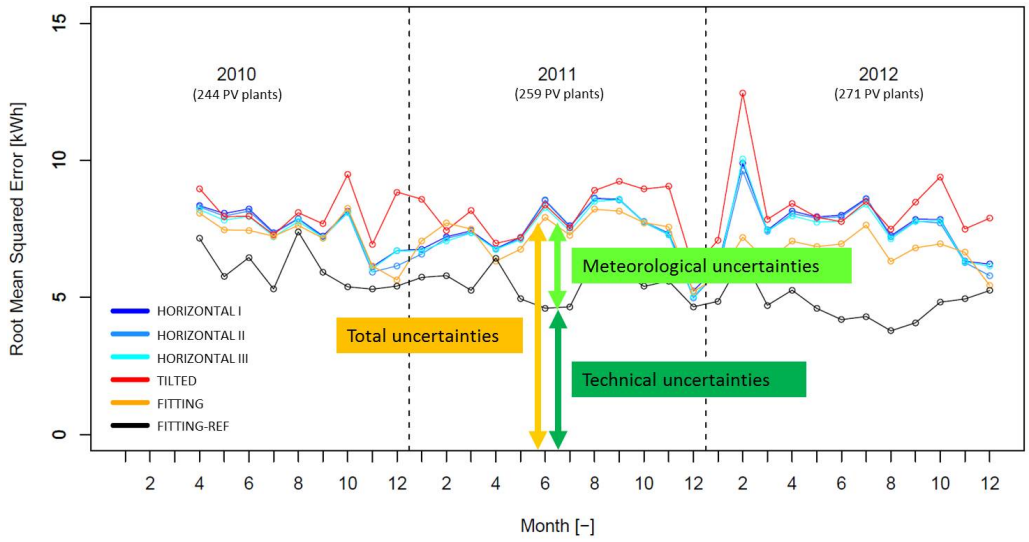


Figure 37: The total PV power production forecast error can be split into meteorological and technical uncertainties, whereas the latter seem to dominate over the former.

On the meteorological side, the dominant uncertainty are the **COSMO irradiance and temperature forecast errors**. As shown with the COSMO irradiance forecast verification (temperature is not verified), the performance of the model in predicting irradiance strongly depends on the weather conditions, which vary both temporally and regionally and thus yield hourly RMSE values of 15 up to over 100 %. The highest forecast errors seem to occur in the flatlands during fog or high fog conditions. Considering the fact that a large part of the PV plants managed by ewz are located in the northeastern Swiss Plateau with a high climatological (high) fog occurence, the (high) fog performance of COSMO is certainly the main problem with regard to PV power production forecasting. However, not only COSMO reveals difficulties in predicting (high) fog - it is rather a general problem of any NWP models.

The **COSMO forecast data interpolation uncertainties** occur during the assignment of the most representative COSMO grid points to the PV plant locations. Even though, this is a commonly used procedure to receive COSMO forecast data for any locations in Switzerland and the COSMO grid space is only around 2 km, there might be especially mountainous regions with large climatological variabilities on a small scale leading to biases when interpolating between grid points. However, since a majority of the PV plants are located in populated areas of the flatlands, these uncertainties are probably only of minor importance.

The third problem on the "meteorological" side might be the **COSMO topography uncertainties**, which also affect mainly the mountainous regions, however. As explained in the data section, COSMO uses an own, strongly simplified topography, which is smoother than the real one and has a maximum slope of 15° only. The former drawback can lead to uncertainties in the horizon lines for certain PV plant locations and thus in their diurnal sunshine duration. The latter aspect introduces uncertainties when using the COSMO global irradiance on a horizontal plane, $I_{glob,h}^{COSMO}$, which actually follows the terrain and is thus only horizontal in a relative sense. Also these uncertainties, however, are assumed to be less relevant than the COSMO forecast errors.

Even though, the FITTING method showed the possibility of the derivation of PV-plant-specific metadata parameters through a parameter fitting based on historical data, there are still remaining **deviations of the assumed or fitted from the real metadata parameters**. Of course, the deviations can be reduced when fitting the parameters PV-plant-specifically instead of assuming the same parameters for all PV plants (HORIZONTAL and TILTED methods). Among the 270 PV plants, however, there is still a number of plants, for which the parameters cannot be fitted properly. The reasons can be errors in the power production measurement data, significant shading by objects, topography, or the PV panel itself (which would require a seasonal fitting, yielding different sets of metadata parameters for different seasons and thus diurnal sun courses), or very small tilt angles (less than 5°) generating difficulties for the fitting algorithm to derive the according azimuth angles. As it is clearly shown in the sensitivity case study, all these remaining deviations can induce significant errors in the PV power production forecasts even if they are small. The most illustrative example for this problem is the PV power production forecast for the case study PV plant, which yields a sensitivity of around 1 kWh per hour when changing the tilt angle by 1° only.

Snow cover on PV panels is probably one of the most important uncertainties on the technical side. Of course, it is mainly a seasonal problem for the mountainous regions. However, certain months between 2010 and 2012 showed that also in the flatlands the weather conditions can be such that snow remains on the PV panels for days or weeks, generating "artificial" errors in the PV power production forecast of over 200 %. The question of how this uncertainty can be overcome is easy: the snow cover has to be

parameterized in the PV power production forecast model somehow, what is not done yet in this study because it did not reveal to be that essential at the beginning.

The second uncertainty caused by the power production rather than the forecast model side are the **power production measurement data errors and uncertainties**. The key problem of these uncertainties is that it is often not possible to recognize whether a certain abnormal pattern in the power production measurement data of a PV plant occurs due to environmental influences as snow cover or shading or due to technical reasons as a breakdown of the PV plant or an error in the measurement data transmission. This distinction, however, is crucial because the environmental influences can theoretically be accounted for in the PV power production forecast model, whereas the technical ones can not.

The **shading by objects, topography, and the PV panel itself** is another external and very PV-plant-specific source of uncertainty. The strength of a statistical as the FITTING method is its (potential) ability to capture at least the significant disturbances through shading for every individual PV plant by implicitly including it through the fitting of the metadata parameters. Therefore, the shading uncertainties are assumed to be of much higher importance within the methods as HORIZONTAL and TILTED, which are based on the same metadata assumptions for all PV plants.

The use of the two models from *Beyer et al.* (2004) and *Wagner* (2014) automatically introduces **PV module efficiency model uncertainties**. The strength of the first model by *Beyer et al.* (2004) is its easy applicability for large-scale PV power production forecasts, since only few parameters have to be estimated. Also are the parameters in the model physically meaningful and partly even given by the manufacturer (as for instance the temperature coefficient α). However, for a statistical as the FITTING method, the model does not seem to be very suitable due to the linear dependency of its parameters, leading to difficulties for the fitting algorithm to derive them. The simplicity of the model also yields a drawback: the heating factor γ seems to be the weakest part of the model since it has to summarize too many aspects influencing the heating of a PV module in one single parameter (as irradiance, wind, installation type etc.). Therefore, the use of one single value for γ seems to be too inaccurate even for a single PV plant, as the sensitivity case study shows. Different values for winter and summer might already improve the performance of the PV power production forecast. The sensitivity case study also shows that the assumptions of a γ of 0.027 °C m W^{-1} for free-standing and 0.056 °C m W^{-1} for roof-integrated PV panels, as they are spread in the literature, do not seem to be applicable in general. The second model by *Wagner* (2014) suits better for a statistical as the FITTING method. Its parameters, however, do not have a clear physical meaning anymore, since they are constructed for the purpose of parameter fitting. How this model performs in fully physical as the HORIZONTAL or TILTED methods (compared to the other one of *Beyer et al.* (2004)), cannot be concluded here since it is only applied in the FITTING method of this study.

Uncertainties in the conversion of the solar global irradiance on a tilted and oriented PV panel are assumed to be minimal since the responsible algorithms have been compared and verified with other commonly used algorithms at MeteoSwiss, revealing quite similar results. Nevertheless, uncertainties can still occur within different PV power production forecasts, because the calculation steps for the diurnal solar cycle are based on empirical equations, of which different ones exist in the literature.

Even though the FITTING-REF method (shown in Figure 37) is based on meteorological measurement data (satellite and surface stations) and thus assumed to be free from meteorological uncertainties, there are indeed also **uncertainties in the satellite irradiance data**. This can be uncertainties in

the algorithms of the HelioMont satellite itself, its data post-processing, or the temporal and spatial interpolation of its data for the PV power production forecast. The sum of these uncertainties would have to be subtracted from the error of the FITTING-REF method (shown in Figure 37) to get the "real" technical uncertainties. To do so, a satellite data verification with surface measurements at each PV plant location would be required. This would, however, go beyond the scope of this study.

6 Outlook and recommendations

6.1 Priority aspects

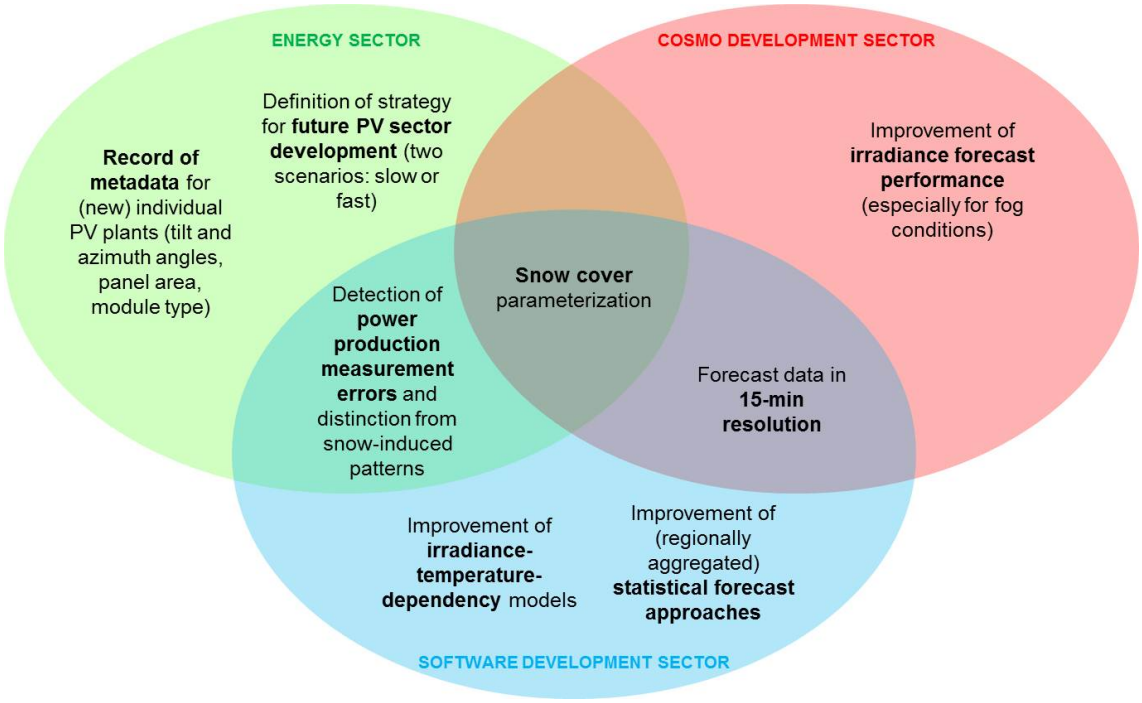


Figure 38: Outlook and recommendations for the development of a PV power production forecast software for Switzerland, divided into the three stakeholders: the energy, the COSMO development, and the software development sectors. The overlapping areas represent aspects, for which two or three sectors are responsible and thus cooperation and knowledge transfer is required.

The main objective of this study was to determine, quantify, and discuss the different errors and uncertainties associated with a PV power production forecast. To a large degree, this objective could be accomplished. Concretely developing an operationally usable PV power production forecast software for the Swiss power market would be part of future projects. Based on the results and conclusions of this study, this final section of the report aims to give an overview of the most important issues, which need to be considered or require further investigation in case of such a software development. Of course, this overview is not complete and might leave out important issues, which are analyzed in any other studies.

The aforementioned overview is given in Figure 38, which is divided into three sectors representing the stakeholders associated with the development of a PV power production forecast software: the energy,

the COSMO development, and the software development sectors. The different issues are assigned to either one, two, or all of the sectors, depending on whether it is an aspect for specifically one sector or whether it is one requiring the contribution and knowledge transfer of two or all sectors. Details about each aspect are provided in the following.

Improvement of irradiance forecast performance (especially for fog conditons)

Even though in total, the technical seem to dominate over the meteorological uncertainties, an improvement of the COSMO forecast of (high) fog and the associated irradiance would improve the performance of a PV power production forecast especially for the flatland regions and in winter. This is an issue, the COSMO development sector is aware of and about to improve in the future. Beside the explicit physical improvement of the COSMO model code, one option would be to apply a model output statistics (MOS) system either specifically on the irradiance forecast parameters or on the COSMO model output in general. With the regular occurence of fog events and thus a large available set of historical fog data, such a MOS approach could probably enhance the fog prediction significantly. Since an improvement and further development of the code of a NWP model or the introduction of a MOS system is a long-term process, a temporary solution for an improvement of a PV power production forecast could be the use of a persistence approach within the PV power production forecast software (what would include the software development sector as well): during typical conditions for fog (as for instance the mornings after clear-sky nights during persistent high-pressure conditions) or high fog (as for instance the cold north-easterly winds in winter called “Bise”) over several days, an algorithm could be implemented in the software to recognize these conditions based on meteorological measurements or the power production measurements in a first step (strong fog probably results in relatively constant PV power production over the whole day without any diurnal course) and using them as power production forecast for the following day in a second step (based on the assumption that the conditions on the following day would remain the same). Of course, this approach would be very elaborate to develop, especially concerning the detection of such conditions. However, already a simple form of it might help to reduce the power production forecast error over several days.

Forecast data in 15 min resolution

This aspect is put in the fields of responsibility of both the COSMO and the software development sectors due to the following reasons: from the COSMO side, irradiance and temperature forecast data have to be made available operationally in a higher temporal resolution than 1 h. Currently, COSMO forecast outputs in 10 min steps could be made available. However, the costs for this resolution would be higher and an interpolation to the 15 min values would still be necessary. If for any reasons, still the use of COSMO forecast data with a temporal resolution of 1 h should be used, it is up to the software development sector to interpolate them to 15 min values. The according approach could be a trigonometric interpolation. Figure 39 shows such an example for irradiance, with the detailed procedure explained in the caption.

Definition of strategy for future PV sector development (two scenarios: slow or fast)

This is the only non-technical but rather economic or political aspect and is important for the general framework a PV power production forecast software has to be embedded in. In particular, there are two general scenarios for the PV sector in Switzerland: either does the number of PV plants increase

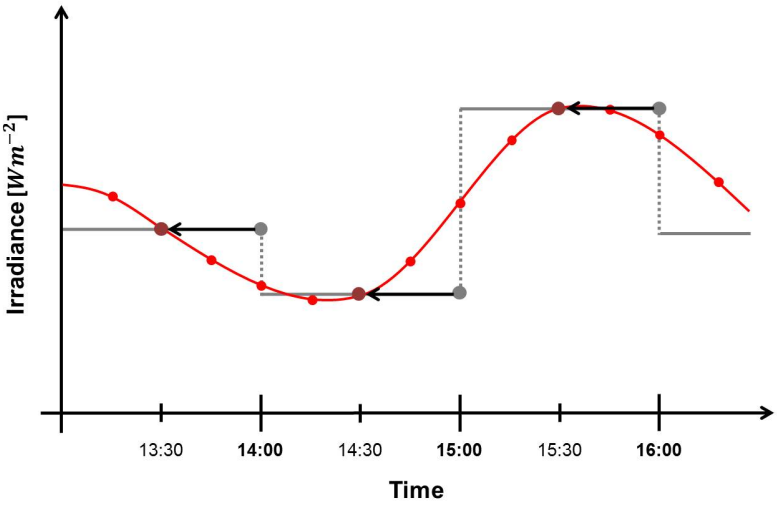


Figure 39: Example for a trigonometric interpolation of COSMO irradiance forecast values from a 1 h to a 15 min resolution: from the three forecast values for 14:00, 15:00, and 16:00 (gray dots), which express the mean irradiance value over the hours between 13:00 and 14:00, 14:00 and 15:00, and 15:00 and 16:00 (indicated by the gray horizontal lines), respectively, the 15 min values should be interpolated. This is done by assigning (indicated by the arrrows) the COSMO output values of 14:00, 15:00, and 16:00 to the ones at 13:30, 14:30, and 15:30 (dark red dots), and then triognometrically interpolating between the latter to get the values every 15 min (red dots and line).

slowly over the next decades or does the energy sector expect a large number of newly constructed PV plants every year und thus a rapid increase. Which scenario is going to be pursued depends on both the economic and political framework conditions associated with the following questions: are the Swiss feed-in tariffs high and thus attractive enough for private persons to install an own PV plant? Do the Swiss energy companies pursue a strategy of buying power mainly from Swiss PV plants or rather from large PV parks abroad? Is the management of the privately owned PV plants all over Switzerland still going to be in the responsibility of the different balance groups or is there a tendency towards a single balance group managing all PV plants? As these questions indicate, the setting of this framework not only affects the balance groups as ewz but also swissgrid and the politics. With the legal decision of the Swiss Federal Council and Parliament to withdraw from the use of nuclear energy step by step within the next decades (referred to as Energy Strategy 2050; *SFOE* (2014)), answers to the aforementioned questions can be expected in the following years. The reason for the importance of the framework for a PV power production forecast software is technical: the aforementioned slow scenario would assure a relatively constant or at least slowly increasing number of PV plants. This means, for most of the PV plants a rather long set of historical power production measurement data would be available. In this case, the use of a statistical PV power production forecast approach (as applied in the FITTING method of this study) would be best. Assuming the fast scenario associated with a high number of new PV plants every year, however, a large part of the PV plants would not have any long time series of historical power production data. This would detoriate the data basis for a statistical approach and hence a fully physical approach would have to be pursued by the PV power production software.

Improvement of (regionally aggregated) statistical forecast approaches

In case of a previously described slow scenario, an improvement and further development of the statistical forecast approach (as applied in the FITTING method) is required. First of all, the results of the

parameter fitting have to be verified more accurately and structured, and not just based on random PV plant samples (as it is done for the FITTING method). This helps to refine the fitting algorithm and to understand its strengths and weaknesses. Furthermore, it has to be analyzed whether an approach fitting the parameters for each individual PV plant (as in the FITTING method) or one fitting a single parameter set for a whole region containing many PV plants (as in the ewz reference method) performs better. An advantage of the latter approach is the fact that through the aggregation it might smooth out small-scale variabilities and thus errors in the fitted parameters, which occur in the former approach. However, the latter approach would require a careful definition of the different regions, in which all associated PV plants would have similar meteorological conditions. Also would it be necessary to assign one single, most representative COSMO grid point to every region. In the Swiss Plateau, this would probably be easier to be achieved, whereas in the Alpine regions the small-scale meteorological variabilities would make the definition of such regions much more difficult. These aspects are reflected in the ewz reference foreacast of this study: even though, the method is based on one single parameter for all PV plants in the city of Zurich (over 100), the power production forecast performs similarly well as the one of the FITTING method, which is based on specific parameters for each individual PV plant. However, the same procedure for a group of PV plants in the mountainous Grisons would probably yield a worse power production forecast performance due to the smaller meteorological homogeneity than in Zurich.

Beside the PV power production forecast for each individual PV plant of the region and for the region as a total, there is a third approach applicable if metadata parameters as for instance the nominal power P_{nom}, the tilt angle β, or the module type are not available for each individual PV plant. An example for such an approach is provided by *Lorenz et al.* (2010): the total area (country) is divided into regions to be forecasted, whereas each region has a set of representative PV plants. The power production forecast is only computed for the representative PV plants and the received result is then up-scaled to get the power production forecast for the whole region. This approach works well if the representative PV plants are selected carefully and if the total number of PV plants is large enough. The latter criterion is fulfilled in the study of *Lorenz et al.* (2010), which is based on over 200'000 PV plants all over Germany. Based on the approach applied in this study, a similar classification is made for Switzerland, as shown in Figure 40: the map at the top shows all around 270 PV plants managed by ewz. The size of the red circles represents the nominal power P_{nom} of the PV plants (the larger the radius, the larger the nominal power). The black squares are the PV plants exemplarily defined as the representative ones (for these plants, ewz could provide additional detailed information about the metadata), whereas the size again indicates the nominal power. An appropriate set of representative PV plants would be one representing the amount of totally installed nominal power in the accoring region as well as possible (which is one criterion defined by *Lorenz et al.* (2010)). To check this criterion for the representative PV plants indicated by the black squares, the two maps at the bottom of Figure 40 are shown: both maps divide Switzerland into grid cells of $1^\circ \times 1^\circ$ (the borders or grid points are indicated by the green crosses). The map on the left-hand side shows the aggregated totally installed nominal power of each region, indicated by the size of the black square. It shows the regions with most of the PV plants: the cantons of Zurich and Grisons. The map on the right-hand side finally indicates whether the selected set of representative PV plants is appropriate to represent the relative distribution of nominal power as shown in the bottom left map. It is calculated as follows: for each region (grid cell), the sum of the nominal power of its representative PV plants is multiplied with the sum of installed nominal power all over Switzerland divided by the sum of installed nominal power of all PV plants of this region. The

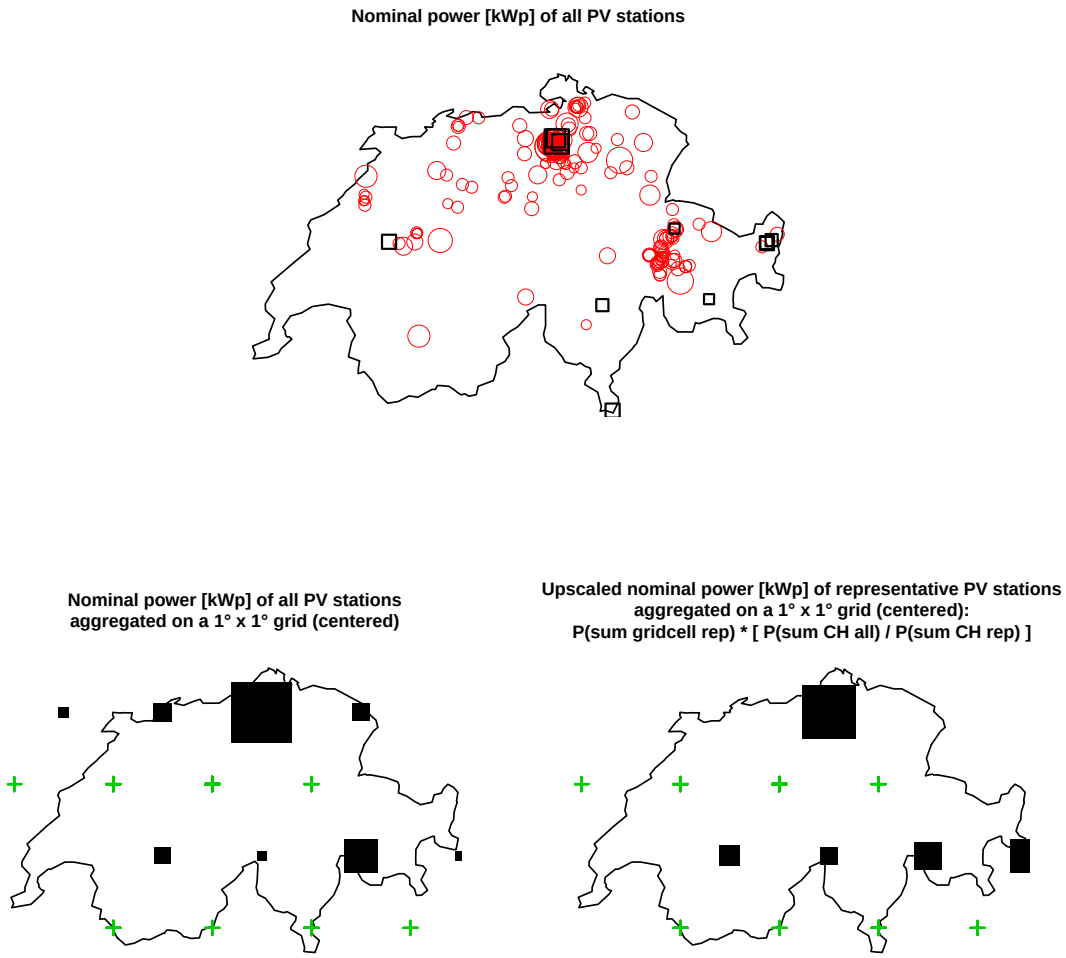


Figure 40: Example for an upscaling approach, which could be applied if detailed metadata are only available for a subset of the PV plants. The explanations are given in the text.

comparison of the left and right maps shows that the set of representative PV plants is not appropriate for all grid cells: for instance in the very eastern part of the canton of Grisons, the representative PV plants overestimate the total relative amount of installed power in this region. The opposite occurs in northwestern Switzerland, where no representative PV plants are located even though there should be some. Figure 40 aims to show, how the up-scaling approach presented by *Lorenz et al.* (2010) could be applied on Switzerland. Due to the relatively small total number of PV plants managed by ewz (around 270), this approach would probably not make sense. However, assuming a Swiss balance group managing many more PV plants all over Switzerland, the approach might become applicable.

Record of metadata for (new) individual PV plants (tilt and azimuth angles, panel area, module type)

In case of a previously described fast scenario, a fully physical PV power production forecast approach not based on any historical power production measurement data would be required. As shown with the methods HORIZONTAL and TILTED, this would, however, require detailed information about the metadata of each PV plant. For this study only the location and the nominal power P_{nom} are known for each PV plant. In order to significantly improve a PV power production forecast software, the balance group managing the PV plants would have to record additional metadata parameters for each individual PV plant - not just for newly installed but also for already existing PV plants. According to the findings of this study, Table 18 lists all the metadata parameters, which are necessary to know, and the ones, which would be good to know for a fully physical PV power production forecast.

Table 18: PV plant metadata parameters, which are necessary to know, and the ones, which would be good to know for a fully physical PV power production forecast.

Parameter	Importance of availability
PV plant location (latitude and longitude)	Necessary
PV panel tilt angle β	Necessary
PV panel azimuth angle ψ	Necessary
Nominal power of the PV plant P_{nom}	Necessary
PV panel area A or PV module efficiency at standard conditions η_{stc}	Necessary
PV module type (including technical data sheet from the manufacturer to extract parameters as the temperature coefficient α, the PV module efficiency at standard conditions η_{stc}, and the nominal operating cell temperature T_{noct})	Necessary
Installation type (free-standing or roof-integrated, type of building etc.)	Good to have
Operation start date (to calculate the age of the PV plant)	Good to have
Information about the maintenance of the PV plant (as if snow on the PV panel is regularly wiped away by the owner etc.)	Good to have

Improvement of irradiance-temperature-dependency (= module efficiency) models

Concerning the module efficiency models applied in this study, there is potential for improvement. In particular, the parameterization of the module temperature T_{mod} and the heating coefficient γ, respectively, might be refined to improve the representation of the different factors influencing the cooling or heating of a PV module. Even though this aspect rather belongs to the field of responsibility of the research field for PV power production forecasting in general, it is assigned to the field of the software development sector in Figure 38 because it might be helpful for the research field to know the requirements of the software development in terms of computational applicability of a PV module efficiency model. Of course, the two applied models applied in this study are only a small part of the ones currently available in the literature. Hence, further literature research is the first step before improving any existing models.

Detection of power production measurement errors and distinction from snow-induced patterns

Independent of whether a statistical or physical PV power production forecast approach is applied, a crucial issue is to develop algorithms detecting erroneous patterns in the power production measurement data of the PV plants. The detection itself, however, is only the first step: in a second step, the algorithm has to be able to decide whether the found patterns (as for instance power production values of zero for several days) are due to environmental reasons as mainly snow cover or due to any technical reasons as a break-down of the PV plant. If snow cover is the reason, then it can actually not be referred to as error in the measurement data but rather reality that has to be captured by the meteorological scheme of the PV power production forecast model. If any technical causes are responsible for the error, it is not up to the meteorological scheme anymore to account for it but to a separate algorithm "telling" the PV power production forecast model not to compute any forecasts for the affected PV plant anymore due to its technical problems. The latter case also helps the responsible energy company or balance group to let repair the affected PV plant. Such detection algorithms have to be developed by the software development sector. However, the energy sector has to support both the development and operation of the algorithm with technical expertise about the different factors, which can influence the PV power production and its measurement. Therefore, this aspect is listed in the interface of the two sectors in Figure 38.

Snow cover parameterization

According to the findings of this study, snow cover is a crucial aspect to be parameterized in a PV power production forecast model in order to significantly minimize forecast errors in the winter months. It is listed in the middle of Figure 38, since it is a complex aspect requiring the contribution of all three sectors: the energy sector can provide information about construction and maintenance of individual PV plants, which can help to understand if and how persistently snow remains on the PV panels. The COSMO development sector is responsible to provide snow forecast and measurement data for the PV plant locations. Combining the information and data from these two sectors, the software development sector then has to include a snow cover parameterization into the PV power production forecast model. Such a parameterization should consist of two main parts: the first one is based on the COSMO snow forecast data (in combination with air temperature data) and aims to predict significant snow fall events and thus limitations in the PV power production for the following days. The second part is based on snow measurement data and pursues a persistence approach: the scheme recognizes the currently lying snow and its height according to snow measurement data (from surface snow stations or maybe also satellites). Depending on the general weather predictions for the day-ahead (based on COSMO), the algorithm then decides whether the snow will remain until the following day. If it does remain, the PV power production measurement data of the current (or the past) day is used as power production forecast for the following day, which is the persistence approach. An exemplary approach to parameterize snow cover can be found in *Lorenz and Heinemann* (2012).

6.2 Further ideas

This subsection provides an (incomplete) list of further ideas that came out of this study or the associated literature research. They are, however, not crucial for the development of a PV power production forecast software but might still help to further understand certain influence factors:

- So far, no extended and detailed **COSMO irradiance forecast performance verification** is available. Even though, this study provides an approximate picture of the seasonal and conditional COSMO irradiance forecast performance based on satellite, no ultimate conclusions can be drawn out of these results. The reason is that the verification is only performed for a set of locations irregularly distributed over Switzerland. Hence, a nationwide verification is used, if possible based on surface station measurements having a significantly smaller error than the satellite. Additionally, it would make sense to verify the irradiance components separately and also to analyze individual regions of Switzerland in order to understand the regional COSMO irradiance forecast performance better.
- To understand the impact of the increasing forecast uncertainty of a **COSMO model run** with increasing lead time on the PV power production forecast, a computation of the latter based on newer model runs than the one at 03:00 UTC would be useful. As described in the data section, this would not make sense from the perspective of a PV power production forecast software embedded in the Swiss power market (because a later model run than 03:00 UTC would not meet the requirements of a balance group). However, it could demonstrate how large the contribution from the limitation by the model run selection to the total PV power production forecast error is.
- The performance of COSMO could be further highlighted by computing the developed PV power production forecast approach based on both COSMO and any **other NWP models** as the one of ECMWF. This would help to detect strengths and weaknesses of COSMO relative to other models with regard to PV power production forecasting.
- To enhance the general understanding of the characteristics of the power production by a PV plant, it is strongly recommended to cooperate with research institutes operating **PV test plants**, which are equipped with additional measurement devices as irradiance, temperature, or snow height measurement stations. This would resemble the bottom-up principle: understanding the PV power production forecast errors on a small scale in order to reduce the same on a required large scale.
- For the improvement of the PV power production forecast approach itself, several further options exist. One option to account for the shading would be to use 360° **horizon lines** for each individual PV plant location. This could be achieved with the help of an already existing tool at MeteoSwiss, which is able to generate the horizon line for any desired location in Switzerland. However, the use of such horizon lines might conflict the topography and the according irradiance of COSMO, what would have to be considered somehow.
- Other possibilites to improve the PV power production forecast are any extensions of the used parameter set or the applied models: the inclusion of the third component of global irradiance, the **ground-reflected irradiance**, might improve the result in certain regions and conditions (as for instance a steep mountainous region with snow-covered ground having a high albedo). Also the diffuse irradiance component has potential for improvement: instead of an isotropic, a more realistic **anisotropic irradiance model** could be used (for instance the one by *Perez et al.* (1987)).
- A once finished software should be **compared to other PV power production forecast softwares available on the market**. This would help to classify the own software within the available ones and thus to recognize potential for further improvement.

- In cooperation with the balance group demanding a PV power production forecast software, it might also be useful to develop an **approach to monetize the PV power production forecast error**, which was one of the original objectives also of this study. However, this would require either a set of assumptions or more detailed information about the (partly confidential) total power production data (not only from new renewable sources as PV) of the balance group in order to decide whether a certain negative or positive error in the PV power production forecast could be compensated by any other power sources or not.

A Appendix

Literature values for η_{STC} and α

Table 19: η_{STC} and α values from 19 different literature studies, summarized in *Dubey et al.* (2013), which are used as basis for the calculation of the mean values $\eta_{stc} = 0.117$ and $\alpha = -0.004\,18\,^{\circ}\mathrm{C}^{-1}$, as explained in the methods section.

T_{mod} [°C]	η_{stc} [-]	α [°C^{-1}]	Comments
25	0.15	0.0041	Mono-Si
28	0.117	0.0038	Average of Sandia and commercial cells
25	0.11	0.003	Mono-Si
25	0.13	0.0041	PVT system
-	-	0.005	PVT system
20	0.1	0.004	PVT system
25	0.1	0.0041	PVT system
20	0.125	0.004	PVT system
25	-	0.0026	a-Si
25	0.13	0.004	Mono-Si
-	0.11	0.004	Poly-Si
-	0.05	0.0011	a-Si
25	0.178	0.00375	PVT system
25	0.120	0.0045	Mono-Si
25	0.097	0.0045	PVT system
25	0.09	0.0045	PVT system
25	0.12	0.0045	PVT system
25	0.12	0.0045	PVT system
25	0.127	0.0063	PVT system
25	0.127	0.006	PVT system
25	0.117	0.0054	PVT system

Abbreviations

COSMO	Consortium for Small-Scale Modelling
ECMWF	European Centre for Medium-Range Weather Forecasts
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EWZ	Elektrizitätswerk der Stadt Zürich
MAE	Mean absolute error
ME	Mean error
MOS	Model output statistics
MPP	Maximum power point
MSG	Meteosat Second Generation
NOCT	Nominal operating cell temperature
NWP	Numerical weather prediction
PV	Photovoltaic
RMSE	Root mean squared error
SMN	SwissMetNet

List of Figures

List of Figures

List of Tables

References

Beyer, H. G., J. Bethke, A. Drews, D. Heinemann, E. Lorenz, G. Heilscher, and S. Bofinger (2004), Identification of a general model for the MPP performance of PV modules for the application in a procedure for the performance check of grid-connected systems, *19th European Photovolatic Solar Energy Conference & Exhibition, Paris 2004*, pp. 1–4.

Dubey, S., J. N. Sarvaiya, and B. Seshadri (2013), Temperature Dependent Photovoltaic (PV) Efficiency and Its Effect on PV Production in the World - A Review, *Energy Procedia*, *33*, 311–321, doi:10.1016/j.egypro.2013.05.072.

Evans, D. L., and L. W. Florschuetz (1977), Terrestrial concentrating photovoltaic power system studies, *Solar Energy*, *20*, 37–43.

Ewz (2013), Elektrizitätswerk der Stadt Zürich (ewz), *http://www.ewz.ch*.

Frei, C. (2014), Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, *International Journal of Climatology*, *34*(5), 1585–1605, doi:10.1002/joc.3786.

Google (2013), Google Maps, *http://www.maps.google.ch*.

Haeberlin, H. (2012), *Photovoltaics System Design and Practice*, 1 ed., 732 pp., Wiley.

Heinemann, D. (2002), *Energy meteorology (postgraduate programme 'renewable energy')*, 11–17 pp., Carl Von Ossietzky University, Oldenburg.

ITACA (2014), The sun as a source of energy, *http://www.itacanet.org/the-sun-as-a-source-of-energy/*.

Lorenz, E., and D. Heinemann (2012), *Prediction of solar irradiance and photovoltaic power*, vol. 1, 239–292 pp., Elsevier Ltd., doi:10.1016/B978-0-08-087872-0.00114-1.

Lorenz, E., T. Scheidsteger, J. Hurka, D. Heinemann, and C. Kurz (2010), Regional PV power prediction for improved grid integration, *Progress in photovoltaics: research and applications*, doi:10.1002/pip.

MeteoSwiss (2010a), Klimabulletin Oktober 2010, *Zürich*.

MeteoSwiss (2010b), Klimabulletin Dezember 2010, *Zürich*.

MeteoSwiss (2011), Klimabulletin November 2011, *Zürich*.

MeteoSwiss (2012), Klimabulletin Februar 2012, *Zürich*.

Nordmann, T., and L. Clavadetscher (2003), Understanding temperature effects on PV system performance, *Proceedings of 3rd World Conference on Photovoltaic Energy Conversion*, *3*, 2243–2246.

Perez, R., R. Seals, P. Ineichen, R. Stewart, and D. Menicucci (1987), A new simplified version of the Perez diffuse irradiance model for tilted surfaces, *Solar Energy*, *39*(3), 221–231.

PVeducation (2014), PVeducation, *http://www.pveducation.org*.

SFOE (2014), Energy Strategy 2050, *http://www.bfe.admin.ch/themen/00526/00527/*.

Spencer, J. W. (1971), Fourier series representation of the position of the sun, *Search*, *2*, 172.

Stoeckli, R. (2013), The HelioMont surface solar radiation processing, *Tech. Rep. 93*, Federal Office of Meteorology and Climatology MeteoSwiss.

Wagner, J. (2014), Final report of the PV-Alps Interreg project (in preparation, personal communication).

Wuerfel, P. (2009), *Physics of Solar Cells - From Basic Principles to Advanced Concepts*, Wiley-VCH, Berlin.

Acknowledgment

I would like to express my greatest gratitude to all the people at MeteoSwiss who supported me during the project and the realization of this report: first of all, I am truly indebted and thankful to both my boss and supervisor Dr. Saskia Willemse and my office mate and co-supervisor Dr. Agnes Richard for supporting me whenever necessary in a very kind way and for being such a great team to work with. I wish to thank Dr. Reto Stöckli for providing his great expertise in solar irradiance modeling and satellite climatology, Jacques Ambühl for supporting me with founded knowledge about COSMO, Dr. Laurent Vuilleumier for sharing with me his expertise in solar irradiance nowcasting, Dr. Christoph Frei for his useful hints concerning the use of surface station measurement data, Francis Schubiger for the inputs based on his knowledge about COSMO irradiance forecast performance, and Dr. Christian Sigg for the discussions about algorithms detecting errors in the power production measurement data. Furthermore, I am obliged to Silvana Baselgia for the initialization of this project with ewz and to Pierluigi Pestrin for providing his expertise about the current and future Swiss PV power market. I also wish to thank Dr. Martin Brändli, Marc Musa, and Dr. André Walser for their great technical support in setting up and maintaining the unavoidable database.

I also would like to thank the whole team of ewz, namely Dr. Silvia Banfi Frost, Romina Schürch, Annina Vinzens, Sergio Taiana, Markus Rappo, Tobias Diekmann, and Prof. Anne Kress for providing historical PV power production measurement data in a high spatial and temporal resolution and for contributing to the project with a crucial technical and economic energy sector perspective.

Finally, I am obliged to all the external contributors: I would like to thank my internship supervisor at ETH Zurich, Prof. Heini Wernli, for his useful scientific inputs, Markus Markstaler (NTB Buchs, Switzerland) for providing me with his expertise about snow cover on PV panels, and to Prof. Hans-Georg Beyer (University of Agder, Norway) for his important information about PV module efficiency models.

MeteoSchweiz
Operation Center 1
CH-8058 Zürich-Flughafen

T +41 58 460 91 11
www.meteoswiss.ch

MeteoSvizzera
Via ai Monti 146
CH-6605 Locarno Monti

T +41 91 756 23 11
www.meteosvizzera.ch

MétéoSuisse
7bis, av. de la Paix
CH-1211 Genève 2

T +41 22 716 28 28
www.meteosuisse.ch

MétéoSuisse
Chemin de l'Aérologie
CH-1530 Payerne

T +41 26 662 62 11
www.meteosuisse.ch