

ANALYSIS OF PV GRID INSTALLATIONS PERFORMANCE, COMPARING MEASURED DATA TO SIMULATION RESULTS TO IDENTIFY PROBLEMS IN OPERATION AND MONITORING

Bruno Wittmer *, André Mermoud *, Thomas Schott †
 * PVsyst SA (Analysis)

Route du Bois de Bay 107, CH 1242 Satigny, Switzerland

† Berner Fachhochschule, Institute for Energy and Mobility Research, Photovoltaic Laboratory (Data providing)
 Jlcoweg 1, CH-3400 Burgdorf

ABSTRACT: Monitoring data of grid-connected photovoltaic installations provides a valuable source to follow the operation and performance of a PV facility. Main failure modes can be spotted from the analysis of only this data. Combining measured monitoring data with detailed simulation results however, provides additional information that can reveal more subtle problems or help to track down the source of a complex malfunction. In this paper we describe how such a comparison can be performed, that would allow the detection of different kinds of problems on the module, string and inverter level.

After some general considerations on how to compare measured data to simulation results, including the corresponding uncertainties, we study how possible discrepancies can point to problems either with the monitoring, the operation or the simulation results. The level of detail in the monitoring data will determine how well this assignment can be done in an unambiguous way.

A tool allowing this kind of analysis is currently being implemented in the PVsyst software and its capabilities to compare monitoring data to the simulation results are being expanded.

Keywords: Evaluation, Monitoring, Performance, Simulation, Software, System Performance

1 INTRODUCTION

The aim of the study is to propose a methodology for comparing measured data of a PV installation with simulation results. The simulation of a PV installation uses as input several meteorological variables, namely the irradiance, the temperature and the wind speed. The irradiance can be given either in the horizontal plane, or in the plane of incidence of the PV modules. In the planning phase of a PV installation the irradiance is normally only known in the horizontal plane and has to be transposed to the plane of incidence. When monitoring the operation of an installation, there is often a measurement available in the plane of incidence. Furthermore, the irradiance has to be decomposed into the direct, diffuse and albedo components. These components are normally not measured separately, and the simulation software has to recur to models to estimate the share of the global irradiance among the three components.

Another set of input parameters needed for the simulation, is the detailed description of the installation. This includes the type of PV modules, the orientation in tilt and azimuth of the modules, wiring resistances, the inverter specifications, etc. The PV simulation software takes all these inputs and simulates the state of the PV installation for every time step (usually one hour). Before calculating the final result of a time step, which is the delivered AC power, many intermediate results are computed, like the optical losses of the incident light, the losses in the PV conversion due to temperature and light intensity, ohmic losses, losses in the conversion from DC to AC, etc. When comparing measured data to the simulation, these intermediate results allow to perform very sensitive crosschecks to ensure a good description by the simulation, but also to spot possible malfunctions of the installation.

In the following we describe a general approach for performing this kind of comparisons, focusing on spotting problems in the measurement accuracy, the simulation results and the operation of the installation.

2 COMPARISON METHOD

When monitoring a PV installation, the measurements usually include the irradiance, the temperature and the generated power. These measurements have all an equivalent variable in the simulation process, to which they can be compared. Figure 1 shows a selection of the most important variables that are used in the simulation. The variables highlighted in red, are the ones which have a corresponding measurement in the examples to follow. The comparison of variables only makes sense for variables that are not a direct input to the simulation.

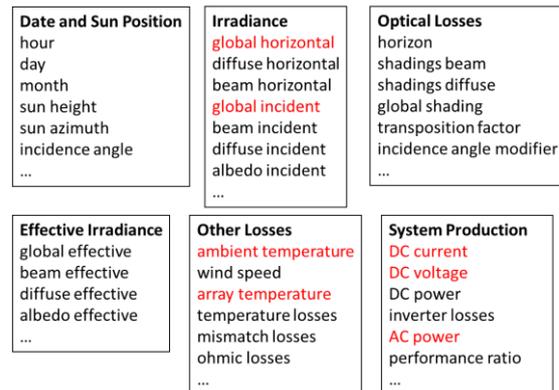


Figure 1: Selection of variables that are used in the simulation. The highlighted variables are the ones that are also measured in the examples of this study.

2.1 Indicators

For the comparison, one first plots the measured values against the simulated ones, the expected result being points that lie on a straight line with slope one and intercept zero. On the left side of Figure 2 this is shown for the AC power generation of the examples. As a measure of the matching between measurement and simulation, one can take several indicators:

Linear Fit Results

Performing a linear regression through the points yields four different indicators for the match. The slope, the

intercept, the residuals and the standard deviation of the residuals. For a good match, the slope and intercept should be close to one and zero respectively. The standard deviation of the residuals should be compatible with the measurement and simulation uncertainties. The residuals themselves, can be used to crosscheck against systematic errors as explained below.

Difference of Measured and Simulated Values

The difference between measured and simulated values corresponds to the residuals with respect to the ideal straight line with slope one and intercept zero. Three different indicators can be obtained: the differences themselves, the average and the standard deviation of the values. The average should be close to zero for a good match, and the standard deviation compatible with the measurement and simulation uncertainties. The individual differences can again be used to check for systematic errors.

Ratio between Measured and Simulated Values

This indicator is equivalent to the previous differences, which are now expressed as a percentage. Again average and standard deviation can be used as indicators. The individual values can be used to check for systematic errors.

As a generalization of these comparisons, it is also possible to use two variables which do not describe the same quantity, but for which a linear behavior is expected. Examples will be given in Section 4.

2.2 Checking for Systematic Errors

The indicators described above are either single values, like the slope, averages and standard deviations, or they are sets of individual values. The sets of individual values can be plotted against other variables (the control variables), to check for systematic errors. The control variables can be either measured values or simulation variables, including the ones that are being used for the comparison. If we take for example the residuals of the linear fit, the scattering around the straight line is caused by the measurement and simulation uncertainties. If these uncertainties were all of statistical nature, one would expect no correlation at all, when plotting the residuals against any control variable. The points should scatter around the x-axis. Any correlation that can be seen in this kind of plot, points to a systematic error either in the measurement or in the simulation.

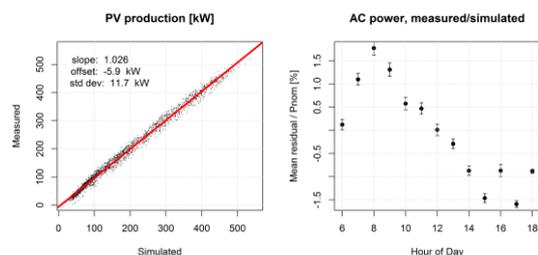


Figure 2: Analysis example for generated AC power. Left: plot of measured against simulated variable (AC power generation) with linear fit. Right: Residuals of linear fit as function of the control variable (Hour of Day).

This kind of correlation plot contains a few thousand points, when simulating a full year, which makes the plot difficult to read. To improve the visualization, the points are grouped according to the control variable. For each group the average, standard deviation σ and number of

data points N is determined. In the plot the average is displayed with a symmetric error bar of size σ/\sqrt{N} , which describes the scattering of the points in the group. In the example of figure 2, the right side shows the residuals as function of the Hour of Day.

3 MEASURED DATA AND SIMULATION

3.1 Studied PV Installation

For the present paper we examined the PV installation Montsoleil, built in 1994, which is located in the Jura Mountains in the northwest of Switzerland. It has a nominal power of 555 kWp, which is provided by 10464 Siemens M55 modules. The modules are mounted with a tilt angle of 50° and are facing two different azimuths (20°E and 35°E). The DC power is converted by a single ABB inverter prototype and injected into the grid at 16 kV. Further details on the installation can be found at [1].

3.2 Simulation

The simulation was performed with the software PVsyst, Version 6.39. The first step is to generate an input file for the simulation containing the meteorological data in hourly values. The simulation needs as input for the irradiance either the global horizontal irradiance or the global incident irradiance, which is also called plane of array (PoA) irradiance. Furthermore, the simulation can work either with the measured ambient temperature and compute the module temperature using a thermal model and a heat transport coefficient U_c , or it can use directly the measured module temperatures. Details on the simulation software can be found in [2].

3.3 Measured data

The data for the present study was gathered by the Photovoltaic Laboratory of the Institute for Energy and Mobility Research at the BFH-TI in Burgdorf, Switzerland. The measured quantities used here are:

- Global horizontal irradiance measured with a pyranometer
- PoA irradiance measured with a pyranometer
- PoA irradiance measured with a reference cell
- Ambient temperature measured with Pt100 sensor
- Reference cell temperature measured with Pt100 sensor
- DC current of PV array
- DC voltage of PV array
- AC power injected into grid

The data is taken every five minutes, and since the simulation proceeds in steps of one hour, mean hourly values were calculated from the raw data. The measurements cover the years 2002 until 2014. The plots in the following show either longterm results for the whole measured time period, or single year results for 2010, which was chosen arbitrarily for this study. More details on the type of measurements along with technical details are described in [3].

3.4 Removing Data Points with Snow Coverage

The PV installation that has been used for this study is located in the Swiss Jura mountains, where snowfall is common during winter. Since the simulation does not consider snow coverage models, it is necessary to remove the data points when the modules are completely or partially covered with snow. As a measure for this, we compare the

PoA irradiance measured by the pyranometer to the DC power produced by the array (figure 3). The pyranometer is heated, and is thus usually free from snow. The points with potential snow cover gather at the bottom of the plot, meaning that PV production is significantly lower than one would expect from the measured irradiance of the pyranometer. We select all points that are within ± 4 standard deviations of a linear regression. This cut is repeated successively, until no more points are excluded. The removed points in figure 3 are highlighted in red, and contain also other outliers that are not related to snow, but rather to measurement errors, electrical failures of the array. With this cut, 5% of the data points are removed from the sample.

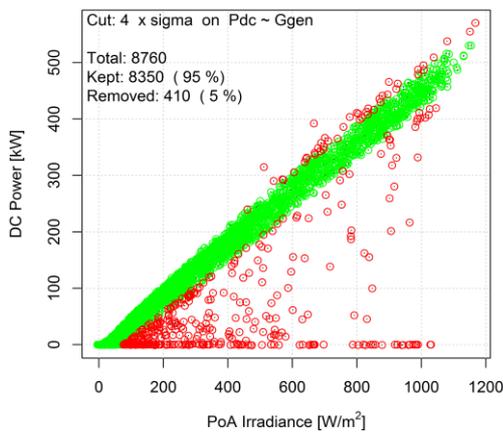


Figure 3: Measured DC production of the array against irradiance from heated pyranometer. The red points are removed from the data set to avoid situations with snow coverage.

3.5 Selecting Clear Sky Data

When analyzing the data, it is often very useful to select data points with no clouds, the clear sky hours. In these conditions one can compute the expected irradiance with a clear sky model, thus obtaining an absolute reference, which does not depend on a measurement. Also the separation of direct and diffuse component have less uncertainties in these conditions. Finally the clear sky conditions can be used to compare monitoring data and simulation for the case that no irradiance measurement is available.

A simple strategy, to select clear sky data, consists in

computing the ratio K_{tcs} between the measured global horizontal irradiance, and the one calculated with the clear sky model. The distribution of the hourly K_{tcs} values for the year 2010 and daylight hours, is shown in figure 4. In this study we apply a cut of $K_{tcs} > 0.9$, to select clear sky values, which will keep around 20% of the hourly values with daylight.

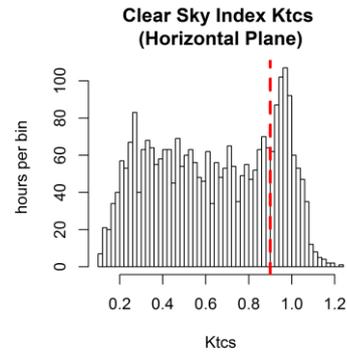


Figure 4: Distribution of hourly K_{tcs} coefficients for the year 2010. Only daylight hours are considered. The dotted line denotes the cut for clear sky conditions.

4 ANALYSIS EXAMPLES

In the following section we present a few analysis examples that were generated with the method described above. Table I summarizes the analysis plots which all follow the logic described in section 2. Plots with results for a single year, are always for 2010.

4.1 Stability of irradiance measurement

In a first example we compare the measured irradiances with the clear sky model from the simulation. The aim of this plot is to examine the stability of the irradiance measurements over the years. The clear sky model is assumed not to change over the years and establishes a constant reference, to which the measurements are compared. To get a meaningful result, it is important to apply the clear sky selection from section 3.5. On top of this, a cut was applied on the time of day, selecting only points from 10:00 AM to 14:00 PM, to avoid distortions coming from large relative errors when measuring small irradiances.

Table I: Overview of the analysis examples.

Analysis Plot	Stability of irradiance measurement	4.1	Ageing of Reference Cell	4.2	Pyranometer Orientation	4.3	Temp. dependence of irradiance measurement	4.4	Fault Detection	4.5	Seasonal variation of DC power simulation	4.6	Simulation of PV production	4.7	Ageing of Installation	4.8
Data Cuts	Clear sky		No snow,		Clear sky		No snow		None		No snow, no faults		No snow, no faults		No snow, no faults	
Variable 1	Measured Irradiance		PoA irradiance from Reference Cell		Measured PoA irradiance		PoA irradiance from Reference Cell		PoA irradiance		Measured DC power		Measured AC power		Measured AC power	
Variable 2	Clear Sky Irradiance		PoA irradiance from pyranometer		Transposed PoA irradiance		PoA irradiance from pyranometer		Measured DC power		Simulated DC power		Simulated AC power		Simulated AC power	
Indicator	Ratio		Ratio		s of residuals		s of residuals		Average ratio		Ratio		Difference		Average ratio	
Control Variable	Year		Year		Tilt/Azimuth		Ambient temperature		Day		Month		Irradiance/ Temperature		Year	

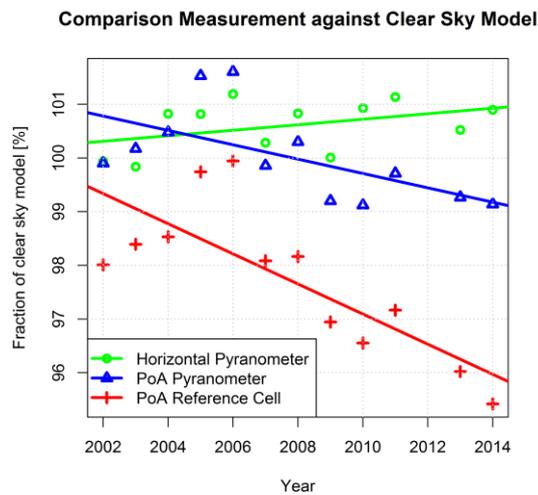


Figure 5: Ratio of irradiance measurement and clear sky model.

In the terminology of section 2, the compared variables are the irradiance measurement and the irradiance of the clear sky model, the indicator being the ratio of the yearly sums of those variables. The control variable is the year. The result of this plot is shown in figure 5 for the horizontal pyranometer, as well as the PoA pyranometer and the reference cell. A clear decrease with time of the PoA measurement responses can be seen. The horizontal measurement is almost stable over time, displaying just a small increase. This plot also shows a strong correlation between all three measurements, especially the two PoA responses follow a common pattern. This pattern could be related to a varying Linke turbidity coefficient, which in this study was taken as constant over all the years. To get rid of this overlaying fluctuations, one can compare the two PoA measurements directly, which is done in the next example.

4.2 Comparison of Pyranometer and Reference Cell

This example is similar to the previous one, but now the comparison is not done between measurement and simulation, but between two measured values: the PoA irradiance as measured by the pyranometer and by the reference cell. As indicator we choose the average ratio between the two values. Figure 6 shows this indicator as function of the year. Since this comparison is not done with the clear sky model, the cut on K_{tes} can be much softer, and it was chosen at 0.4. No cut was applied on the time of the day. From the plot one can see, that the reference cell measures systematically lower values than the pyranometer. After 7-8 years of operation, the ratio between the two values starts to decrease constantly, which is probably due to the ageing of the reference cell. If so, we don't understand well why this has not occurred earlier. The reference cell is made of the same cell type as the power producing panels are. It was chosen to have the same characteristics as the panels, to allow for a representative comparison.

Further detailed studies could be performed, to determine if the fluctuations that overlay the general trends in figure 6. Could be related to the Linke turbidity coefficient.

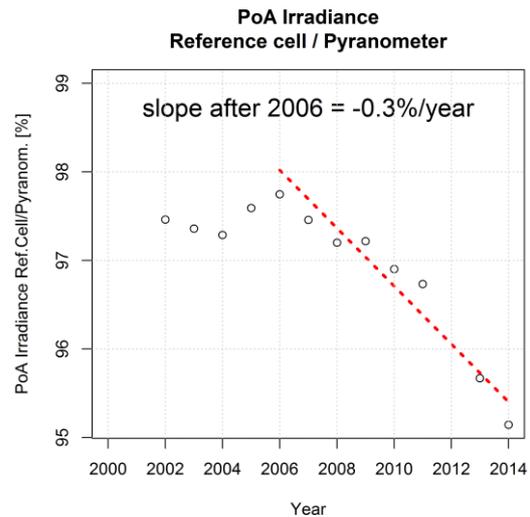


Figure 6: Ratio of irradiance measurement with reference cell and pyranometer.

4.3 Orientation check of the pyranometer

Comparing the transposed horizontal irradiance to the measured irradiance in the collector plane allows to find the orientation of the PoA measuring device. The comparison is done between the PoA irradiance as calculated with the transposition and as measured with the pyranometer. As this check is only performed under clear sky conditions, the transposition is essentially based on the solar geometry (low diffuse), therefore not model-dependent. As indicator, we take the standard deviation of the linear fit residuals. The result is shown in figure 7, where this indicator is plotted against the tilt and azimuth for the transposition. The indicator is minimized at a tilt of 46° and an azimuth of -23° , while the nominal orientation of the pyranometer is at a tilt of 50° and an azimuth of -26° (26° E). A similar check could be performed for the assessment of the true orientation of a full array; but in our case this is not possible because there are 2 different orientations, mixed in the same measurement.

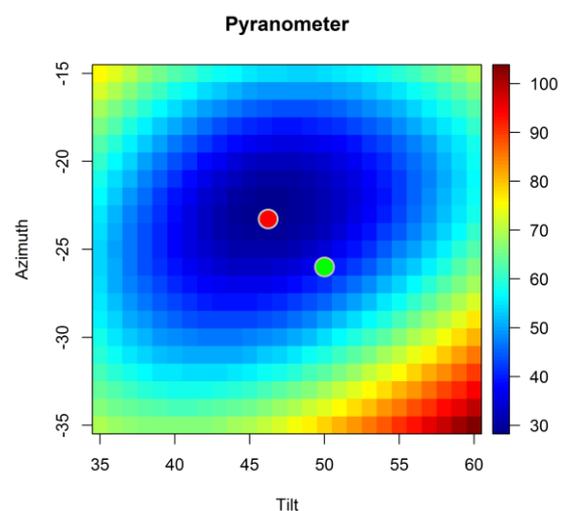


Figure 7: Standard deviation in W/m^2 of the residuals of a linear fit between measured and simulated PoA irradiance. The red point is the fit result, the green one the nominal orientation.

4.4 Temperature dependence of irradiance measurement

It is known, that the irradiance measurements depend on the temperature. The relative temperature dependence of a reference cell and pyranometer is of the order of a few $10^{-4}/^{\circ}\text{C}$, leading to errors around 1% for a ΔT of 30°C . To check the temperature dependence of the irradiance measurement, we compare the two PoA irradiance measurements. As indicator we choose the residuals, and as control variable the ambient temperature. The result is shown in figure 8. The systematic dependence of the irradiance measurements is seen, leading to deviations up to 20 W/m^2 over the whole temperature range. This behavior is a combination of both irradiance measurements, and cannot be assigned unambiguously to either the reference cell or the pyranometer.

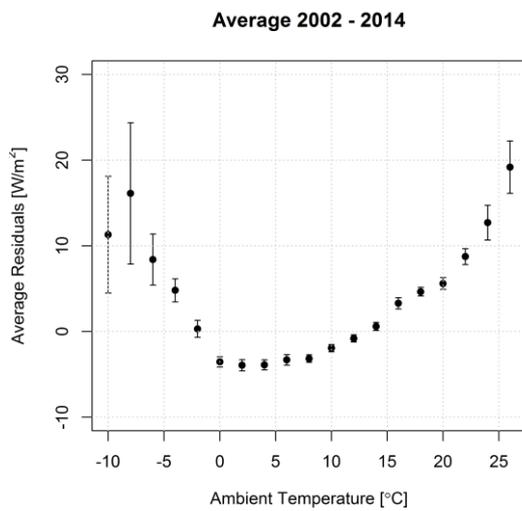


Figure 8: Average residuals from the comparison of the two PoA irradiance measurements. The shape is defined by the temperature behavior of the pyranometer and the reference cell

4.5 Fault detection

A straightforward way of checking for malfunctions is to compare the measured and simulated generated power, and to display the result as a function of time. The simulation here provides a valuable comparison reference, since it approximates the ideal running conditions for any external conditions at every hour. Figure 9 shows an example for the month of July 2010 in daily values. In this period, the logbook of the operator mentions five incidents, where the system was off. These five incidents correspond to the lowest values in the diagram. Even the cases where the system was off for less than one hour are detected in the data.

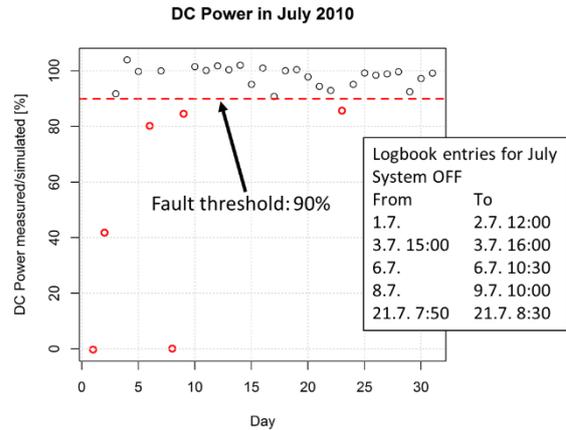


Figure 9: Generated DC power per kWm^{-2} PoA irradiance for the month of July. The events reported in the logbook correspond to the lowest entries in the plot.

4.6 Seasonal dependence of DC power simulation

As a crosscheck for an accurate simulation, one can compare the measured and simulated DC power generation and check if there is a seasonal dependence over the year. The average of this plot over all measured years is shown in figure 10. It can be seen, that the simulation systematically yields smaller values in summer than in winter. This is an indication, that the thermal coefficient U_c , describing the heat transfer between modules and the environment, has not been chosen correctly in the simulation and has to be checked. This kind of plot allows to distinguish the thermal effects, which act differently along the year, from losses that have no seasonal dependence, like ohmic losses or module degradation.

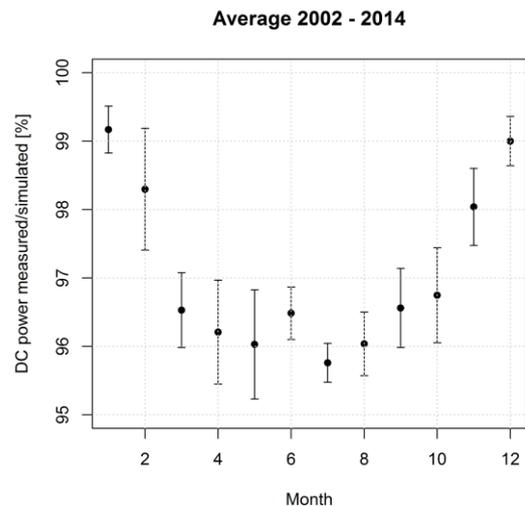


Figure 10: Ratio of measured and simulated DC power as function of the month.

4.7 AC Power injected into Grid

In this example we compare measured and simulated AC production of the PV installation. Figure 11 shows the difference between both in hourly values, plotted against the PoA irradiance and the ambient temperature. The value has been normalized to the nominal power of the PV array, to get a figure that is independent of the system size. The

average difference is at 1%, and that the RMS value is 2.2% of P_{nom} . These figures give a measure of the general quality of the simulation. Both distributions are also flat, meaning that there is no systematic error caused by irradiance or temperature.

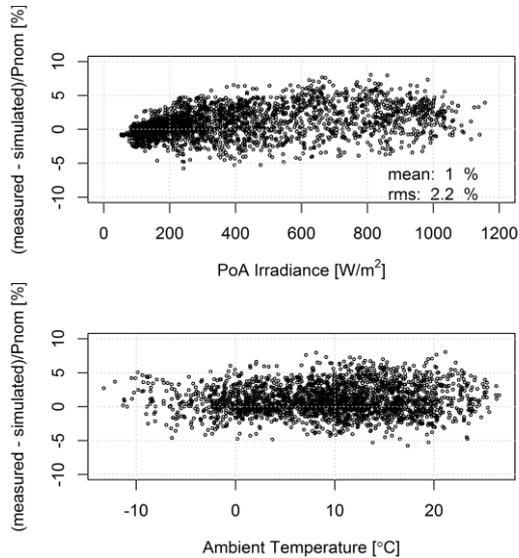


Figure 11: Difference of measured and simulated power injected to the grid.

4.8 Long term AC power generation

The indicator in this example (figure 12), is the average ratio between the measured and simulated AC production. Assuming that the irradiance measurement calibration is stable over time, this gives a measure of the degradation of the PV installation over the years. Only the trend in this plot is meaningful, since the absolute value of the ratio depends on choices made for the simulation that aimed to match the results for the year 2010.

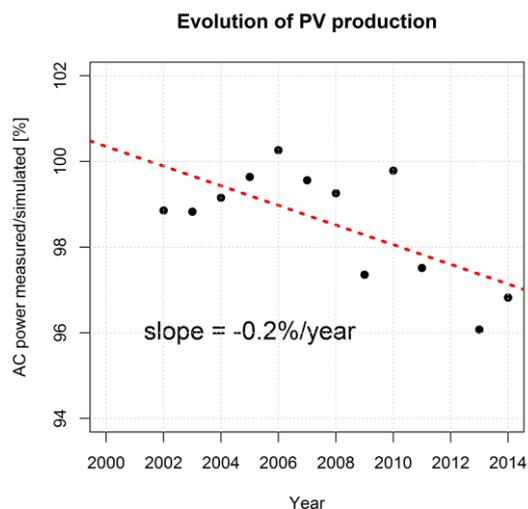


Figure 12: Ratio between measured and simulated system production. The simulation parameters were adapted to best match all measured values of 2010.

5 CONCLUSION

In this study we presented a general approach to compare measured and simulated data of a PV installation. The aim of this method is to validate the measurements, spot faults in the operation of the installation, detect possible problems in the simulation parameters and get an overview of the system performance.

For the analysis plots, measured and simulated values are compared. The indicators that are used to describe the matching of the two, can be the differences, the ratios or the residuals of a linear fit. These indicators are plotted against time to detect faults, or to display the long term evolution of measurements and system performance. The indicators can also be plotted against other control variables like the ambient temperature or the irradiance, to reveal systematic errors.

In section 4 we showed several examples of analysis plots that demonstrate the strength of this approach in detecting faults and systematic errors and to provide concise overviews of the long term evolution of the installation.

One important aspect in this approach is that it operates on hourly values to compute differences and RMS values, which makes it possible to extract dependences that would get lost when comparing yearly or monthly averages.

The method that is described here, will be used in a tool that is being implemented in the PVsyst software. The aim of this tool is to simplify the comparison of monitoring data to simulation results.

5 ACKNOWLEDGMENTS

The authors would like to thank Prof. Urs Muntwyler and Thomas Schott from the Bern University of Applied Sciences, for providing the measured data that was used in this study.

6 REFERENCES

- [1] BFH Website, Montsoleil PV installation
<http://www.pvtest.ch/index.php?id=403&L=2>
- [2] PVsyst Website
<http://www.pvsyst.com/en/>
- [3] BFH Website, Fine measurement principle
<http://www.pvtest.ch/index.php?id=2667&L=2>