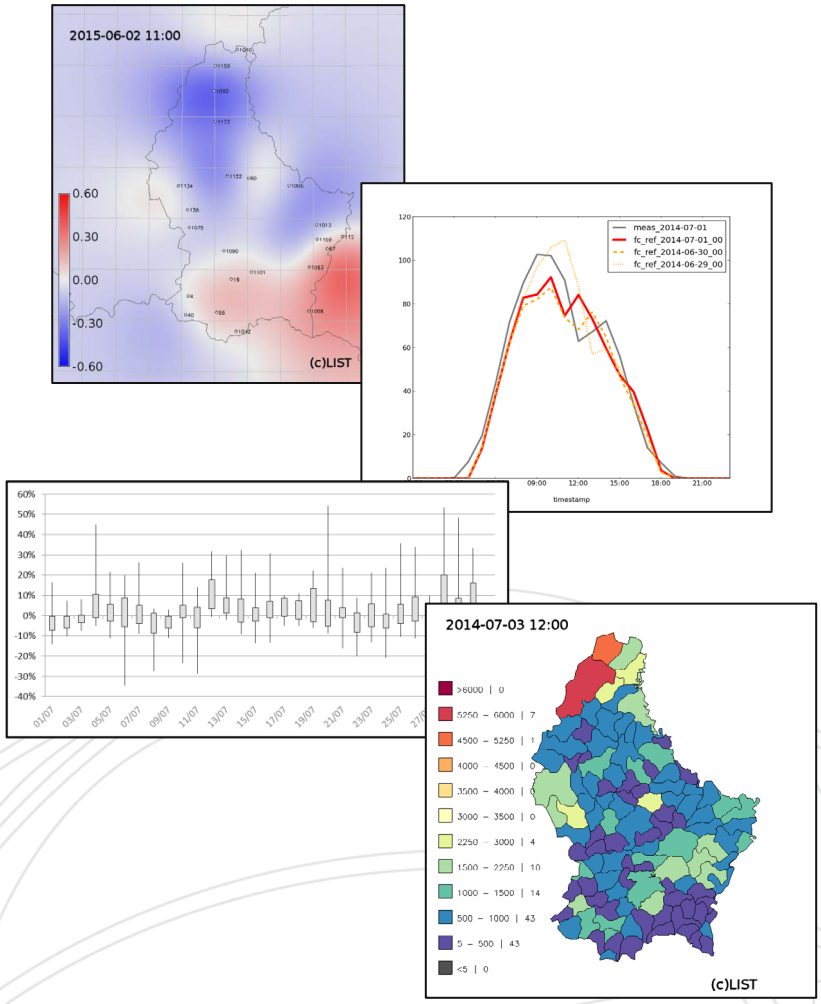
LUXEMBOURG
INSTITUTE
OF SCIENCE
AND TECHNOLOGY

PV-FORECAST



FINAL REPORT

Full Titel: **Development of a dynamic, self-adapting photovoltaic power forecasting model**

Subject: **Final report**

Date: **10.08.2016**

Luxembourg Institute of Science and Technology (LIST) ERIN Department
5, avenue des HautsFourneaux | L-4362 Esch/Alzette
Tel: (+352) 275 888 – 1 | LIST.lu

Contents

Table of images

Abbreviations

AC	alternate current
AEV	Administration de l'Environnement
CMV	cloud-motion vector
csv	comma separated values files
DC	direct current
dt	day time values
ECMWF	European Center for Medium-ranged Weather Forecasts
ERIN	Environmental Research and Innovation
GHI	global horizontal irradiation
GIS	geographical information system
kWp	kilo Watt peak – nominal power of the photovoltaic module
LIST	Luxembourg Institute of Science and Technology
NWP	numerical weather prediction
PV	Photovoltaic
ssrd	surface solar radiation downwards

Summary

In view of European and national climate protection objectives and the depletion of fossil energy resources, the shares of electricity from decentralized, renewable energy sources are rising. The reliable management of our electricity supply and grids as well as the containment of increasing price volatility on the electricity market, will depend on the ability to handle fluctuating renewable sources, such as wind and solar. The forecasting of the dynamics of photovoltaic (PV) power production is therefore crucial for the integration of high shares of photovoltaic into our energy system and market.

Within the PV-Forecast project, the LIST developed a forecasting model for Luxemburg, able to predict the expected regional PV power up to 72 h ahead. The model works with solar irradiance forecasts, based on numerical weather predictions from the European Center for Medium-Range Weather Forecasts (ECMWF) in hourly resolution and a spatial resolution of 0.125° x 0.125°.

Using a set of physical, technical models, the algorithm is able to predict the expected hourly power production of a collection of PV reference systems, distributed over Luxembourg. For development purposes, power forecasts have been calculated based on historical irradiance forecast data (for 2014 and 2015) and have been compared to power measurements of these systems, available for the same period. In order to reduce systematic errors, the model for each reference system has been calibrated, which led on average to a reduction of the error by -2.66% (at a mean deviation before the calibration of -5.96 %).

The performance of the forecasting algorithm on level of the reference systems has been evaluated, using commonly known evaluation criteria, also described in this report. The mean bias of the forecasts for all reference systems over the two years is low, 1.1% (low bias is an indicator of low systematic error). The root mean square error (RMSE) in an indicator of the performance on hourly basis, which gives higher weighting to larger deviations from the real value and is considered a suitable indicator in power forecasts. The average RMSE lies around 7.4% - a good value for single site forecasts.

The approach behind the PV-Forecast project was to combine the purely model based predictions, as presented above, with a feed-back loop from the measurements of reference systems, in order to adapt the short term forecasts. To this aim, two different approaches have been tested.
The approach of “error persistence” is based on the assumption that a deviation of the power forecast for a time step would (with a certain probability) persist to the next time step, or even a few time steps ahead. Therefore, the forecasts were adapted, based on the error of the previous hour (or 2 – 3 hours before). The results show a reduction in bias, as well as RMSE for the “1h error persistence approach”, here for reference system nr. 80: bias dropped from 1% to 0.14%; RMSE improved from 6.57% to 5.81%. For the “2h error persistence approach”, the bias improved to 0.4%, but RMSE did not generally improve. It became clear, that an error persistence based forecast adaptation might improve the forecasts only in a very short forecast horizon (in our case 1h), but we proposed a further improvement of the approach which seems to be promising.
The second approach is based on the assumption that inaccuracies in irradiance forecasts are to a large extended due to inaccurate forecasting of cloud movements (in direction and speed). Based on the power forecast deviations of previous time steps, error maps were plotted, visualizing the regional distribution of errors. Based on sequences of error maps from the previous time steps, it was tested if errors propagate over the forecast region constantly in speed and direction, due to inaccurate cloud movement predictions. If so, their movement could be predicted and the forecast would be adapted. The analysis did show only few days during the period of two years, which would fit the approach. This leads also to the thesis, that this approach could be more promising at higher time resolutions for the irradiance forecast (e.g. 15 min).

Finally, the developed forecasting model does show good performance for power predictions of the reference systems and is able to upscale these predictions to the national level. The model currently predicts up to 72 h ahead, updated once per day, hourly expected PV power productions per commune.

1 Introduction

1.1 Background & Objective

The share of decentralized and fluctuating energy sources, such as wind power and photovoltaic (PV), is constantly increasing and will represent a major part of the future energy mix. This will bring new challenges for grid operators and energy provider and will ask for new approaches to feed-in and store energy. Precise predictions of the energy production from these sources will be crucial to ensure grid stability and will be essential information for the trading of energy on the energy market.
The focus for prediction models that serve this purpose is the challenge is to give a precise prediction of the dynamics of the energy production of the whole set of PV systems in an area, thus the temporal feed-in profile from PV within a specified region (e.g. a balancing group, a city, a region).

The idea behind this research project was to investigate the possibilities for precise PV power predictions for a region by a combined approach of predictive modeling of the PV power and the monitoring of reference systems.

The specific objectives of the project were defined as follows:

1) Establish a photovoltaic power forecasting model for Luxemburg
2) Prove whether forecasting accuracy (and spatial resolution) of a photovoltaic forecasting model could be increased by combining the modeling approach with online measurements of reference systems
3) Identify possible and optimized forecasting time horizons

Beside the scientific findings, the project delivers a PV power prediction model for Luxemburg, adapted to the regional situation, which could be implemented as an online, real-time tool in a next development step. The project can thereby contribute largely to the future integration of PV into the local power grid and energy mix.

1.2 Description of the approach

The approach of the PV power forecast model is a combination of geo-referenced irradiance and ambient temperature forecast data from numerical weather prediction (NWP) models of the European Centre for Medium-Range Weather Forecast (ECMWF) with measurement data of reference PV systems distributed over the region (Figure 1). The irradiance forecast data, which is being retrieved from the ECMWF web servers, is processed in order to obtain the irradiance in plane of the PV modules. This is done for a number of given PV systems that serve as references and for a matrix of predefined orientations and inclinations representing the whole set of PV systems in the country. The reference PV systems, of which measured PV power in a temporal resolution of 15 minutes is available, are distributed over the whole country of Luxembourg. Based on a general model and individual system profiles for each reference system representing their technical characteristics, a power forecast is being generated. The predicted power of the reference systems is compared to their measured generated power, with the aim to set up a feed-back loop that enables the adaptation of the short term forecasts, based on prediction errors of previous time steps.

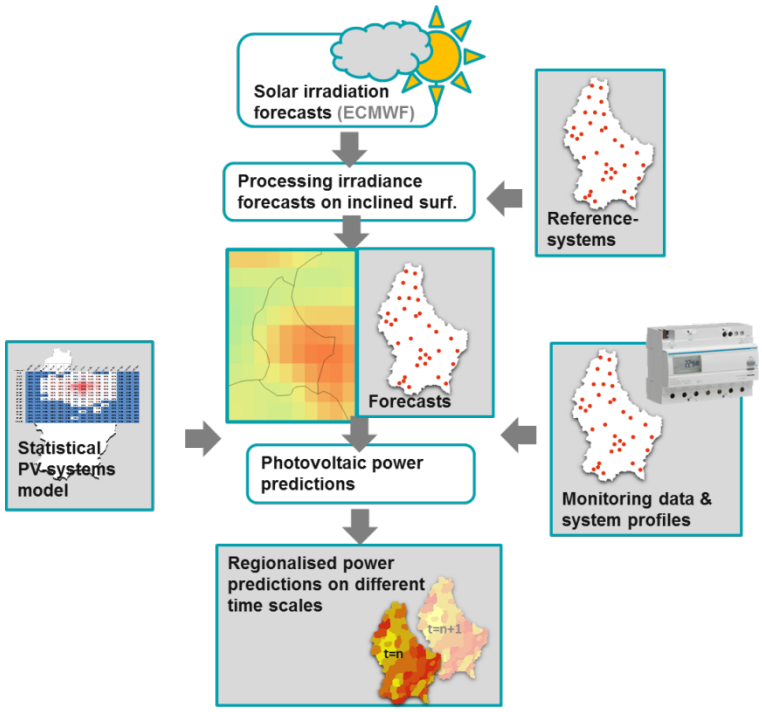


Figure 1 - scheme of the forecasting approach, combining modeling data and statistical information (left hand side) with a feed-back loop from PV reference systems (right hand side)

A statistical model of the spatially distributed PV installations in the whole country is then used in order to model the generated power for three days on an hourly basis over the whole region. The spatial resolution of the ECMWF irradiance forecast data is 0.125° x 0.125°, but the PV power forecast are based on individual forecasts for the whole set of PV systems in Luxembourg and could thus be aggregated on different levels (currently done on communal level).

2 Description of the forecasting model

2.1 Irradiance forecasts from ECMWF (G_H)

Irradiance forecast data are automatically extracted from the European Centre for Medium Ranged Weather Forecasts (ECMWF), London. The used parameter is “ssrd”, which stands for surface solar radiation downwards and can be, to a reasonably good approximation, considered to be what would be measured by a global pyranometer at the surface (Hogan 2015). This value represents the global horizontal irradiance, hence the sum of direct and diffuse irradiance.

$$G_H = \text{global horizontal irradiation} \qquad [\mathrm{Wh/m^2}]$$

The currently available spatial resolution of the forecast data is 0.125 ° x 0.125 °, which corresponds to approximately 9 km in width and 14 km in height.

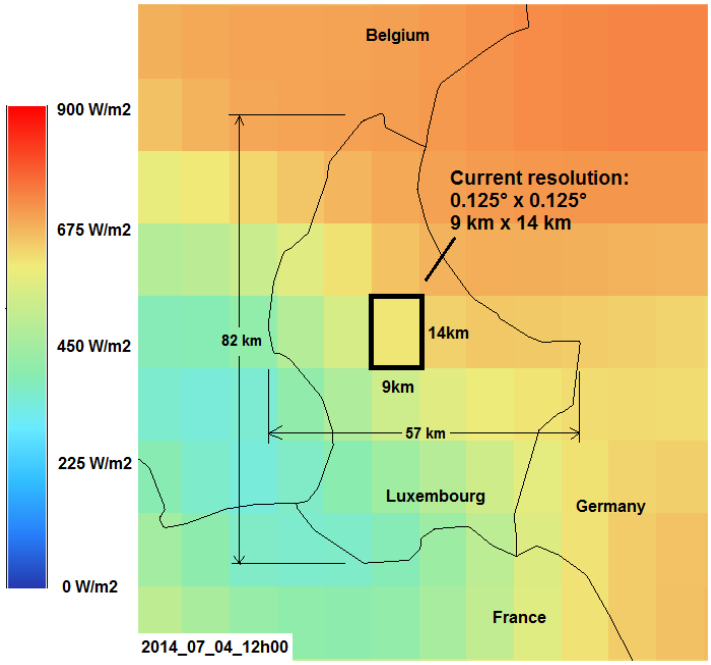


Figure 2 - Current available resolution for forecasting data from ECMWF compared to the size of Luxembourg

Currently, these irradiance forecasts at ECMWF are being published twice a day, starting from 00:00 and 12:00, and are available with a 8 hours delay only, which is partially due to calculation time. Therefore, the intra-day part of the forecast starting at 12:00 o’clock, being available from 20:00 on only, is of no practical usage for solar power forecasting. In the following, it has been decided to only use the forecasts starting at 00:00, available from 08:00 o’clock each day. Each forecast covers a range of 72 hours ahead in 1 hour resolution. Unfortunately, 15 min. time steps are not available, which would be more suitable to power forecast applications. If only the 00:00 o’clock forecast is being used, this results in 3 forecast horizons covering each day:

- an intra-day forecast 0-24h, published the same day (08:00)
- a day-ahead forecast 24-48h, published the previous day (08:00)
- a 2-days-ahead forecast 48-72h, published 2 days before (08:00)

For the development and research on the forecasting algorithm this delay is not of importance as we work with historical forecasts, but once the method should be used, it needs to be checked if either earlier forecasts are available for paying customers, or if other forecasting services provide faster forecasts.

2.2 Techno-physical model

2.2.1 Global irradiance on inclined surfaces (G_ψ)

The global horizontal irradiance is the sum of the direct or beam irradiance, reaching the horizontal surface in direction from the sun, and the diffuse irradiance, being the reflected part coming from the surrounding environment (clouds, mountains etc.). In order to be able to calculate the irradiance reaching a tilted, oriented surface (such as PV modules), information on the direct and diffuse solar irradiance is necessary, but often not available, if not measured. There are different methods for the calculation of the inclined global irradiation depending on the quantity and quality of available parameters. A rather simple method with a relatively good accuracy was published in 1998 by (Olmo, et al. 1999). A recent study from (Khalil et Shaffie 2013) points out the simplicity and accuracy of the method developed by Olmo et al. The advantage of this method is that it requires only the global horizontal irradiation GHI, the sun position (elevation to the horizon and incident angle to the inclined surface) and the albedo of the environment. During the development phase of this project, it has been decided to use this approach described below to estimate the irradiance on inclined surfaces and the approach has been validated using our own measurement data from LIST's PV test field and meteo station (PV-Lab) (see 3.1).

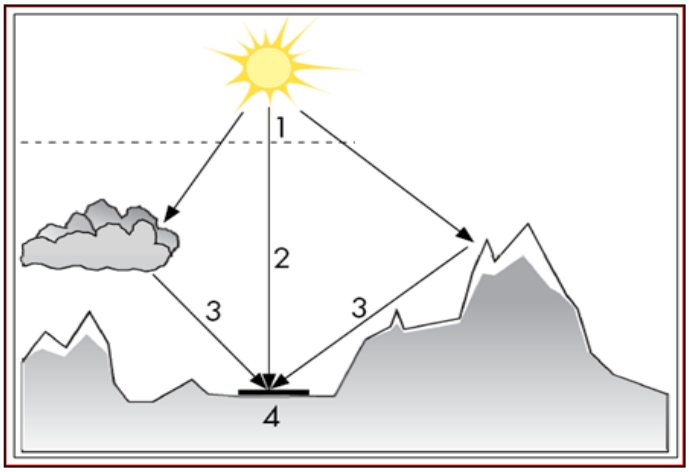


Figure 3 – scheme showing direct irradiance outside the atmosphere (1), direct irradiance at the surface (2), diffuse irradiance (3) and global irradiance (4)

Global irradiance G_ψ on an inclined surface can be calculated using the following formula from (Olmo, et al. 1999).

$$G_\psi = G_H^{(-k_t(\psi^2-\psi_H^2)} F_c \qquad (1)$$

k_t = Clearness index	[/]
ψ = incident angle	[rad]
ψ_H = elevation angle	[rad]
F_c = ground reflected radiation	[/]

The incident angle ψ is the angular distance subtended by the normal to the inclined plane and the sun. The elevation angle ψ_H is the angular distance between the normal direction to the horizontal plane and the sun's position. Often, the elevation angle is measured from the horizontal plane to the suns direction, but here it is measured from the normal of the horizontal plane to the suns direction.
The clearness index k_t is a function of the global horizontal irradiance G_H and the extra-terrestrial solar radiation I_o (depending on the distance to the sun).

$$k_t = \frac{G_H}{I_o} \qquad (2)$$

I_o = extra-terrestrial solar radiation	[W/m_2]

The extra-terrestrial solar radiation I_o can be calculated based on the distance to the sun, which is changing over the year, but can be estimated depending on the day of the year using the following functions.

$$I_o = 1367 * (\frac{R_{av}}{R})^2 \quad (3)$$

Where R_{av} is the mean sun-earth distance and R is the current sun-earth distance depending on the day of the year.

$$(\frac{R_{av}}{R})^2 = 1.00011 + 0.034221 * \cos\beta + 0.00128 * \sin\beta + 0.000719 * \cos(2\beta) + 0.000077 * \sin(2\beta) \quad (4)$$

$$\beta = \frac{2\pi\, n}{365} \quad (5)$$

$n = \text{day of the year}$

Ground reflections F_c are determining the Albedo caused fortification of the global irradiation on the inclined surface.

$$F_c = 1 + \rho * \sin^2(\frac{\psi}{2}) \quad (6)$$

ρ = Albedo [/]

The Albedo factor can be taken from literature, taking into account the amount of reflections from the ground before the inclined surface. In our current model a standard value is used:

ρ = 0,35 [/]

2.2.2 Reflection losses on module surface

In a first attempt, the "ASHRAE incidence modifier" method has been used to consider reflection losses on the PV modules surface. But the known weakness of this model, the false representation of reflection losses at high incident angles, has proven to be problematic for the forecasting performance. Therefore the following physical model for the incident angle modifier was implemented:

$$I\,AM_B = \left(\frac{\tau(\psi)}{\tau(0)}\right) \quad (7)$$

$\tau = \text{transmittance}$
$\psi = \text{incident angle}$
$\tau(0) = \text{transmittance when normal to the sun}$
$\tau(\psi) = \text{transmittance at incidence angle}$

$$\tau(\psi) = e^{-(\frac{KL}{\cos(\psi_r)})}\left[1 - \frac{1}{2}\left(\frac{sin^2(\psi_r - \psi)}{sin^2(\psi_r + \psi)}\right) + \left(\frac{tan^2(\psi_r - \psi)}{tan^2(\psi_r + \psi)}\right)\right] \quad (8)$$

$K = \text{glazing extinction coefficient, typical in PV:}$	4	$[m^{-1}]$
$L = \text{glazing thickness, typical in PV:}$	0.002	[m]
$n = \text{index of refraction of the cover glass, typically for PV:}$	1.526	[/]
$\psi_r = \text{refraction angle}$		

Standard values used here were taken from literature of (De Soto, Klein et Beckman 2006)

$$\psi_r = \sin^{-1}\left(\frac{1}{n}\sin(\psi)\right) \quad (9)$$

$$\tau(0) = \exp(-KL)\left[1 - \left(\frac{1-n}{1+n}\right)^2\right] \quad (10)$$

The actual losses are then:

$$\text{Reflection losses} = 1 - \text{I AM}_B \quad [\%]$$

2.2.3 PV-module performance at operating conditions

In the following, PV modules performance can be described as a function of the global irradiance in plane of the module G_ψ and the module operation temperature T_{module}.
T_{module} can be approximated by (Lorenz, et al. 2011):

$$T_{module} = T_a + G_\psi * \gamma \quad (11)$$

T_a = ambient temperature [°C]
$\gamma = 0.02$ (free standing) or 0.056 (BIPV) [/]

This simplified approach could alternatively be replaced by the "Sandia Module Temperature Model" or the "Faiman Module Temperature Model". So far, the PV-Forecast algorithm works well with the approximation above.

The modules efficiency at operation temperature η_{Tm} can be calculated with the following simplified approach:

$$\eta_{Tm} = \eta_{STC} * (1 - (-\frac{K_T}{P_{MPP}} * (T_{module} - 25°C)) \quad (12)$$

η_{Tm} = efficiency at operation temperature [%]
η_{STC} = efficiency at standard test conditions [%]
K_T = temperature coefficient for module power [W/°K]
P_{MPP} = nominal power of the module [W]

2.2.4 Effective PV power per surface area

Resulting on all the previous calculation steps, the effective photovoltaic power per m^2 can be calculated:

$$P_{eff\,A} = G_\psi * I\,AM_B * \eta_{Tm} \quad (13)$$

$P_{eff\,A}$ = effective power of PV modules per m^2 [W/m^2]

On level of the PV module, the effective power is then calculated based on the dimensions of the module.

$$P_{eff\,mod} = P_{eff\,A} * w * l \quad (14)$$

$P_{eff\,mod}$ = effective power on level of PV module [W]
w * l = width and length of the PV module [m * m]

2.2.5 Module performance degradation and mismatch effect

In its current version of the model, performance degradation over time and mismatch effects are taken into account by a lump sum factor. The following factors are considered:

$c_{degr\,1st}$ = degradation loss factor 1st year= 2.5 [%]
$c_{degr\,f}$ = degradation loss factor following years= 0.5 [%]

c_{mm} = miss match loss factor = 2.5 [%]

Degradation loss factors are considered different for the first and the following years. 2.5% losses are taken into account during the first year, followed by 0.5% per year for the subsequent years. In the future it might make sense to distinguish degradation loss factors depending on the cell technology, but up to now, all the reference systems are crystal silicon technologies.

2.2.6 PV module power and part load operation mode

The resulting power on level of the PV module is then calculated by:

$$P_{mod} = P_{eff\ mod} * \left(1 - c_{degr\ 1st} - \left(n_t - (n_c + 1)\right) * c_{degr\ f} - c_{mm}\right) \quad (15)$$

P_{mod} = power on level of the PV module [W]
n_c = year of construction [/]
n_t = year of today (day of forecast) [/]

Some of the following calculation steps require to know at which share of the nominal power the PV modules are currently being operated. This is defined as part load operation mode:

$$m_{part\ load} = \frac{P_{mod}}{P_{MPP}} \quad (16)$$

$m_{part\ load}$ = part load operation mode [%]
P_{MPP} = nominal power of the PV module [W]

2.2.7 Wiring losses – direct current cabling

Wiring losses in between the PV module strings and the inverter, hence on direct current level (DC), are calculated based on cable sections and cable lengths for the PV arrays nominal power (MPP). This value [%] is considered when the PV array is operated at MPP, while part load behavior is taken into account with this simplified approach:

$$c_{DCwire} = (m_{part\ load})^2 * c_{DC\ MPP} \quad (17)$$

c_{DCwire} = factor DC losses at operating conditions [%]
$c_{DC\ MPP}$ = factor DC losses at MPP [%]
$m_{part\ load}$ = part load operation mode [%]

2.2.8 Inverter losses

Inverter efficiency is also changing with the current part load mode of operation. The reference systems description contains characteristic points of the efficiency curve of the inverter. Depending on the part load operation mode, the part load efficiency of the inverter can be interpolated.

$$\eta_{inv} = f(m_{part\ load}) \qquad (18)$$

η_{inv} = part load efficiency of the inverter [%]

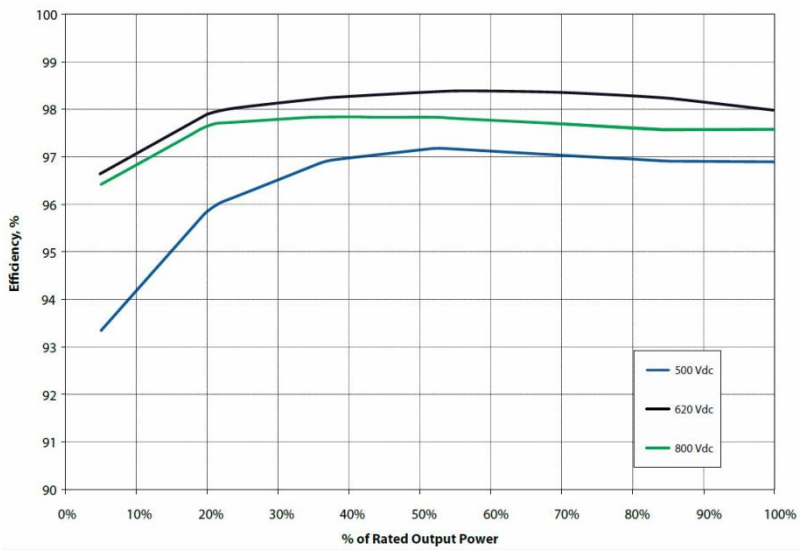


Figure 4 – example for an inverter efficiency curve of a "power-one AURORA TRIO-27.6-TL-OUTD" (power one s.d.)

Following table is an example for the characteristic points of the inverters efficiency curve:

$m_{part\ load}$	5%	10%	20%	40%	60%	80%	100%
η_{inv}	96.7%	97.1%	97.9%	98.3%	98.4%	98.3%	98.0%

Table 1 - efficiency curve of an inverter with a European inverter efficiency of 98%

2.2.9 PV power per inverter

The resulting power on level of the inverter is then calculated by:

$$P_{inv} = P_{mod} * n_{mod} * (1 - c_{DCwire}) \qquad (19)$$

2.2.10 Wiring losses – alternate current cabling

Wiring losses after the inverter, hence on alternate current level (AC), until the point of injection, are consider the same way as the DC wiring losses:

$$c_{ACwire} = (m_{part\ load})^2 * c_{AC\ MPP} \qquad (20)$$

c_{ACwire} = factor DC losses at operating conditions [%]
$c_{AC\ MPP}$ = factor DC losses at MPP [%]

2.2.11 Aggregation of subsystems (per inverter) to reference system level

The model calculates the expected PV-power per Inverter (P_{inv}) for each subsystem of a reference system. The reference systems are divided into subsystems, representing a number of inverter setups of the same type. Inverter setups may vary, taking into account several parameter:

- Varying amount of PV modules per inverter
- Different Inverter type
- Different orientations and/or inclinations of parts of the PV generator
- Different PV modules

The resulting PV power per reference system is the sum of each of its subsystems.

2.2.12 Simplified model for the calculation with synthetic system profiles

Due to the difficulties described under 2.4.1.3, it was not always possible to get all data on the reference systems necessary to use the above described approach. Therefore a simplified procedure has been developed using standard values and publically available information to estimate the expected performance of the reference systems. These PV reference systems are called synthetic system profiles in the following chapters.

Orientation and inclination of the synthetic profiles of the PV reference systems were estimated using different publically available data sources, described under 2.4.2. Therefore, certain parts of the calculations for the synthetic profiles can be done accordingly to the above described approach, e.g. irradiance in plane of the modules, module temperature, degradation, mismatch and reflection losses.

The whole calculation of the PV systems performance is done for an artificial PV module of 1 kW_p power with the following standard values:

parameter	Value [unit]	description
η_{STC}	17 [%]	efficiency at standard test conditions
K_T	-0.5 [W/°K]	temperature coefficient for module power (for 1kW_p P_{MPP})
A_{kWp}	6.7 [m^2/kW_p]	In contrast to the detailed model, for the synthetic profiles calculation are done per kW_p of a mean system. Based on mean efficiencies of common PV modules, this surface demand per kW_p was estimated. w = l = 2.588 [m]
P_{MPP}	1 [kW_p]	PV module power and part load operation mode follows the approach described above, using an artificial PV power of 1 kW_p
$c_{DC\ MPP}$ = $c_{AC\ MPP}$	0.5 [%]	Cabling losses, AC as well as DC
η_{inv}		Inverter losses are considered using the values given in Table 1
P_{inv}	1 [/]	PV power after the inverter P_{inv} follows standard approach using a single artificial module.

Table 2 – standard values used in the synthetic PV reference system profiles

2.3 Statistical part of the model – representing PV systems in Luxembourg

In order to be able to upscale the PV power forecasts on the whole territory of Luxembourg, statistical information on the entirety of PV systems in the country are necessary. As the forecasts should be regionalized and as irradiance conditions vary over the whole forecasting area, the nominal power and location of the individual PV systems are needed (at least statistically). Further, for dynamic forecasting orientation and inclination of the installations are of importance: e.g. a PV installation, oriented towards east starts producing earlier in the day (morning), compared to an installation oriented west (which delivers less in the morning but more in the afternoon). Depending on the inclination of the PV modules, a PV installation with a different orientation delivers the maximum power at another time of the day/year (when the sun is most perpendicular to the surface of the PV module).

The statistical model (see Figure 5) takes into account these specificities of the PV installations in Luxembourg and is built-up by:

- nominal installed power [kWp] in this category
- orientation [°] of the PV modules
- inclination [°] of the PV modules
- cell technology (information and matrix available, but currently not used in the model)
  - mono-crystalline silicon
  - poly-crystalline silicon
  - other technologies

elev/orient	-180°=-150°	-150°=-120°	-120°=-90°	-90°=-60°	-60°=-45°	-45°=-30°	-30°=-15°	-15°=15°	15°=30°	30°=45°	45°=60°	60°=90°	90°=120°	120°=150°	150°=180°
0°=5°	0	0	0	0	6	3	30	36	0	0	0	0	0	0	0
5°=10°	0	0	0	11	0	31	34	238	109	50	0	40	0	0	0
10°=15°	0	0	0	0	38	38	104	614	43	137	87	0	0	0	0
15°=20°	0	0	0	137	139	351	534	1'268	628	704	202	153	18	0	0
20°=25°	0	0	0	96	272	532	2'235	3'077	1'389	1'113	351	174	96	0	0
25°=30°	0	0	0	55	144	542	1'730	2'922	1'717	859	295	238	36	0	0
30°=35°	0	0	0	108	79	543	1'271	3'685	1'358	784	257	201	19	0	0
35°=40°	0	0	0	72	118	283	530	1'750	382	283	297	226	79	0	0
40°=45°	0	0	0	23	40	55	176	553	151	121	74	56	13	0	0
45°=50°	0	0	0	11	0	43	39	234	54	76	61	18	0	0	0
50°=55°	0	0	0	0	0	0	16	17	0	0	0	0	0	0	0
55°=60°	0	0	0	0	0	0	4	0	0	19	4	0	0	0	0
60°=70°	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0
70°=80°	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
80°=90°	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 5 – Distribution of the nominal power in the AEV dataset [kW_p] by orientation & inclination of PV installations in Luxembourg

All these anonymized compiled data have their origin in two data sources: AEV (Administration de l'Environnement – Environmental Agency) and the grid operators Creos (most important electrical grid operator in Luxembourg) and Sudstroum (grid operator in Esch-sur-Alzette).

2.3.1 Statistical data on inclination and orientation of PV systems

Specific data on orientation and inclination of individual PV systems is currently not registered, neither by utility companies / energy providers nor by grid operators. This means that precise data on the entirety of installations in the country is not available.

Nevertheless, the Administration de l'Environnement (AEV) has a dataset on subsidized photovoltaic installations in Luxembourg, according to the applicable regulation. Although, these regulations changed several times during the years, the AEV dataset contains mainly data from small PV installations in Luxembourg, due to limitations on the maximum eligible nominal power in the subsidy regulation.
The used dataset from AEV contains basically data from 7'242 subsidy requests corresponding to 42'720 kW_p, installed in the years 2001-2012 – anonymized and without locations. A subsidy request is not necessarily equal to a PV system, as these could be split up on several owners. These data were extracted from subsidy request forms and were partial differently interpreted by the installers and

therefore contained some obviously misinterpreted data. The whole data set required a manual check and cleaning, which resulted in a remaining dataset of:

- 6'307 subsidy requests (87.1% of 7'242)
- total nominal power of 37'857.91 kW_p (88.6% of 43 MW_p before cleaning) with
  - 17'455 kW_p of mono-crystalline PV modules
  - 19'751 kW_p poly-crystalline modules and
  - 651 kW_p of other types (expected to be mainly amorphous) equal to 1.7%

The "average" orientation is 3.2° (nearly south) and the "average" elevation is 26.6°. With a total nominal power in the cleaned dataset of 37,9 MW_p, compared to a total installed power in Luxembourg of 116.2 MW_p, at the end of 2015 (ILR s.d.) the dataset represents 32.6 % of the installed power. As the sample size is high (6'307 subsidy requests) this can be considered statistically robust.

The use of this dataset, which is without alternative, implies two assumptions:
First, it is assumed that the distribution of orientations and inclinations across the country is equal, as this is not regionally distinguished. Second, orientations and inclinations of systems outside the AEV data set (not subsidized) are similarly distributed. Again, considering the relatively high sample number, these are considered valid assumptions.

Recently, a new AEV dataset from January 2016 with data from 2001 – 2014 covering 8'669 subsidy requests was received but not yet not used. A file describing the procedure of cleaning the original data is available.

2.3.2 Statistical distribution matrix on inclination and orientation

With all the 6'307 subsidy requests out of the cleaned AEV dataset, statistical distribution matrix (Figure 5 and Figure 6) with 225 different possible combinations of orientations and inclinations was generated: 15 different orientations (South = 0°) and the 15 different elevations (by steps of 5° or 10°) Figure 6.

elev/orient	-180°=-150°	-150°=-120°	-120°=-90°	-90°=-60°	-60°=-45°	-45°=-30°	-30°=-15°	-15°=15°	15°=30°	30°=45°	45°=60°	60°=90°	90°=120°	120°=150°	150°=180°
0°=5°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5°=10°	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.6%	0.3%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
10°=15°	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.6%	0.1%	0.4%	0.2%	0.0%	0.0%	0.0%	0.0%
15°=20°	0.0%	0.0%	0.0%	0.4%	0.4%	0.9%	1.4%	3.3%	1.7%	1.9%	0.5%	0.4%	0.0%	0.0%	0.0%
20°=25°	0.0%	0.0%	0.0%	0.3%	0.7%	1.4%	5.9%	8.1%	3.7%	2.9%	0.9%	0.5%	0.3%	0.0%	0.0%
25°=30°	0.0%	0.0%	0.0%	0.1%	0.4%	1.4%	4.6%	7.7%	4.5%	2.3%	0.8%	0.6%	0.1%	0.0%	0.0%
30°=35°	0.0%	0.0%	0.0%	0.3%	0.2%	1.4%	3.4%	9.7%	3.6%	2.1%	0.7%	0.5%	0.1%	0.0%	0.0%
35°=40°	0.0%	0.0%	0.0%	0.2%	0.3%	0.7%	1.4%	4.6%	1.0%	0.7%	0.8%	0.6%	0.2%	0.0%	0.0%
40°=45°	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.5%	1.5%	0.4%	0.3%	0.2%	0.1%	0.0%	0.0%	0.0%
45°=50°	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.6%	0.1%	0.2%	0.2%	0.0%	0.0%	0.0%	0.0%
50°=55°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
55°=60°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
60°=70°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
70°=80°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
80°=90°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Figure 6 – PV power distribution [%] by orientation & inclination of PV installations in Luxembourg

It can be seen, that most of the PV installations are south-oriented and have a horizontal elevation of around 30°. For 146 cells (value: 0.0%, colour: dark blue in Figure 6 and greyed out in Figure 7) out of the 225 cells, no corresponding data exists in the dataset.

In order to reduce the calculation time, all those constellations from the remaining 79 combinations were chosen, which represent more than or equal to 0.2% of the power. This selection results in 57 possible combinations and corresponds to 97% of all the samples in the dataset (see figure below):

elev/orient	-180°–150°	-150°–120°	-120°–90°	-90°–60°	-60°–45°	-45°–30°	-30°–15°	-15°–15°	15°–30°	30°–45°	45°–60°	60°–90°	90°–120°	120°–150°	150°–180°
0°–5°	0.00%	0.00%	0.00%	0.00%	0.02%	0.01%	0.08%	0.10%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
5°–10°	0.00%	0.00%	0.00%	0.03%	0.00%	0.08%	0.09%	0.63%	0.29%	0.13%	0.00%	0.11%	0.00%	0.00%	0.00%
10°–15°	0.00%	0.00%	0.00%	0.00%	0.10%	0.10%	0.27%	1.62%	0.11%	0.36%	0.23%	0.00%	0.00%	0.00%	0.00%
15°–20°	0.00%	0.00%	0.00%	0.36%	0.37%	0.93%	1.41%	3.35%	1.66%	1.86%	0.53%	0.40%	0.05%	0.00%	0.00%
20°–25°	0.00%	0.00%	0.00%	0.25%	0.72%	1.40%	5.90%	8.13%	3.67%	2.94%	0.93%	0.46%	0.25%	0.00%	0.00%
25°–30°	0.00%	0.00%	0.00%	0.14%	0.38%	1.43%	4.57%	7.72%	4.54%	2.27%	0.78%	0.63%	0.10%	0.00%	0.00%
30°–35°	0.00%	0.00%	0.00%	0.28%	0.21%	1.43%	3.36%	9.73%	3.59%	2.07%	0.68%	0.53%	0.05%	0.00%	0.00%
35°–40°	0.00%	0.00%	0.00%	0.19%	0.31%	0.75%	1.40%	4.62%	1.01%	0.75%	0.78%	0.60%	0.21%	0.00%	0.00%
40°–45°	0.00%	0.00%	0.00%	0.06%	0.11%	0.14%	0.46%	1.46%	0.40%	0.32%	0.19%	0.15%	0.03%	0.00%	0.00%
45°–50°	0.00%	0.00%	0.00%	0.03%	0.00%	0.11%	0.10%	0.62%	0.14%	0.20%	0.16%	0.05%	0.00%	0.00%	0.00%
50°–55°	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.04%	0.04%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
55°–60°	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.05%	0.01%	0.00%	0.00%	0.00%	0.00%
60°–70°	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
70°–80°	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
80°–90°	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%

Figure 7 - final statistical matrix with 57 possible combinations of orientation and inclination (97%)

2.3.3 Statistical data on nominal power and location of PV systems

Most complete data on nominal power of PV systems is available on side of the grid operators. The main dataset used in the project is originating from Creos. It contained nominal photovoltaic power installed per village / city, last updated in June 2016. This dataset comprises 4'512 PV installations with a total nominal power of 109'875.96 kWp. Recently, a dataset from Sudstroum was received, but not yet implemented in the forecasts (containing 1.16 MW_p).

Obviously, once the forecast model would be operationally implemented, regular updates of this dataset are necessary, as the nominal power of PV systems is continuously increasing. For comparison, the Creos dataset contained: 102'120 kW_p in its version of 1.6.2014 and 92'737 kW_p (29.5.2013).

A distribution of the nominal power per commune as of June 2016 can be seen in Figure 8.

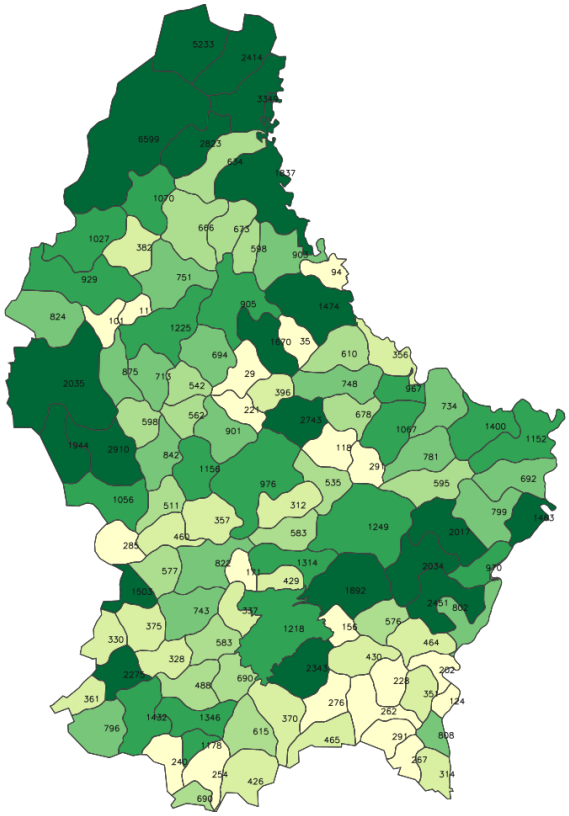


Figure 8 – installed PV power per commune

2.4 Data acquisition and setup of reference system profiles

2.4.1 Choice of suitable PV reference systems

It was foreseen to implement a number of reference PV installations into the model, spread equally over Luxembourg. The objective was to obtain a real time feedback from PV systems in order to be able to adapt the short term forecast based on their measurements. At the beginning of the project, it was decided not to install new measurement equipment, but to rely on the smart metering data from Creos which were already measured in a good temporal resolution (every 15 minutes). At that time, Creos had 173 PV installations with smart meters in their portfolio.
Creos and LIST informed each of these customers about our common project and requested their feedback. Only seven customers did reject to be part of this project. A unique number was given to each PV installation in order to anonymize the data. From that time on, only this number, the locality, installed power and GPS-position was used internally.

2.4.1.1 Desk audit of potential reference systems

From the remaining Creos customers, each of the PV installation was analyzed using Geoportail (LUREF Luxembourg 1930 / Gauss data in www.geoportail.lu)[1].

Hereafter the explanations with the example of the installation 0068 http://g-o.lu/telw4:
First, a view of the PV installation in Geoportail. The object is the yellow marker at the top of the picture:

Figure 9 - satellite view in Geoportail for the concerned PV installation (yellow marker)

The surroundings on that image were checked in order to see possible obstacles in direction of the sun (mountains, buildings, antennas, pylons etc.). With a closer view, the orientation of the PV installation can be checked by the measuring tool of “Geoportail”. As most of the installations are medium to large sized, they were not in the AEV dataset, therefore no information on orientation (and elevation) was available.

[1] old version has been used which is still reachable with the link http://mapv2.geoportail.lu/.

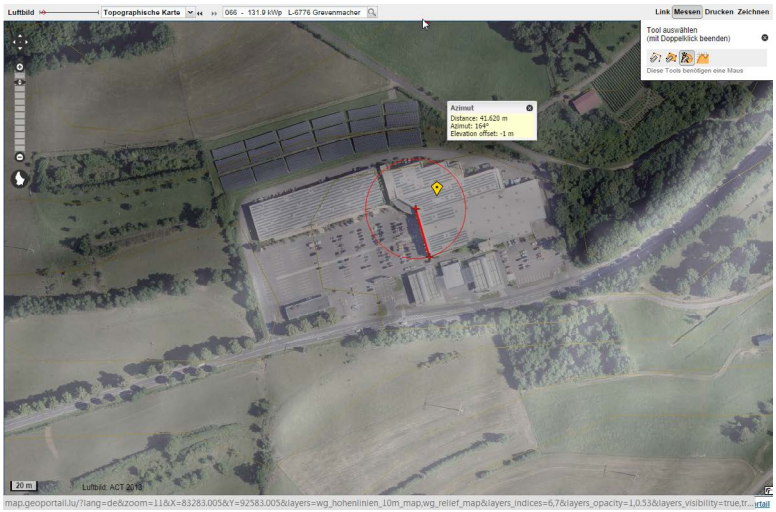


Figure 10 - measuring the orientation of the PV installation by Geoportail tool

Some PV installations were distributed over several roofs with different inclinations and orientations. At the beginning, these installations were sorted out, but as suitable installations were scarce, some had to be included again at a later stage. Mainly in the North of Luxembourg, suitable PV installations were rare.

In a next step, several cross-sections (measuring the height profile along a line from the installations) in different directions were created in Geoportail. This gives an indication if an obstacle (e.g. a hill) might cause shading on the PV modules. Similar analysis were done for several potential sun positions.
Only partial shading during wintertime or during summertime (in the early morning or late evening) was tolerated.

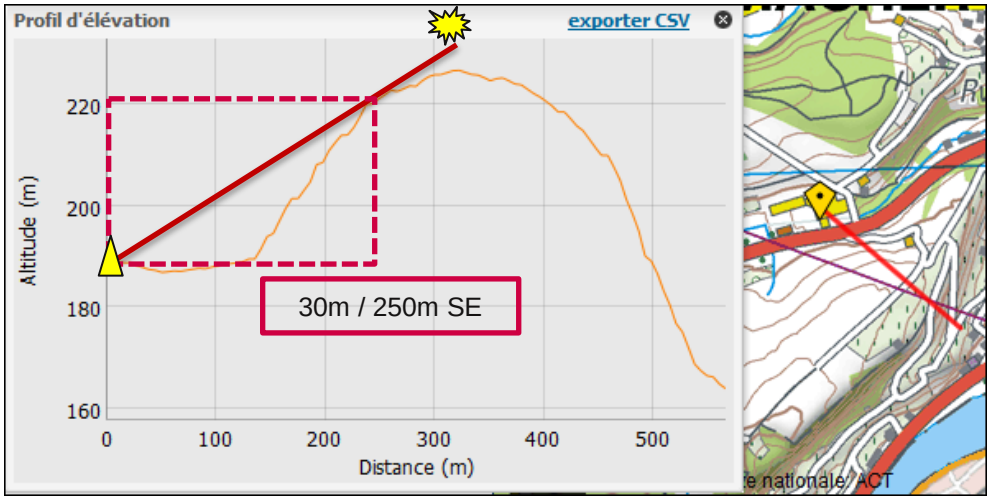


Figure 11 – analyzing potential shade effect of obstacles surrounding PV reference systems

Further insights could be gained from Google Maps, as date of the images and resolution were different as compared to geoportail.lu.

Finally, about 80 PV installations to be considered as possible reference systems remain after this check.

2.4.1.2 On site visits of potential reference systems

A checklist with more than 70 items was elaborated in order to standardize the visit of a PV installation. Each installation had to be characterized according to their technical parameters, such as type of PV modules, number and type of inverters, cross section and lengths of all the cables, inclination and orientation of the PV arrays, age of the installation, location and array set up - to mention only the most important parameter.

The acquired data for feeding the model was based on own measurements and/or investigations on site, plans from the owners or the installer, information asked by phone calls or own estimations. After a site visit, a technical model was set up for each PV system.

Finally about 80 PV installations were visited.

2.4.1.3 Difficulties of system profiles, assessment and data acquisition

Characterising the PV reference installations turned out to be challenging. The problems that have been faced could be divided in two groups of factors: technical and human factors.

Technical factors - most of the PV installations with metering data were medium to large-sized (up to 999 kW_p). Thus it was time consuming to characterise the installation (e.g. checking if the number of PV modules matches the plans, estimate the cable length with their corresponding cross-section, measuring several angles of the inclination of the roofs and PV module).
Further, the installations are

- not built as designed,
- spread over several roofs (different orientations / inclinations) - must be treated as several sub-systems,
- partly not working / have been modified - a technical "history" was necessary, but exact dates of modifications were often not available.

The owner often does not have the necessary information and refers to the installer and vice versa. In several cases the installer in not in business anymore. Also, if the installer has no maintenance contract for the system, the costs arising from a site visit or handling the requested information is not covered, so some refused to support the LIST in that request.

Human factor – in some cases, the owners could not reached, the installer forgot to inform LIST about a maintenance visit or owner / installer seem to have other objections to a site visit.

Limitations concerning the current implementation of metering devices – The PV systems currently measured frequently by Creos deliver their data with one day delay. Even in the standard setup of the currently being rolled-out smart-meters, it is not foreseen to deliver "real-time" data from the PV installations, e.g. every quarter of an hour, to a central data storage.
Therefore, the adaptation-loop comparing the generated energy of a reference system with its individual forecast is, in the current setting, not possible, but would be feasible with a specific adaptation of the future smart-meters.

2.4.2 Synthetic system profiles for PV reference systems

Due to above explained problems in gathering detailed information on the reference systems, the approach of synthetic system profiles was developed (see also 2.2.12.). The idea behind the approach was to use standard values and to analyze publically available data sources to gather the necessary information to run a synthetic profile for the PV system, which describes the behavior in a PV system sufficiently to be calibrated with the help of the measurement data.
The list of previously identified potential PV reference systems, where measurement data is available, was screened once again. With the help of several additional data sources, a priority list of systems was evaluated to estimate the suitability of the system and the available information to be used as synthetic system profile. The full desk audit of potential synthetic reference systems was documented in an internal logbook.

Following data sources were used:

- **List of monitored PV systems** available at side of the grid operator Creos
- **Geoportail.lu** – Aerialphoto / Orthophots from the years 2001, 2004, 2007, 2010, 2013 and the provided measurement tool of geoportail.lu: enables the identification of the surface area of a PV installation, the orientation, rough estimation of year of construction and evtl. adaptations or expansions of the systems.
- **Google maps** – recent and historical aerial photos
- **Google street view** – if available – potential estimation of inclination of the PV system

- **Statistical information on PV systems** supported by the financial incentives administrated by the Administration de l'Environnement (not publically available)
- in specific cases, additional **visits at the installations** were done
- **measurement values** from smart meters

Nominal power and measurement values are provided by Creos for research purposes. Orientation, inclination, year of construction and module surface of the systems is derived from several GIS data and orthophoto sources mentioned above. Where this information is not available or not plausible as compared to data given from Creos, or a comparison of module surface with the nominal power given and max. power measured is not rational, the system dropped out of the list of potential synthetic system profiles.
After a first round of desk audits of systems from the priority list, the regional distribution of the identified reference systems was checked and further systems in specific regions were additionally taken into account.
This whole process was relatively time consuming and couldn´t be automated, but finally provided a set of 15 synthetic system profiles.

2.4.3 Full list of PV reference systems used

Finally, after several attempts of site visits, telephone contacts, online research and desk audits, a set of 8 detailed reference system profiles and 15 synthetic profiles is available.

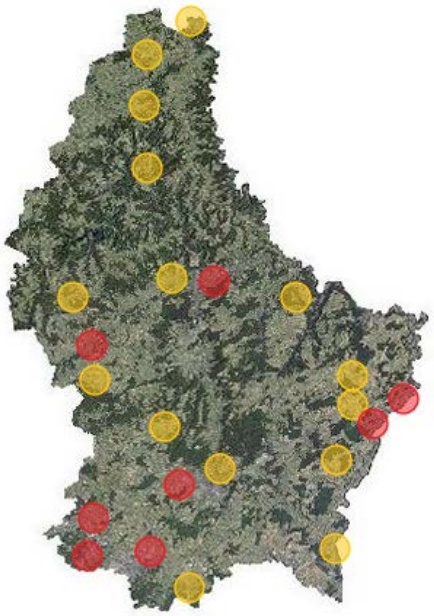


Figure 12 - map showing 8 technically detailed reference systems (red) and 15 synthetic reference systems (orange)

The spatial distribution of the reference systems is displayed in Figure 12. Additional checks of the list of measured PV systems with the intention to integrate additional systems in areas with low coverage were done and partially successful, but still the coverage in the North-East, North-West and in the middle of the country could be improved. Unfortunately, there are currently no suitable systems available.

2.5 Adaptation of system profiles and estimation of calibration factors

After the completion of the model and the set of reference systems, the forecasting system needs a first run of the model in order to evaluate the suitability of the model and the chosen or acquired parameter for the individual reference system. The techno-physical model represents the theoretical, optimal behavior of the reference system. As irradiance measurements on site of the reference systems were in most cases not available, it was not possible to calibrate each reference system independent of the forecast data. Irradiance forecast are known to be relatively precise on days of clear sky conditions. Therefore, each forecast of the reference systems has been compared with the measurement curve for specific days (in March and July 2014) where forecast and real measurement showed cloudless conditions.

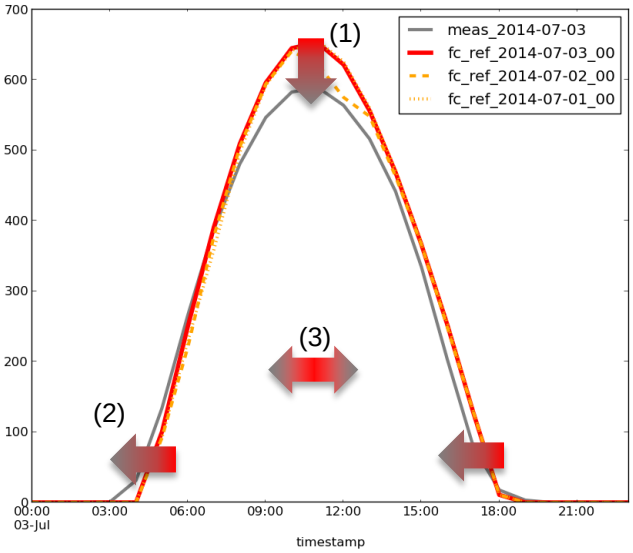


Figure 13 – example of deviations of the "clear sky"- curve of reference system nr. 0004 from the measurement under cloudless conditions

The shape of each curve can be influenced (1) in its height (by calibration factor mainly) or (3) in its width (by adapting the inclination) and can be shifted (2) e.g. towards earlier hours (by turning the orientation angle eastwards). Adaptations in orientation and inclination have been used scarcely. The calibration factors have been chosen after an analysis of the relative monthly error ($\varepsilon_{M\ dt}$), considering day time values only, for the months March and July 2014.

Ref.nr	**0004**	**0080**	**0138**	**0067**	**0112**	**0040**	**0016**	**0058**	**1006**	**1008**	**1010**	**1013**
Calib.	0.93	0.92	0.97	0.94	0.95	0.93	0.93	0.92	0.93	0.93	0.93	0.96

Ref.nr	**1042**	**1052**	**1063**	**1075**	**1090**	**1101**	**1122**	**1134**	**1159**	**1169**	**1173**
Calib.	0.99	0.97	0.92	0.97	0.98	0.91	0.92	0.91	0.94	0.88	0.92

Table 3 – Calibration factors chosen for the reference systems

From the chosen calibration factors, it is obvious that the model generally overestimates the expected PV power. All calibration factors are below zero, ranging from 0.88 minimum up to 0.99.

The effect of applying these calibration factors and adapting the systems parameter to better fit the curves for clear sky days, has been evaluated and documented under 3.3.

2.6 Development of an approach to establish a feedback-loop in the forecast

PV power forecasting is relying on the accuracy of the solar irradiance forecasts and can thus, if solely based on those predictions, never be more accurate than the underlying meteo forecast. Although, the solar radiation forecasts improved a lot during the last decade, they are still the main source of uncertainty, as the physical description of the PV systems is comparably straightforward. In solar radiation forecasting by numerical weather prediction models (such as the forecasts delivered by the

ECMWF), mathematical equations describe processes in the atmosphere and the models are being fed with measured parameter from the recent past and current observations. The models predict near future developments in the atmosphere, such as cloud movements and cloud building on different heights, which influence strongly the solar radiation reaching the earth's surface. These cloud movements and transformations on several heights need to be precisely predicted in their speed, direction and thickness, in order to estimate their regional effect on solar radiation. On larger areas, inaccuracies in cloud movements balance out more strongly and the accuracy of the forecast increases.
Within this project, we wanted to test if and to which extend measurements from existing PV systems in the forecasting region might help to improve the short-term forecasting accuracy to level out irradiance forecast errors.

Two approaches have been followed in order to create a feedback loop based on online PV power measurements (comparable to smart metering):

- **Error persistence** : First analysis of a comparison of the forecasts for single PV systems to the measured power, did show persisting trends (over few hours) of over- or underestimation of the real power. Although this is not generally the case, there was enough evidence to test whether a correction of the forecast, 1h to a few hours ahead, based on the assumption of a persisting error would increase the accuracy of the power forecast. (results see 3.7.1)

- **Error movement vectors** : The thesis behind this approach is, that a main source of error in solar irradiance forecasting arises from inaccurate forecasting of cloud movements (direction and/or speed). A cloud front moving into the forecast area over time might thus lead to over- or underestimations along its front-border, depending on whether it is moving faster or slower than predicted (or inaccurate in direction). The errors, which can be derived for the single PV systems used as references, could thus be visualized on maps and might show graphical patterns moving over the forecast area with time. If clear movement patterns could be identified, existing methods used to predict cloud movements, could be used to forecast the propagation of error movements. This should be validated within this project and evaluated to which extent a forecasting of these error patterns is possible and leads to more accurate predictions. (results see 3.7.2)

2.7 Upscaling procedure

Upscaling, in the context of PV power forecasts, normally describes the process of extrapolation from forecasts for a specific representative set of PV systems to the entirety of systems in the forecast region. The approach described here is a bit different – described step wise:

The following procedure is performed for each hour and could respectively be performed for different forecast horizons:

1. Each forecast cell (see Figure 2) gives a forecasted global horizontal irradiation (0-72h ahead)
2. Based in the "error persistence method" ("error movement vectors"-method not feasible, see 3.7.2), the forecast could be adapted to improve short-term accuracy of the forecast
3. For each forecast cell, a synthetic 1 kW_p PV system is described. This model uses the same equations, parameter and assumptions as the synthetic reference systems profile. Each of these synthetic 1 kW_p systems consist of 57 subsystems, representing the different orientations and inclinations, which are most common in the entirety of PV systems in Luxemburg (2.3). The nominal power of each subsystem is weighted accordingly to the share of PV power installed in the country within its class of orientation and inclination. Based on this model, the power forecast for each subsystem is summed up for each forecast cell, representing a power forecast factor [kW/kW_p] for this cell and hour.
4. A list of PV systems in the grid of Creos (incl. their coordinates) and Sudstroum is used to multiply the nominal power of each PV system in Luxembourg, with the power forecast factor from its specific forecast cell
5. The individually forecasted PV power of each system is assigned to a commune (could also be another aggregation level) and is summed up. The result is an hourly PV power forecast per commune, as seen in Figure 29.

2.8 Description of the current implementation of the mathematical model

The mathematical model, as described above, is currently implemented in several Phyton scripts. This includes the automation of the following steps:

- Retrieving solar irradiance and ambient temperature forecasts from ECMWF web server
- Processing of solar irradiance on tilted surfaces
- PV power forecasts for the reference systems
- Calculation of deviations between forecasts and measurements (in csv-files and plots)
- Processing of error maps on hourly basis
- Upscaling of the forecast on PV systems in Luxembourg
- Aggregation (currently on communal level)

Phyton is an open source / public license programming language, widely known, which facilitates the transfer to potential users and its implementation on their side without cost intensive software requirements.

The analysis and assessment of the forecast performance was done manually, using specifically developed Excel-templates. Those analyses were necessary in the development phase, only. The processing and visualization of geo-referenced data was done using an open-source, public license GIS tool: Open GRASS GIS. The adaptation of the forecasts by a feed-back loop from metering data, based on the error persistence approach, is currently still done manually, but its automation would be straightforward.

Parts of the calculations are computationally quite intensive and took several days, as we were calculating the forecasts for the three forecast horizons for the full two years of historical data and were producing csv-files and plots for the performance analysis. The upscaling procedure is also calculation time consuming – as an example:
For each of the effective 34 forecasting cells for Luxembourg (see Figure 2) was created a statistical matrix, representing the orientations and inclinations of relevance for Luxembourg - 57 combinations. This results in a total number of 1’839 combinations (34 x 57), which are calculated as virtual PV subsystems. The forecasts were calculated for these 1’839 virtual PV subsystems for each 24 hours of a day and during 2 years (730 days) with the corresponding meteorological forecasting data from ECMWF: Hence,
1’839 PV subsystems x 24 hours x 730 days resulted in 10’739’760 calculations of the full physical model for only one forecast horizon!

In its practical implementation (leaving out the performance analysis and only forecasting 72h ahead), calculation time is obviously much lower, so that is should not become a bottle neck in the regular forecasting. Anyhow, parts of the upscaling procedure have already been parallelized in order to be run on servers using a number of CPUs in parallel to reduce calculation time.

3 Project results

3.1 Evaluation of the physical model to process the irradiance data

A comparison of the calculated values for an inclined surface based on GHI measurements and the chosen approach with the measured inclined global irradiation (both measurements done at our own meteo station and PV test field) shows the relatively good fitting of the trends of both curves, measured and calculated (Figure 14). However it can be seen that there are slight deviations of the calculated (red) values above measured values (blue), which cause small overestimations of the calculated inclined irradiation for clear days.

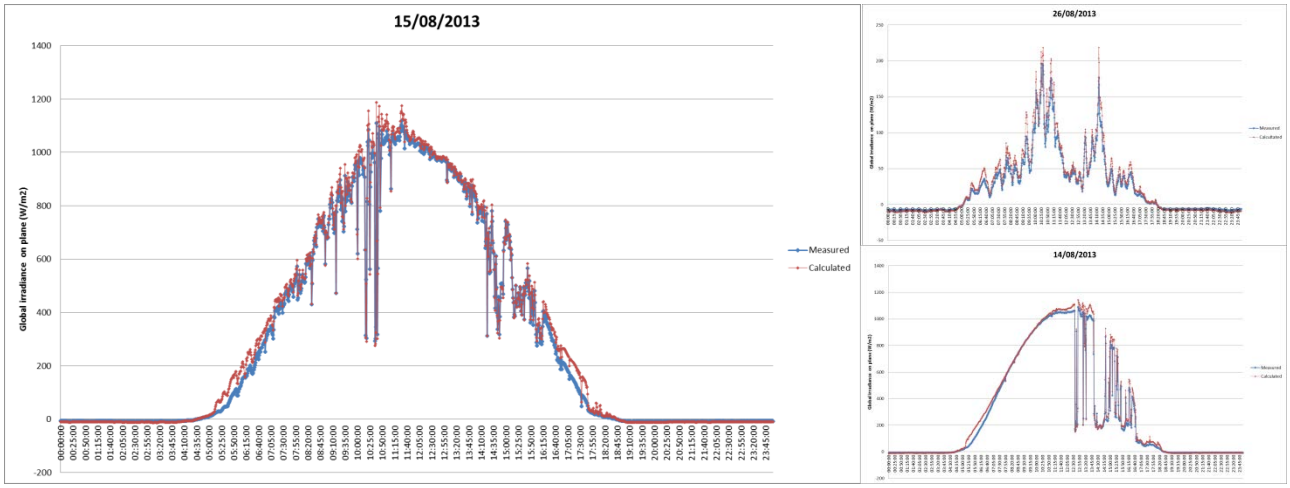


Figure 14 – examples of the validation of the calculated global irradiance on an inclined surface (red) (based on GHI) and the measured global irradiance on an inclined surface (blue)

Comparing the pyranometer measurement values at inclined orientation of 30° elevation and 0° azimuth with the calculated values based on GHI measurements, it can be seen, that the variance remains low even under different cloud conditions over a month. Figure 15 shows the data of July 2014, which was mainly clear (six very cloudy days, five very clear days and twenty predominantly clear days). It can be seen, that the calculated values have a low scatter but a systematic deviation in relation to the measured values probably due to the high daily irradiation during the measurement period.

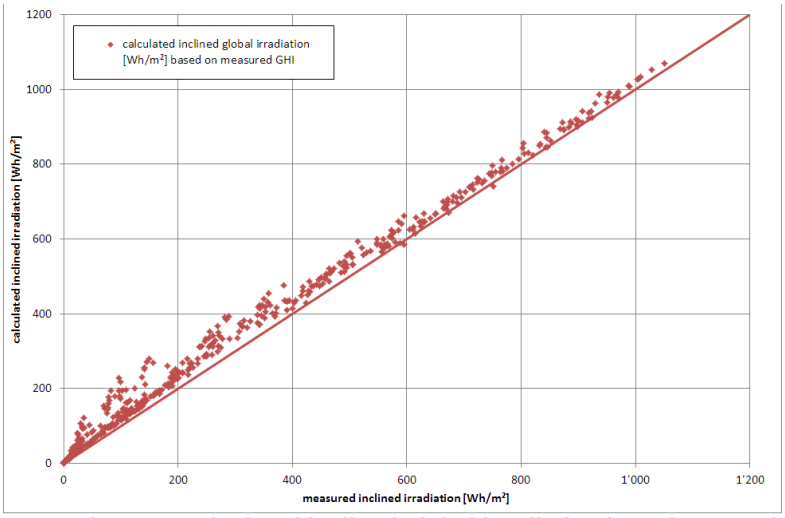


Figure 15 - July 2014, calculated inclined global irradiation based on GHI (y-axis) over measured inclined global irradiation (x-axis)

3.2 Validating PV-model performance

In order to validate the quality of the techno-physical part of our PV system model, independent from the quality of the irradiance forecasts, we used measured "irradiance in plane" from reference cells of three reference systems, for which such data was available. The measured irradiance in plane was used as input for our model instead of irradiance forecasts and the modelled energy yields were compared for one month (July 2014) with the measured energy yield data from smart meters (see Figure 16 as an example).

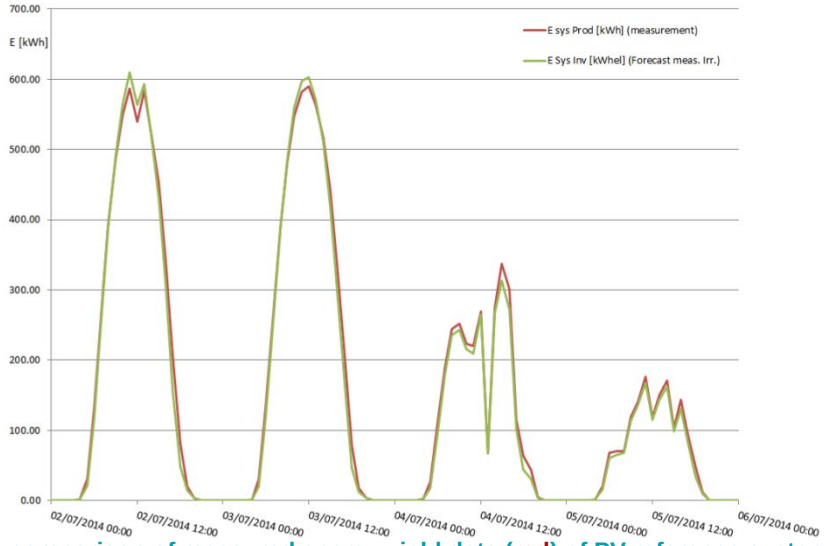


Figure 16 - comparison of measured energy yield data (red) of PV reference system (nr. 0004) to modelled energy yield (green), using the irradiance measurements of reference cells as model input (here only four days in July 2014 are shown)

The model shows good correlations with the measured data throughout the whole months. As an illustration, four days (two with clear weather conditions, two with overcasted sky) have been chosen demonstrating the good consistency of the modelled PV yield with the measured data. We observe only slight overestimations of the model at clear weather conditions and slight underestimations at cloudy conditions (more details can be found in the publication (Minette, Koster et O'Nagy 2014) in the Annex).

3.3 Efficiency of the calibration factors and adaptation of system profiles

As explained in chapter 2.5, each system profile has been adapted based on characteristic curves for cloudless days in March and July 2014, respectively calibrated based on the monthly normalized error. Figure 17 illustrates the effect of calibration and a slight adaptation of the orientation of the system nr. 0004. The resulting forecasts after this modification fit better the measured curve in both months. In July 2014, there is still a visible overestimation of the PV power, but further increase of the calibration factor would lead to an underestimation in March. Similar effects could be observed for other reference systems, therefore the chosen approach was kept and no second round of calibrations was done, for the time being.

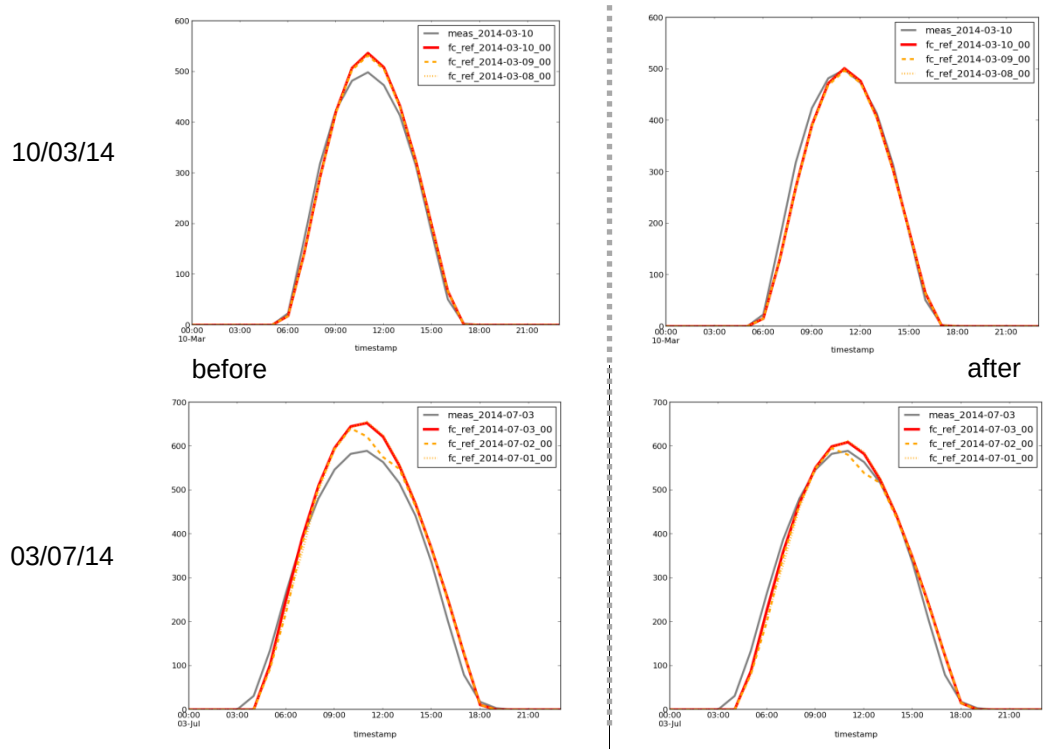


Figure 17 – Comparison of the forecast (red – 24h forecast) to the measurement for two clear weather days (10.03.14 on top and 03.07.14 at the bottom) before (left) and after (right) calibration and adaptation

Generally, the calibration was found to be a very effective way of reducing this error on level of reference systems. Figure 18 depicts that the calibration led to a reduction of the monthly normalized error $\varepsilon_{M\,dt}$ (see also 3.4.1) for all reference systems, with only few exceptions for March 2014. The improvement ranges from a 0.46 % (system 0142 in July '14) to a 8.5 % difference (system 1173 in July '14). The few exceptions are concerning the error reduction in March for the systems 0138, 1042, 1075 – here the March forecasts before calibration were already quite well. The calibration led in two of those cases to a slight underestimation of the March production values.
On average, the calibration reduced the error in these two months by -2.66 % at a mean deviation before the calibration of -5.96 % - hence, the mean deviation after calibration was -3.3 %.

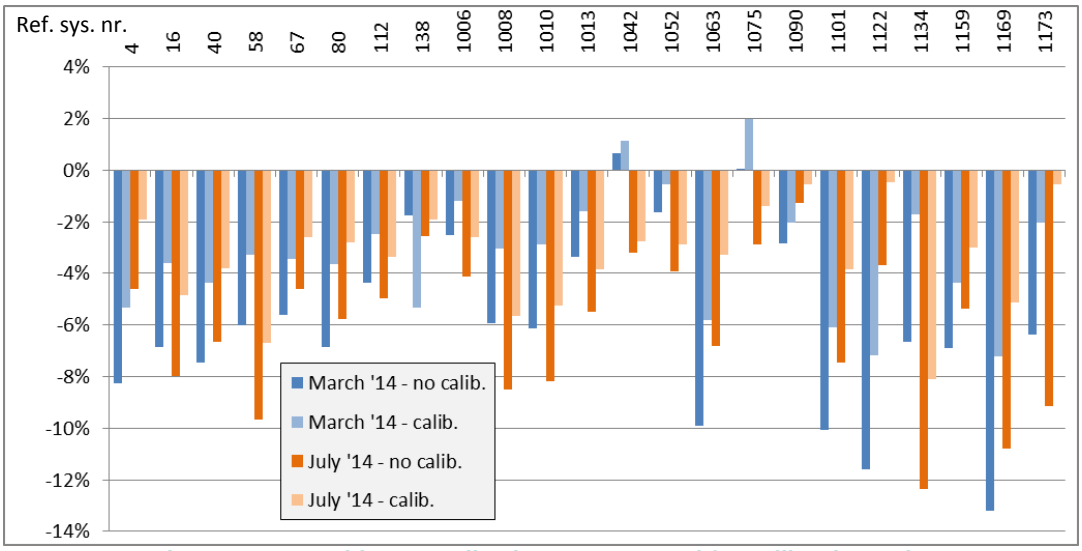


Figure 18 – monthly, normalized error $\varepsilon_{M\,dt}$, used for calibration only, and its reduction after calibration

3.4 Performance evaluation of the forecasts on reference system level

In chapter 3.4 the forecasting performance is assessed – meaning the actual accuracy of the forecast. In the case of PV power forecasts, the focus is on the hourly performance of the forecasts as an important aspect for the grid operators and utility companies. The accuracy on forecasting the PV production on monthly basis or daily sums is normally higher, as compared to hourly values, where accuracy in time is a challenge. Furthermore, the accuracy of forecasts for specific single sites (single PV systems, such as the reference systems in our approach) is lower than for regional forecasts, as in larger forecast areas local phenomena (e.g. cloud movements) can level out.

Although the final output of this PV power forecast model described here are regionalized forecasts, only accuracy on single site forecasts are evaluated. The simple reason for this is the lack of hourly measurement data for the entirety of PV systems in the forecasting region to which we could compare our forecasts.

Although the calculations have been done for three forecast horizons (0-24h, 24-48h, 48-72h), the focus is on intra-day forecast, 0-24 h ahead. From literature it is know, that NWP models (such as the ECMWF data used here) perform best for 1-3 days ahead irradiance forecasting, but on very short term intra-day forecasts, other algorithms might perform better (cloud-motion vectors for example) (Kleissl 2013). The approach described here, using smart metering data of PV reference systems, tries to enhance NWP based power forecasts on this short-term intra-day time scale.

3.4.1 Introducing evaluation criteria

In order to be comparable to other evaluations of forecasting approaches, the following evaluation criteria have been chosen in analogy to the literature (Lorenz, et al. 2011).

To evaluate the accuracy of the forecasts for the individual PV reference systems, each hourly value has been compared to the measurement value. The error has been normalized to the nominal power of each reference system and is given as:

$$\varepsilon(t) = \frac{P_{pred}(t) - P_{meas}(t)}{P_{nom}(t)} \qquad (21)$$

P_{pred} = predicted power of the PV system [kW]
P_{meas} = measured power of the PV system [kW]
P_{nom} = nominal power of the PV system [kW_p]

The root mean square error (RMSE) is a common term in the evaluation of forecasting algorithms for solar irradiance (Kleissl 2013) as well as for power forecasts in wind and solar. RSME is supposed to be suitable; specifically for power predictions and their application in utility companies, as large error are disproportionally problematic, as stated by (Lorenz, et al. 2011).

$$RMSE = \frac{1}{\sqrt{N}} \sqrt{\sum_{t=1}^{N} \varepsilon(t)^2} \qquad (22)$$

The mean value of the error (bias) is further interesting to evaluate the performance and to identify systematic errors in the forecasts:

$$bias = \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t) \qquad (23)$$

During the development of this project, another evaluation factor has been introduced which was used to estimate calibration factors (introduced under 2.5). For this purpose, the relative monthly error ε_M has been calculated, normalized to the maximum power measured in the month for a specific PV system.

$$\varepsilon_M = \frac{1}{N} \sum_{t=1}^{N} \frac{P_{meas}(t) - P_{pred}(t)}{P_{M\,max}} \qquad (24)$$

$P_{M\,max}$ = maximum power measured in the month under evaluation [kW]

Another important aspect in the evaluation of solar power forecasting is the aspect of handling night time values. Irradiance forecast and real production are zero during the night. Thus, forecast and measurement completely fit and the error is zero. If night time values (hourly errors) are taken into account when estimating the evaluation criteria, the results show better performance obviously, only by trivial night time forecasts. For this reason it has been decided to evaluate both, the performance including night time values and taking into day time values only. The evaluation criteria are marked with the suffix “dt” if it considers “day time” values only:

$RMSE_{dt}$ = root mean square error, considering day time values only
$bias_{dt}$ = bias, considering day time values only
$\varepsilon_{M\,dt}$ = monthly normalized error, considering day time values only

The normalized error ε for each hour, as the base value for the other evaluation criteria, can potentially be larger on days of high solar irradiation and thus high PV power. Hence, the mean power “mean P” within a certain time span is an important reference value. Also “mean P” is normalized to the nominal power for reasons of comparability.

mean P = mean PV power of a system within a reference period
mean P_{dt} = mean PV power of a system within a reference period, considering day time values only

3.4.2 Forecast performance of the hourly values – on monthly basis

In order to evaluate the performance of the forecasting model, the error on hourly forecast values ε for each reference system, as compared to the measured value, has been calculated and normalized to its nominal power (considering the intra-day forecast, 0-24h ahead). The hourly, normalized error ε is evaluated on a monthly basis over 2 years, 2014 and 2015, in the following chapter.
In Figure 19 the performance evaluation criteria for reference system nr. 0067 are exemplarily depicted - further graphs of all reference systems can be found in the annex. RMSE and bias, related to the left axis, illustrate the accuracy of the forecast for the respective reference system and can be set in relation to the mean power for each month (right axis).

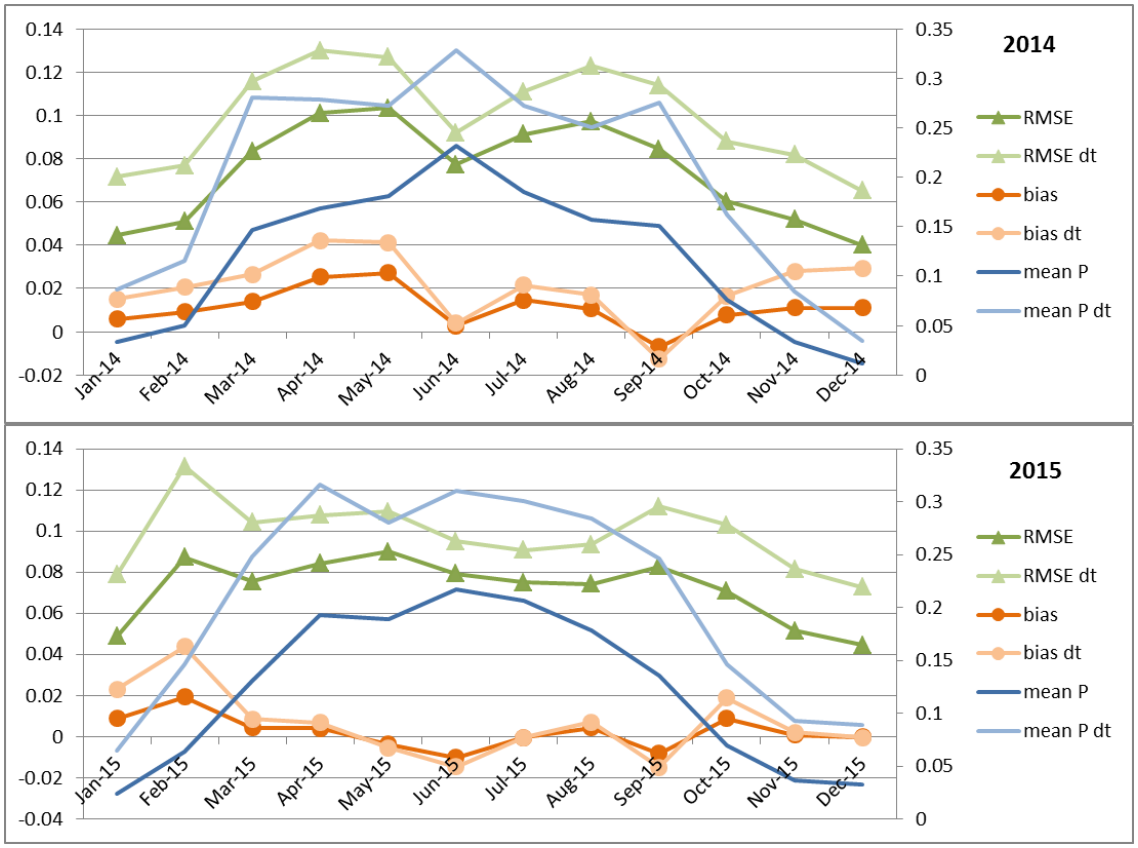


Figure 19 – evaluation criteria of the hourly performance for the two years 2014 and 2015 (here for ref. sys. Nr. 0067)

Mean monthly power, normalized to the nominal power of the system and shown on the right axis, gives the average power the system delivered in the respective month. The value is given in order to set the other evaluation criteria into relation with the mean power, as deviations from the forecast could be larger in months with relatively high irradiations. The curves of "mean P" and "mean P_{dt}" show the expected seasonal variation, peaking through summer time. The system shown here has a typical curve, as compared to the other reference systems. During June, the system reaches its peak production when "mean P" lies around 0.2 of its nominal power, meaning 20%, (30%, respective 0.3 if only day time values are considered). On first sight this seems relatively low, but as the systems reach their nominal power only a few hours a month, even during summer, this is a normal value that can be validated by literature.

The "bias" evaluates the actual mean error of the forecast, without specific weighting. A low bias means that there is low systematic error in the forecast – the system is neither over- nor underestimating the actual PV power constantly. Anyway, there can be large deviations in the single hourly forecasts that might compensate each other and are not visible in bias only.
Over the two years, the monthly bias ranges from 2.7% to -1.0% - a representative value for the set of reference systems. Bias shows no clear seasonal deviation over the two years, which can be confirmed by the other systems. If curves of the different systems are compared (see Annex), similar bias curves can be observed. This hints to a bias originating from the irradiance forecast for the specific months and confirms the suitability of our model throughout the seasons.

The root mean square error RMSE represents a mean error, weighting larger deviations much stronger than small deviations. The RMSE shown in Figure 19 ranges from values around 0.04 (4%) in January and December, up to 0.1 in April/May 2014. The graph shows representative curves for the set of reference systems, generally increasing during months of high solar power, as RMSE is specifically sensitive to large deviations which occur more frequent in this period.

Obviously, in February 2015 the RMSE shows an untypical increase for this period, which is not related to a technical problem on this specific reference system, as the same increase can be observed for the other PV systems. This effect is due to snow cover in February 2015 (confirmed by MeteoLux). Hence, the solar irradiance forecast is predicting the irradiance independent from the snow cover, but the PV-systems throughout the country underperform due to snow cover. This is a well-known weak point of PV power predictions based on irradiance forecasts only.

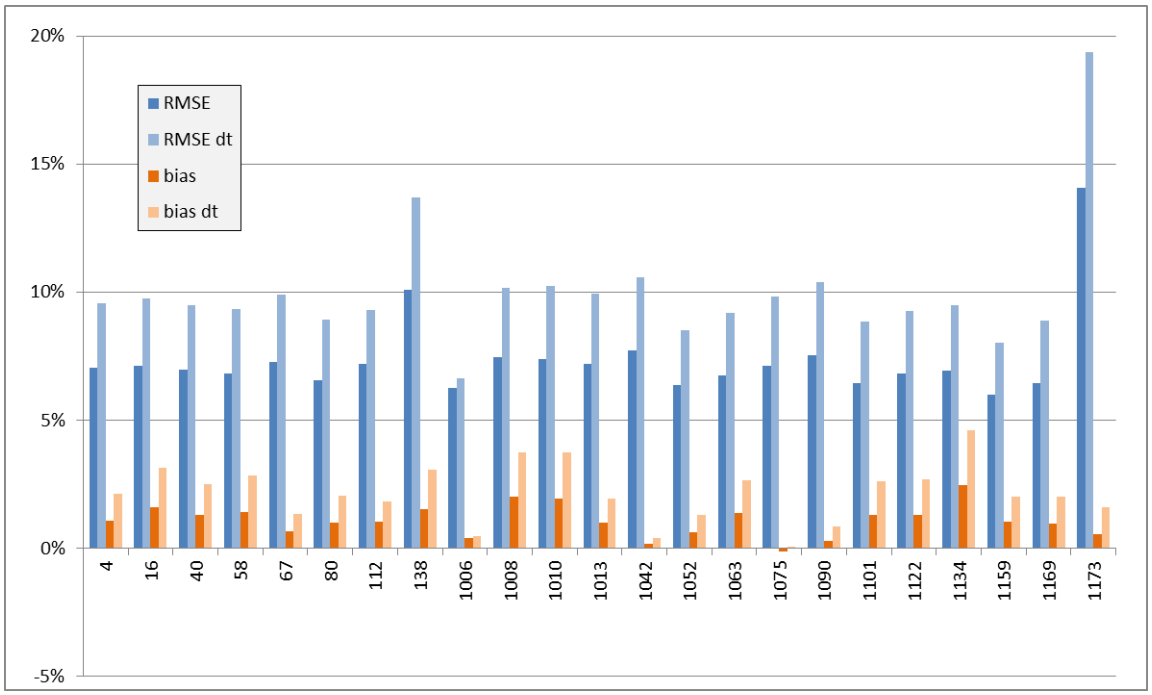


Figure 20 – mean evaluation criteria over 2 years (2014 & 2015) for the reference systems

Comparing all reference systems over the two years (Figure 20), the mean performance in terms of RMSE and bias is relatively similar with few exceptions. The mean bias over all systems is 1.1% ($bias_{dt}$ = 2.2%), while the values might range from -0.12 % (nr. 1075) up to 2.45 % (nr. 1134). Generally, the bias is positive in the range of 1%, which means an overestimation of the systems expected PV power. Considering bias, there are no extreme exceptions from that trend, but it could be checked if (for some systems, e.g. nr. 1134) a stronger calibration factor might reduce the bias.
On average, the RMSE over all systems lies at 7.4% ($RMSE_{dt}$ = 10.0%) and ranges from 6.00% (nr. 1159) to 14.09% (nr. 1173). Except for system nr. 0138 and nr. 1173, all systems perform comparably similar and their RMSE lies around 6.9%. The reason for the comparably high RMSE for the two outlying systems is unknown, yet – further analysis ongoing.

In comparison to other literature data for the accuracy of single site forecasts (not regional forecasts), the values given above seem reasonable and the model seems to work comparably well (e.g. (Kleissl 2013)).

3.4.3 Forecast performance of the hourly values – on daily basis

Although RMSE and bias give already a good impression of the accuracy of the forecast over a larger time scale, only the daily forecast curves give real insight in the daily performance. Therefore, similar plots as Figure 21 are available for all reference systems and every day in 2014 and 2015.

The plots of Figure 21 exemplarily show six days in July 2014 (01.07 – 06.07.) and their respective curves for the three forecast horizons (red / orange / yellow) and the measured production (grey) of system nr. 0067. Obviously, the forecasts fit relatively well the real production on clear days (02.07. & 03.07.). Larger deviations occur on overcasted days – this observation correlates with the reported accuracy of the irradiance forecasts from literature.
The plot for 05.07.2014 shows a rather cloudy day, resulting in relatively large hourly deviations from the real power, although the mean production fits well. But as the mean power on such cloudy days is relatively low, the normalized error (see Figure 22) remains in a good range.

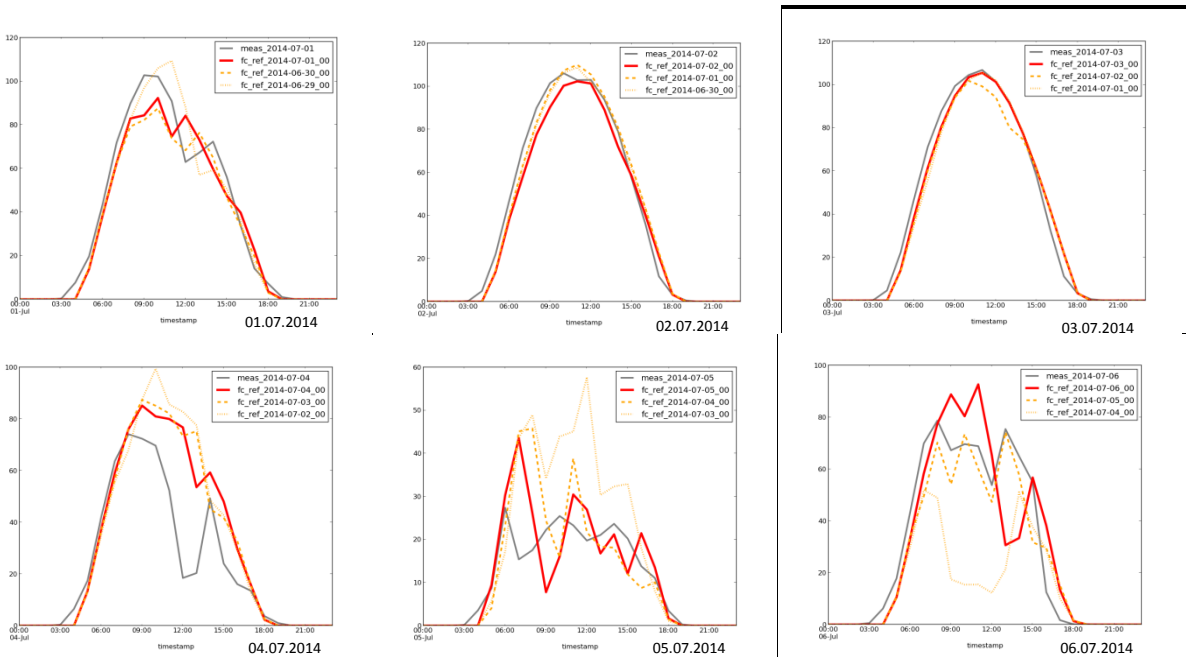


Figure 21 – example for system nr. 0067, six days in July 2014, showing the correlation of the three forecast horizons (0-24h in red line / 24-48h in orange, dashed line / 48-72h in yellow, dotted line) and the measured values (grey line)

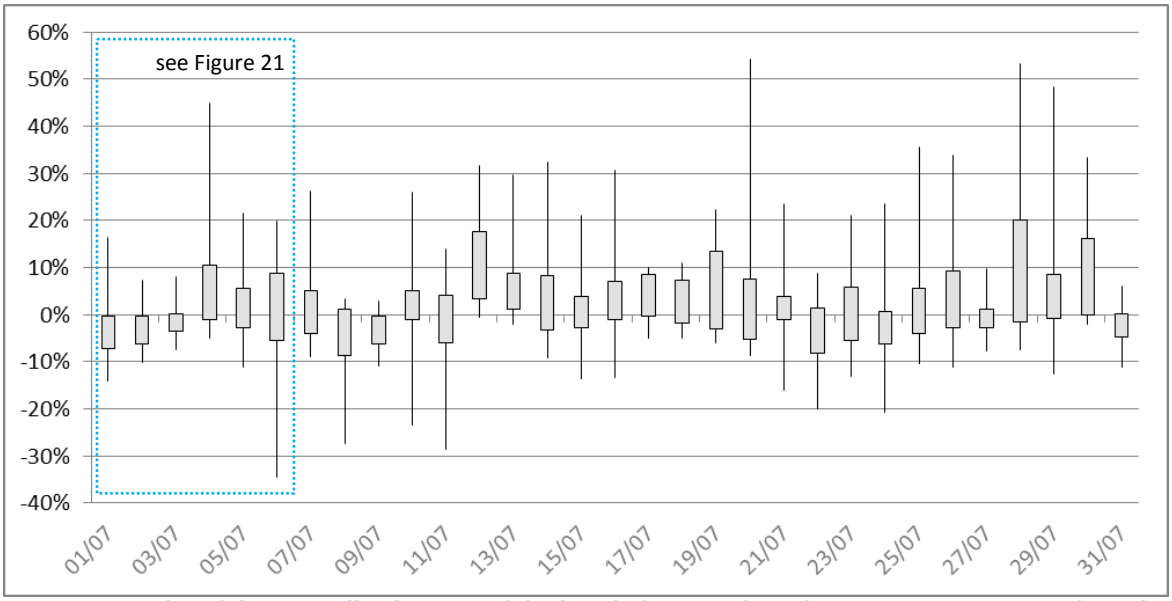


Figure 22 – Boxplot of the normalized error ε of the hourly forecast for reference system nr.0067 for July '14

The boxplot in Figure 22 illustrates the normalized error ε_{dt} and its variation on the hourly value. The grey boxes for each day depict 50 % of the forecast values around the median. The thin lines above and below the box show the upper, respectively lower 25% of the single values. Boxplots allow to give a quick overview on the quality of the hourly forecasts over a full month – they illustrate the bias as well as the scatter of the majority of the values and the extreme outliers. These plots are available for each month in 2014 and 2015 for each system.
It can be seen that for the large majority of hourly forecasts, the normalized error lies within a range of +/- 10%. But, single hourly forecasts can, in extreme cases, deviate from the real power in a range of more than 50% of the nominal power.

3.5 Comparison of the performance of the synthetic profiles of reference systems to the detailed system profiles

Synthetic system profiles were introduced due to a lack of detailed technical information on the PV systems to be used as references. As these system profiles work with standardized and estimated parameter, the synthetic reference systems were expected to have larger deviations to their actual measured power. But surprisingly, the evaluation of the forecast quality for the reference systems shows no real difference in their performance (see Figure 20 - synthetic profiles have ID numbers above 1000). The concept of synthetic reference system profiles might therefore be further supported in the future, since the effort to gather detailed technical information on the respective system is relatively high. With the roll-out of smart meters, the amount of potential reference systems will increase rapidly and also smaller and simpler systems will become available. This will reduce the problems which arise from the characterization of large scale PV systems (see under 2.4.1.3). In view of this development, this concept might become more promising.

3.6 Evaluation of the three forecast horizons – 24h / 48h / 72h

Within this project, the focus was on the short term and intra-day forecast. Nevertheless, the forecasts were calculated for three different forecast horizons: 0-24 hours ahead, 24-48h ahead and 48-72h ahead. A comparison of the accuracy of the three forecast horizons was made on monthly level. Figure 23 shows examplarily a boxplot for system nr. 0067, in October 2014, depicting the normalized error, considering day-time values only. The box illustrates the 50% of the normalized error values for the whole month around the median (middle line within the box). The upper and lower lines show the range of the more extreme upper, respective lower 25% of the error values – the end of the line representing the most exterme outlier.

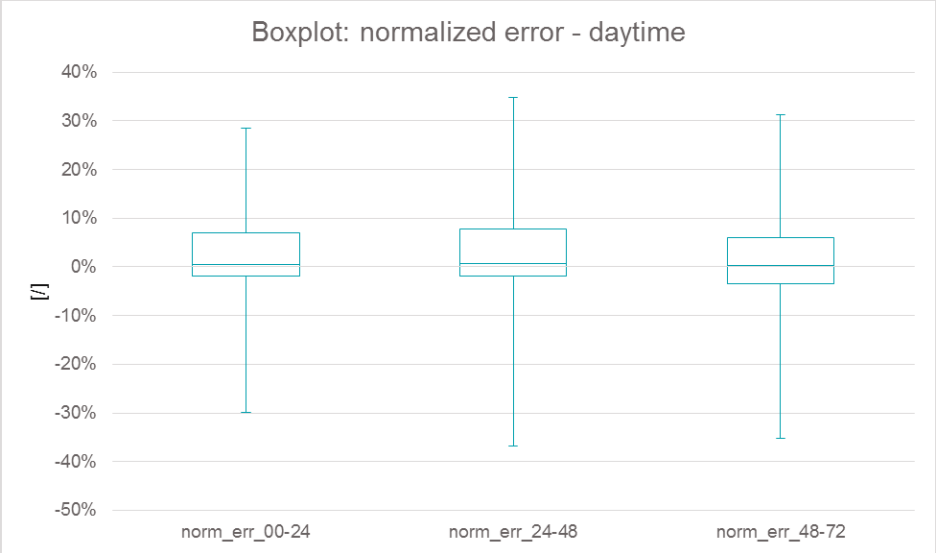


Figure 23 – boxplots of the normalized error ε_{dt} for ref. sys. Nr. 0067, showing the three forecast horizons (in Oct. 2014)

This months was chosen to be relatively representative for the whole set of 24 months for all references systems. We see no clear tendency in an increased accuracy for shorter forecast horizons, which could actually be expected to be more accurate. For this specific month, the median is around 0% and the 2nd and 3rd quartile (the 50% around the median) lie within a range of -3.5 % to 7.8%, relatively close to each other. Only a very small decrease of the error can be observed for forecast horizon 0-24h, when compared to longer horizons – specifically concerning the outer extremes of the forecast horizon 24-48h.

This trend can be approved when comparing the forecast horizons on monthly basis over the full range of 2 years. Exemplarily, again the plot for system nr. 0067 (Figure 24) is shown here. In this plot, the monthly normalized error $\varepsilon_{M\,dt}$ is depicted, which was developed for calibration purpose only and can be used to compare forecast horizons with each other, but should not be considered for estimating forecast

accuracy over the full period. It can be seen, that the best performing forecast horizon might change from month to month and is not necessarily always the shortest (the 24h) forecast.

	Forecast horizon 0-24h	Forecast horizon 24-48h	Forecast horizon 48-72h
Mean $\varepsilon_{M\ dt}$	-4.02 %	-4.67 %	-4.39 %

Table 4 – mean $\varepsilon_{M\ dt}$ over 2 years for all reference systems, comparing the three forecasts horizons

This observation can also be approved when calculating the mean monthly error $\varepsilon_{M\ dt}$ over the full period of two years for all reference systems. The 24h Intra-day forecast shows the best performance, but the differences are minimal (compare Table 4).

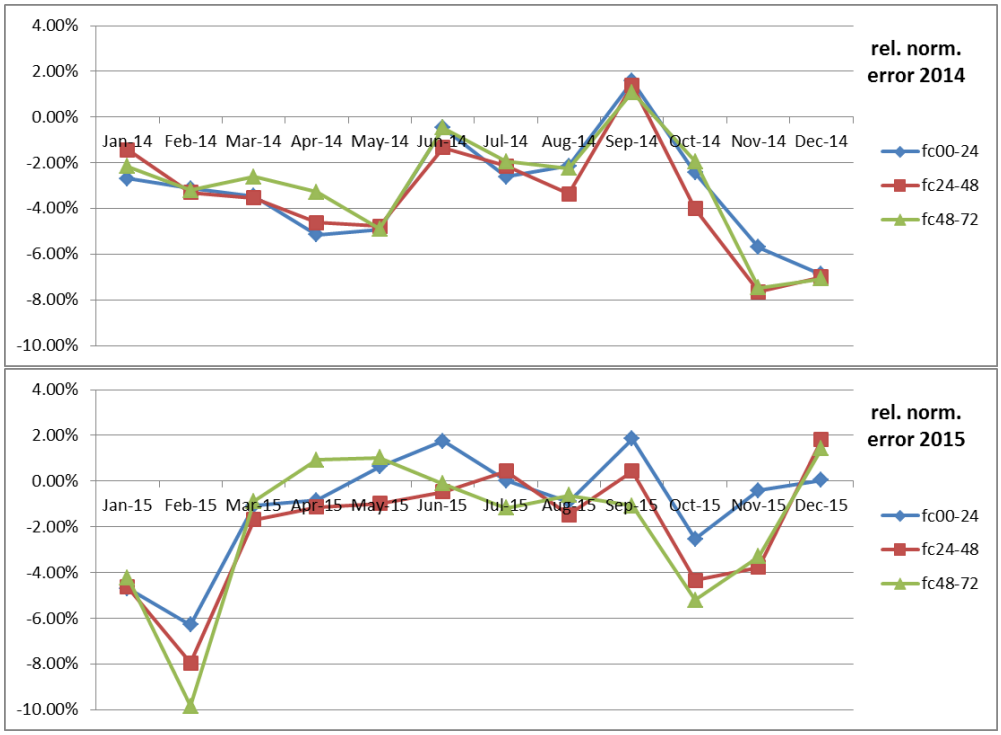


Figure 24 – monthly error ε_M normalized to maximum measured power of the respective month, for system nr. 0067, showing the three forecast horizons

3.7 Performance evaluation of the history-based forecast adaptation

3.7.1 Forecast adaptation based on error persistence

The analysis of the error ε over time for single reference systems revealed that for a considerable amount of days, the forecasts tended to "constantly" over- or underestimate the real power for a time span of several hours. Although, there are also days where the deviation for single reference systems swing from under- to overestimation hour by hour, the approach of error persistence should be tested.
In order to adapt the forecast for a specific PV system, the error ε for each time step was calculated and the deviation from the real measurement was considered to be persistent over a certain time range:

Example: Forecast adaptation by 1h error persistence for the time t_0

$$P'_{fc\,to} = P_{fc\,t0} + \varepsilon_{abs\,t-1} \qquad (25)$$

$P'_{fc\,to}$ = adapted power forecast for t_0
$P_{fc\,t0}$ = original power forecast for t_0
$\varepsilon_{abs\,t-1}$ = absolute error for t_{-1}

This approach was not expected to deliver appropriate forecast adaptations for longer time spans, but was test for 1 to 4 hours ahead.

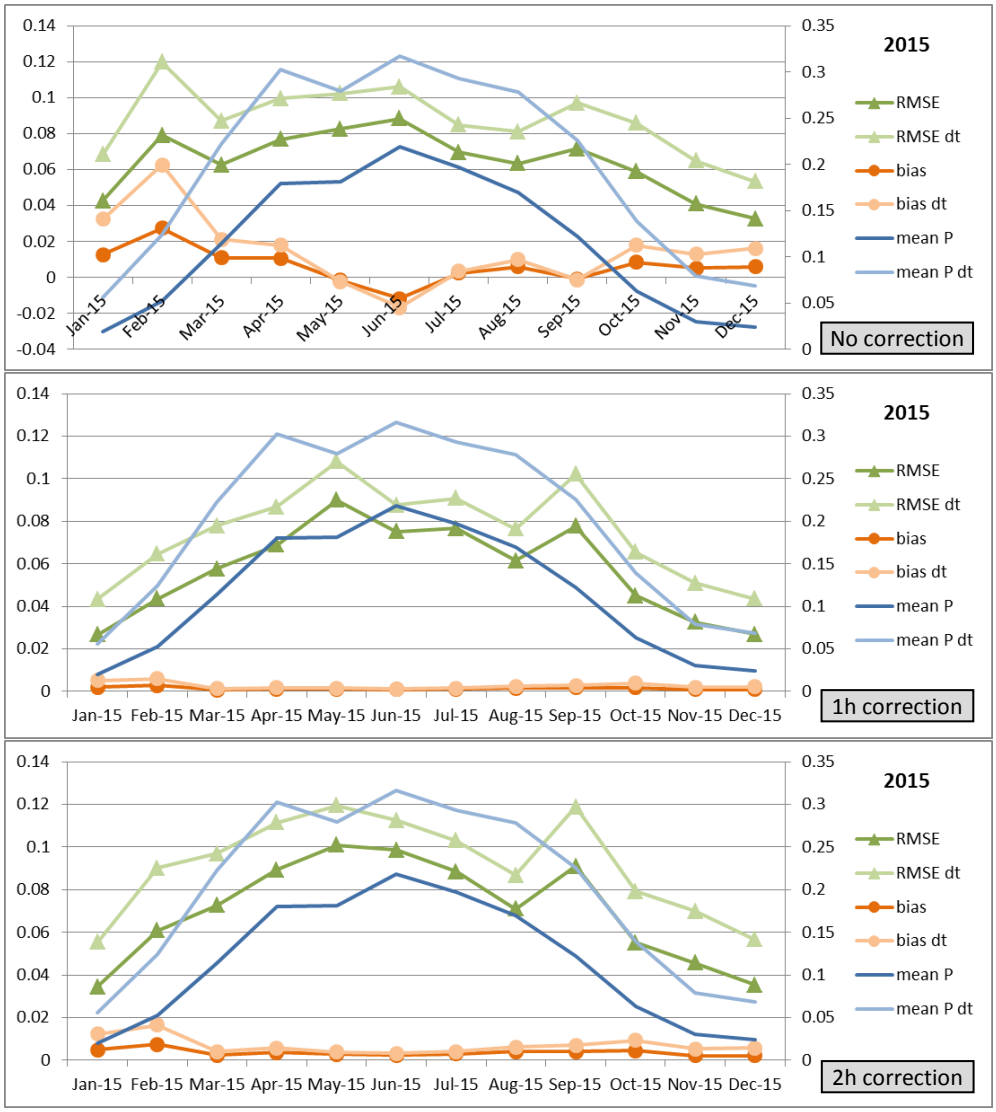


Figure 25 – evaluation criteria on forecast accuracy for system nr. 0080 in 2015 without correction based on error persistence (top), based on 1h error persistence (middle) and 2h error persistence (bottom)

Shown in Figure 25 are the results for a forecast adaptation based on 1 hour- and 2 hour-error persistence. With this simple approach, it was possible to reduce bias very effectively, which is not surprising. Both, the 2 hours ahead and 1 hour ahead error persistence adaptation reduced the bias to a minimum. Evaluated over the two years, exemplarily for system 0080, the bias dropped from 1 % (2.05% for $bias_{dt}$) to 0.14 % (0.27% $bias_{dt}$) for the 1 hour ahead error adaptation and to 0.4% (0.74% $bias_{dt}$) for the 2 hours ahead adaptation.
RMSE is less well improved, since this simple approach does reduce systematic error as well as short term persistent over- or underestimations, but does reduce outliers which influence RMSE to a greater extent. Anyway, for the 1 hour ahead adaptation, mean RMSE over the full two years does decrease from a value of 6.57% (8.95% for $RMSE_{dt}$) to 5.81% (7.62% for $RMSE_{dt}$). The 2 hours ahead forecast adaptation based on error persistence did not generally improve the RMSE. Figure 25 shows months with lower RMSE as well as higher RMSE in other months for the 2 hours ahead adaptation.
For longer time periods, 3 hours ahead or 4 hours ahead forecast adaptations, there is no improvement of the forecast possible– contrarily, the forecast quality drops as relatively high absolute errors from hours of high power (noon) would be shifted to hours of potentially lower peak power.

The very short term error adaptation based on error persistence works well for times of relatively constant over- or underestimations – which is obvious. This effect is visible in Figure 25 for February 2015. The strong increase in bias and RMSE for Feb.'15 of the original forecast is due to snow cover on the PV modules, which is not represented in the model. This systematic overestimation by the forecast model is being effectively compensated by the error persistence adaptation – which could be a suitable application for this approach.

Generally it became clear, that an error persistence based forecast adaptation might improve the forecasts only in a very short forecast horizon (in our case 1h). A more advanced forecast adaptation by error persistence would be possible if we take into account the potential performance of the PV system under clear sky conditions and set the error into relation to the maximum possible power at each time step.

3.7.2 Forecast adaptation based on error movement vectors

In order to assess the possibility of identifying error movement vectors (as described under 2.6), the individual normalized error ε for each time step needed to be estimated and visualized on a map. The data points representing the error at each reference system were calculated and spatial referenced, while the points in between were interpolated. For each hour of the two years under survey, a map similar to Figure 26 is available.

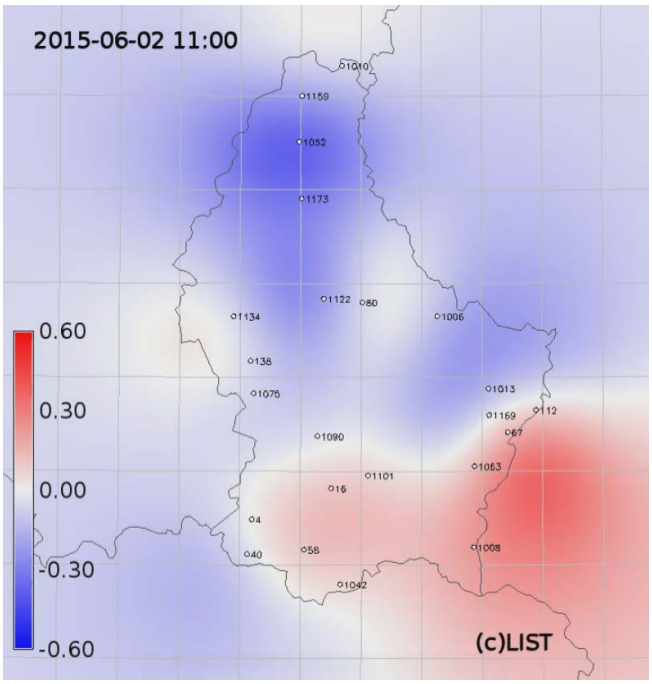


Figure 26 – map illustrating the regional deviations of power forecast from the measurement of reference systems

The error map above shows the distribution of deviations over the region. The forecast did, for the PV reference systems in the South and South-East of Luxembourg, overestimate the PV power (red), while for the North and middle of the country, the power was underestimated (blue).
The individual hourly error maps are sequentially concatenated to create a video sequence to analyse the changes of the occurring error distribution over time. Monthly video sequences were screened in order to evaluate the possibilities to identify and track error movements on specific days.

Figure 27 shows a subsequent series of pictures from such a video sequence. It shows areas of over- and underestimation during the individual hours, but unfortunately, these areas change in most cases relative drastically from one hour to the next. Areas of underestimated forecasts, as for example for South, East and West of Luxembourg at 13:00, might change into overestimation at 14:00 – without previous indication. Trends become seldom visible and are in most cases not persistent. To be able to detect trends in error movement automatically, the progressing areas need to be relatively clearly identifiable and trends would need to persist at least over two time steps to identify speed and direction. Forecasting those trends would only bring a clear advantage, if the trends continue further over the two time steps, in best case constantly in speed and direction. Unfortunately, this is in the large majority of days not the case.

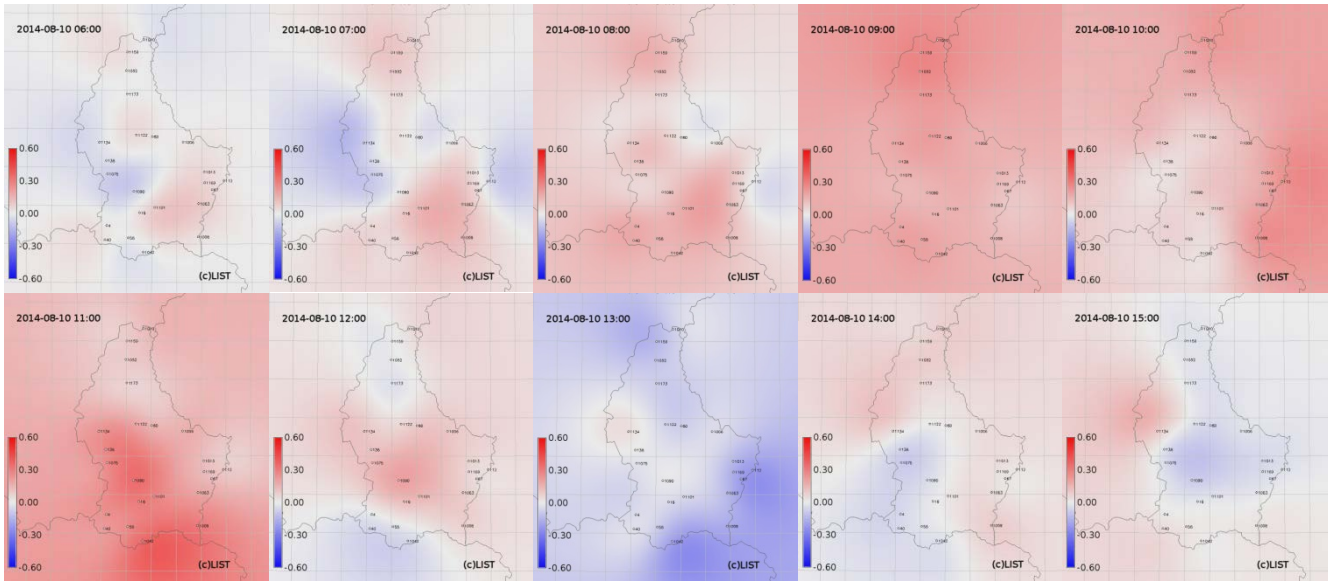


Figure 27 – picture series from an error map video sequence for the 10.08.14 06:00 until 15:00

For only few days, as e.g. shown in Figure 28, error movements are relatively clearly identifiable. At 05.09.15, the forecasts were relatively well suiting around 08:00 (indicated by pale colours) – the following hours show underestimations of the forecasts in the southern part of Luxemburg and overestimations in the North. The area of overestimation sweeps over the forecast area from North to South within 4-5 hours. Such movements could be identifiable and might be forecasted in the short term future, but even here, an improvement of the forecast might only be possible 1-3h ahead.

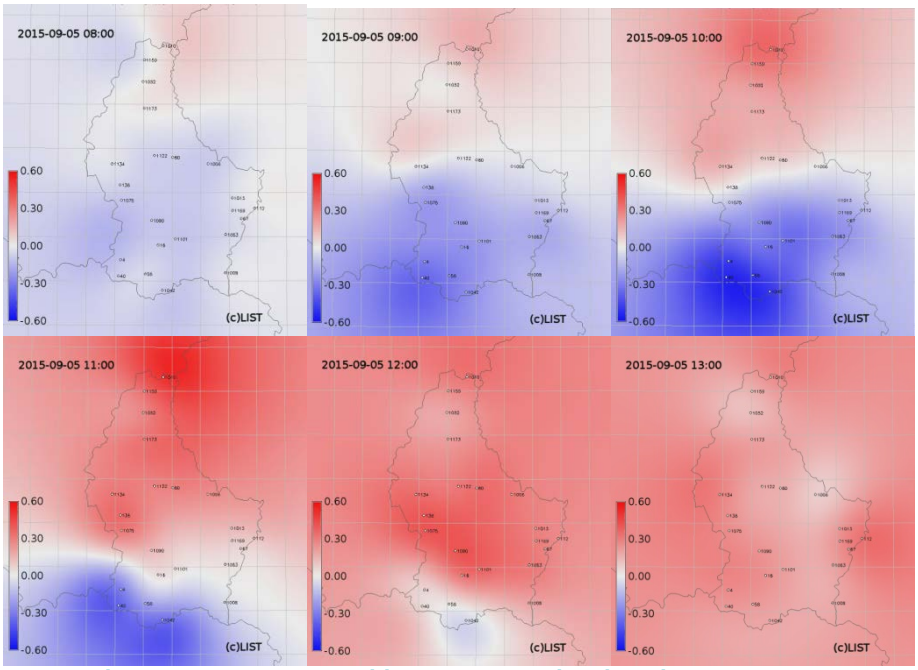


Figure 28 – error maps video sequence showing clear movement of areas of different deviations

The conclusion of the analysis of video sequences of error maps is therefore negative, concerning the approach of forecast improvements by “error movement vectors”. Some days show forecast error patterns where the described method might make sense, but most of them not. The single error maps change too drastically from one time step to the next. This leads also to the thesis, that this approach could be more promising at higher time resolutions (e.g. 15 min), which would increase the potential set of pictures that serve as basis to identify movements. The technical possibility to apply this approach would surely increase with the resolution in time, but the absolute forecast horizon for which the method might improve the forecast remains limited (1-3h ahead).

3.8 Results of the upscaling to the regional scale (Luxembourg)

The final result of the whole PV power forecasting algorithm is a dynamic and regionalized power forecast. Based on the upscaling procedure (2.7), using a list of all PV systems installed in the grid of Creos (and later also Sudstroum), an artificial power forecast for each system is generated. These powers could be aggregated on different regional or technical scales, if the information would be available, e.g. per street, per village, per commune or per transformer station. Currently, the expected PV power is aggregated on communal level, as shown in Figure 29.

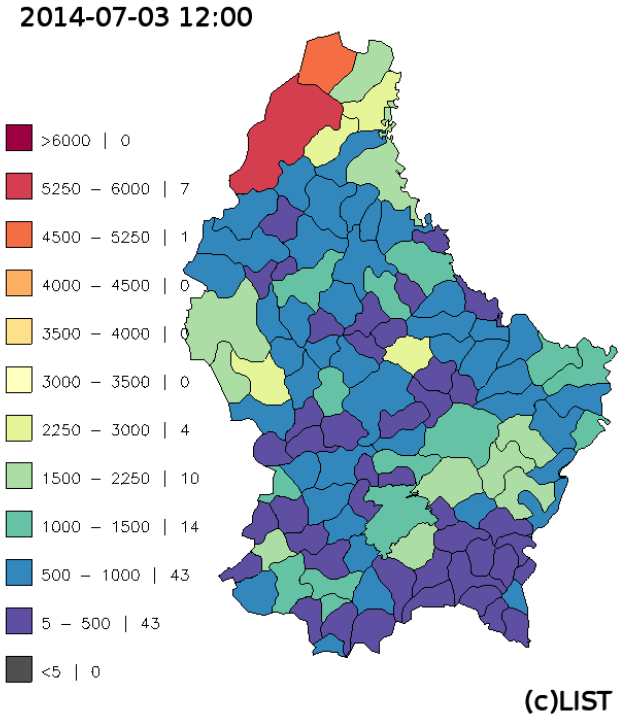


Figure 29 – forecasted PV power for 03.07.2014 12:00, aggregated per commune

The visualization of the predicted PV power on communal level gives a rather divers picture, as the differences in installed power per commune are relatively high (compare Figure 8 on page 19). Therefore, the high differences in between the actual PV production of the communes are mainly due to the installed nominal power and only to a minor extend caused by different weather conditions. The commune of Wincrange for example, illustrated in deep red on the map, has by far the largest installed PV capacity with 6'599 [kW_p] nominal power and therefore sticks out of nearly every map, relatively independent of the irradiation.

More meaningful is the dynamic dimension of the forecast. Based on these single pictures, video sequences have been produced to illustrate the daily variation in PV power over the two year periods. Figure 30 give a good example for the 03.07.2014.

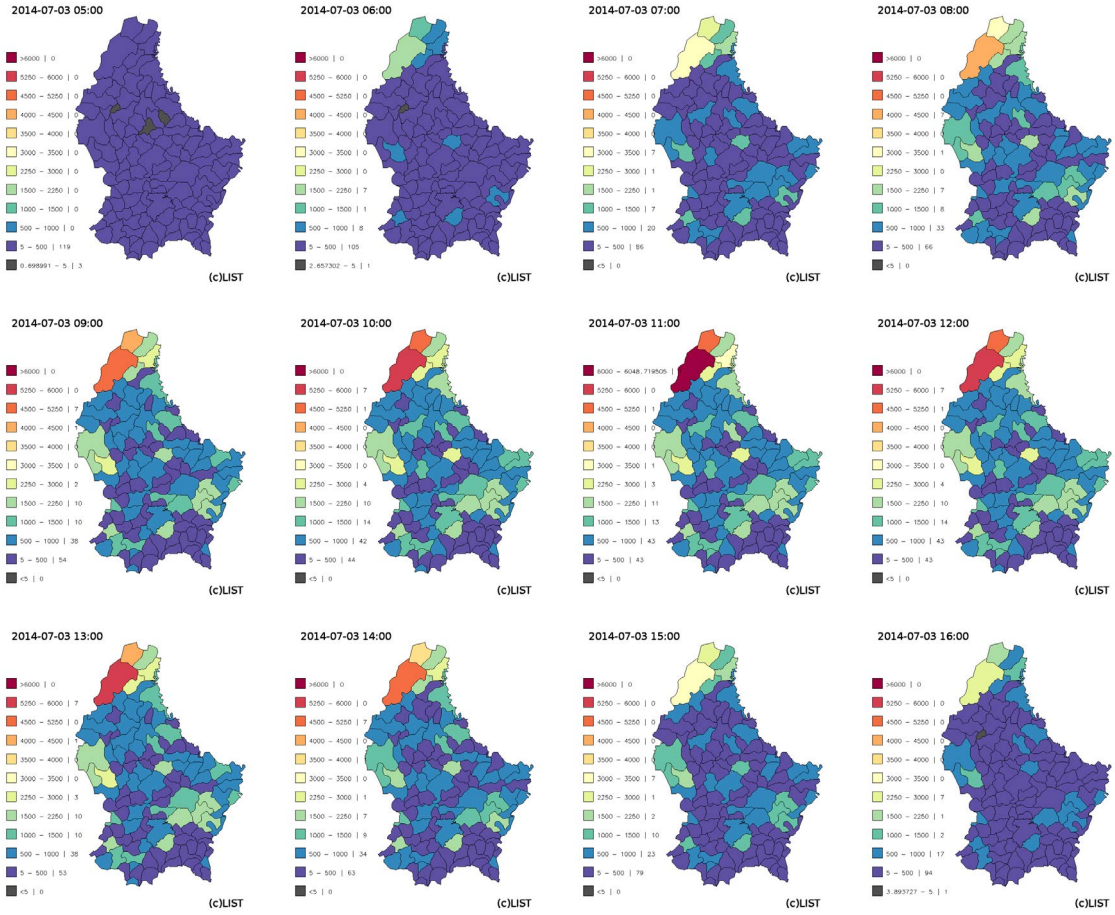


Figure 30 – sequence of hourly maps, illustrating the regionalized forecast of expected PV power, aggregated on communal level

4 Outlook

Throughout the development and evaluation of the model, several aspects have been discussed that might be adapted and improved in a future development of the forecasting scheme.

Parts of the difficulties in finding suitable reference systems were linked to the large size of the PV systems equipped with smart meters. With the current deployment of the smart meters in whole Luxembourg, more suitable reference PV systems will be available to act as a future reference installation. It should then be possible to choose small installations of about several kW_p, which are much less complex and would enable a more precise characterisation and modelling. As our results show, that the concept of synthetic system profiles is leading to comparable performance as the detailed system profiles, this would enable us to increase the amount of reference systems at relatively small effort. The resulting higher spatial resolution could also lead to larger opportunities in adapting the forecast by the currently abandoned approach of error movement vectors.

As the irradiation forecasts are only available in hourly time steps, an increase of the temporal resolution on forecast side is only possible by interpolation. The measurement values of reference systems are already available in 15 min time steps. As explained under 3.7.2, this might increase the potential for forecast adaptations by the approach of “error movement vectors”, although the absolute forecast horizon, which the approach could eventually improve, would remain the same.

A more advanced forecast adaptation by ”error persistence” would be possible if we take into account the potential performance of the PV system under clear sky conditions and set the error into relation to the maximum possible power at each time step.

Further, some calibration factors could evtl. be improved by a more detailed analysis of the individual forecast results as the calibration procedure was kept the same for all the systems. Also, for some few

systems, it might be advantageous to apply seasonal changing calibration factors – but for most systems no clear dependency of the bias on seasons is observed.
Furthermore, it depends on the specific purpose for which the forecasts model would be used, if it might make sense to change the aggregation level of the up-scaled power forecast. For an energy provider / retailer the regionalized forecast would even not be necessary, so the forecast can be aggregated on the level of their customers. For grid operators, if the necessary information on the localisation of individual PV systems is available, the forecast could be aggregated on level of transformer stations or street level.

5 Communication and Dissemination

Communication and dissemination takes a considerable amount of time in a research project. The project itself, its approach and its intermediate and final results are communicated on several levels:

Technical committee
Two technical committees were held, where the current state of development, the approach itself and the data availability and requirements were discussed with the local stakeholders / experts. The first technical committee was more focused on the presentation of the approach and the exchange on the availability of the necessary data on PV-systems in Luxembourg.
In the second meeting, intermediate results were presented and the LIST got feedback from the stakeholders on their specific interests in forecasts in the role of a grid operator, energy provider/retailer or green electricity producer. The stakeholders stated that the presented modelling approach (without the adaptation based on reference systems) would already be a good progress as compared to their current practice.

A third and final technical committee is planed after the holiday season, in order to present the final outcomes of the project and to further discuss the potential and challenges of an implementation of the forecasting algorithm on side of the grid operator Creos and the local main energy provider Enovos.
(Date, to be confirmed: 20.09. / 21.09. / 27.09.2016)

Conference Proceedings & Fairs
The project and its intermediate results were presented at the European PV Solar Energy Conference **"PV SEC" 2014 in Amsterdam**, in front of the leading international experts on PV power forecasting. The team received positive feed-back and got very interesting suggestions for its further development.

Figure 31 – PV-Forecast project presented at the PV SEC 2014 in Amsterdam

LIST presented the project at their booth at the **"Hannover Faire" 2015**, incl. a short film on two energy related projects at LIST – PV Forecast was one of them.

Publication targeting the general public
The publication of intermediate results has also been made available to the interested public in Luxembourg, via the engineering focused magazine "revue technique" in 2015.

Scientific peer-reviewed journal publication
A publication in a peer-reviewed scientific journal is under preparation. Currently, the team is planning to submit in the “Applied energy” journal.

Website & social media
On the website of the LIST (and Tudor, before) the project has been presented. Further, the short film, including the PV-Forecast project, was made available on some social media platforms (e.g. youtube.com)

6 Bibliography

De Soto, W., S.A. Klein, and W.A. Beckman. "Improvement and validation of a model for photovoltaic array performance." *Solar Energy*, 2006: 78-88.

Hogan, Robin. *Radiation Quantities in the ECMWF model and MARS.* London: ECMWF, 2015.

ILR. *Institut Luxembourgeois de Régulation.* n.d. www.ilr.lu (accessed 08 09, 2016).

Khalil, Samy A., and A.M. Shaffie. "Performance of Statistical Comparison Models of Solar Energy on Horizontal and Inclined Surface." *International Journal of Energy and Power (IJEP)*, 2013: 8-25.

Kleissl, Jan. *Solar Energy Forecasting and ressource Assessment.* Oxford: Academic Press, 2013.

Lorenz, Elke, Thomas Scheidsteger, Johannes Hurka, Detlev Heinemann, and Christian Kurz. "Regional PV power prediction for the improved grid integration." *PROGRESS IN PHOTOVOLTAICS: RESEARCH AND APPLICATIONS*, 2011: 757-771.

Minette, Frank, Daniel Koster, and Oliver O'Nagy. "PV-FORECAST: REGIONALISED FORECASTING OF ENERGY PRODUCTION FROM PHOTOVOLTAIC AND THEIR DYNAMICS BY A COMBINED APPROACH OF MODELLING AND REAL TIME MEASUREMENTS OF REFERENCE SYSTEMS." *European PV Solar Energy Conference 2014.* Amsterdam: WIP Munich, 2014. 8.

Olmo, F.J., J. Vida, I. Foyo, Y. Castro-Diez, and L. Alados-Arboledas. "Prediction of global irradiance on inclined surfaces from horizontal global irradiance." *Energy*, 1999: 689-704.

power one. "TRIO-27.6-TL Datasheet." n.d.

Annex

Two separated additional files represent the annex of this report and are delivered with it:

A) Annex document, containing the evaluation criteria plot over 2014 and 2015 for all reference systems
B) The publication : Minette, Frank, Daniel Koster, and Oliver O'Nagy. "PV-FORECAST: REGIONALISED FORECASTING OF ENERGY PRODUCTION FROM PHOTOVOLTAIC AND THEIR DYNAMICS BY A COMBINED APPROACH OF MODELLING AND REAL TIME MEASUREMENTS OF REFERENCE SYSTEMS." *European PV Solar Energy Conference 2014.* Amsterdam: WIP Munich, 2014. 8

PV-FORECAST: REGIONALISED FORECASTING OF ENERGY PRODUCTION FROM PHOTOVOLTAIC AND THEIR DYNAMICS BY A COMBINED APPROACH OF MODELLING AND REAL TIME MEASUREMENTS OF REFERENCE SYSTEMS

Frank Minette, Daniel Koster, Oliver O'Nagy
Centre de Recherche Public (CRP) Henri Tudor,
29 avenue John F. Kennedy, L-1855 Luxembourg-Kirchberg
frank.minette@tudor.lu, daniel.koster@tudor.lu, oliver.onagy@tudor.lu

ABSTRACT: The share of decentralized, fluctuating and not demand side controlled energy sources, such as wind power and photovoltaic (PV), is constantly increasing and will represent a major part of the near future energy mix. Precise predictions of the energy production from these sources will be crucial to ensure grid stability and reliable energy trading. Accurate PV power predictions are not yet widely implemented, but are gaining more importance. The prediction accuracy for larger regions is already very promising, but the drawback is the accurate forecast for smaller regions.
The PV-Forecast project will improve the forecasting performance at higher spatial resolutions by using a combined approach of predictive modelling of the PV power and the monitoring of reference systems in order to increase the accuracy for (spatially) higher resolutions. The approach described in this paper isn't fully developed yet, but the results presented seem to be promising to that end, that the site specific forecasts for the reference systems show good trends, compared to measurement values. The absolute errors normalized to the nominal power of the systems, for clear weather conditions, do lie under 10%. At cloudy conditions, the forecast performance is substantially lower, due to inaccuracies in the irradiance predictions, as shown in this paper.

Keywords: forecast, power, reference PV systems, statistical model, monitoring

1 PURPOSE OF THE WORK / INTRODUCTION

The idea behind this research project is the investigation of possibilities for precise PV power predictions for a small region or a small country by a combined approach of predictive modelling of the PV power and the monitoring of reference systems. Beside the scientific findings, the project would deliver a PV power prediction model for the Grand Duchy of Luxembourg (2'586 km²), adapted to the regional situation, which could be implemented as an online, real-time tool in a final development step. The project would thereby contribute largely to the future integration of PV into the local power grid and energy mix.

2 STATE OF THE ART

Over the last 10 years, the solar irradiance modelling progressed fundamentally and developed application oriented approaches e.g. for energy applications [5]. Beginning from purely meteorological models for solar irradiance prediction, the model developers added physical models that represent PV systems and simulate their performance under the modelled irradiance and temperature conditions. In order to evaluate and calibrate the models, measurement data from real PV systems are being used [3], [4].
Recent studies proved the feasibility of regional PV production forecasts based on intra-day and day-ahead irradiance forecasts given by the ECMWF [3], [7]. The forecast accuracy for single PV installations might vary and is not the purpose of these modelling efforts. The accuracy for larger regions is already very promising (13% for the region of Germany) [4], which proves that these kind of models can generally provide the necessary forecast data. The results show, that local temporal mismatches of predictions and real measured production data, balance out over larger areas. This means, in reverse, that predictions for smaller limited areas show higher inaccuracies. It is especially this last point (small area, i.e. Luxembourg) which is the focus of our work.

3 TECHNICAL APPROACH

The following chapter will explain the technical approach behind the forecasting model shown in Fig. 1.

3.1 Overview of the technical approach

The approach of the PV power forecast within the "PV-Forecast" project is a combination of geo-referenced irradiance and ambient temperature forecast data from numerical weather prediction (NWP) models of the European Centre for Medium-Range Weather Forecast (ECMWF) with measurement data of reference PV systems distributed over the region. The irradiance forecast data, that can be retrieved from the ECMWF web servers, are processed on our side in order to obtain the irradiance in plane for a matrix of predefined PV system orientations and inclinations and for a number of given PV systems that serve as references. For these reference PV systems, of which measured PV power in a temporal resolution of 15 minutes is available and that are distributed over the whole country of Luxembourg, an individual irradiance forecast in plane of the PV modules is being considered. Based on individual models for each reference system, representing their technical characteristics, a power forecast for those installations is being generated. The predicted power of the reference systems is compared to their measured generated power with the aim to obtain the base data for a self-adapting model, in the end.

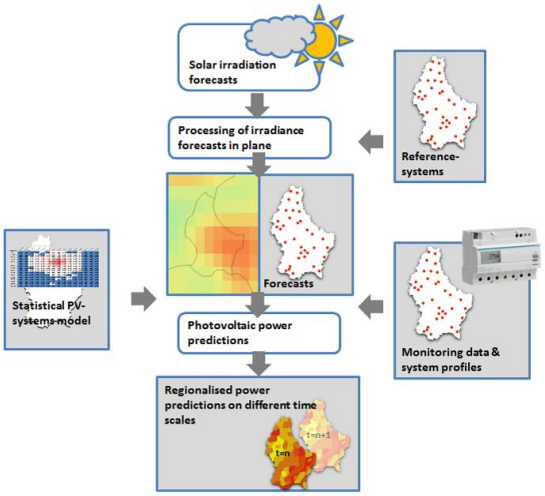


Fig. 1: scheme of the forecasting approach, combining modeling data and statistical information (left hand side) with a feed-back loop from PV reference systems (right hand side)

A statistical model of the spatially distributed PV installations in the whole country is then used in order to model the generated power for three days on an hourly basis over the whole region. Currently, the spatial resolution of this PV power forecast for whole Luxembourg, is the same as the ECMWF irradiance forecast data (0.125° x 0.125°).

3.2. Solar irradiation forecast

For the meteorological part, the PV power forecast is based on a combination of geo-referenced irradiance and ambient temperature forecast data from numerical weather prediction (NWP) models, elaborated by the European Centre for Medium-Range Weather Forecast (ECMWF). Recent studies proved the feasibility of regional PV production forecasts based on intra-day and day-ahead irradiance forecasts given by the ECMWF [3], [7]. The solar radiation data used here is the "surface solar radiation downward" (ssrd, parameter 169 available in grib or netcdf format). This parameter is the incident solar radiation or the so called global horizontal irradiance (GHI) [9].
The solar radiations as well as the temperature forecasts are available at an hourly base up to 72 hours (three days) in advance and with lower time resolution from 72 to 240 hours (10 days). These forecasts are updated twice a day. The forecast beginning from midnight 00:00 is available 8 hours later (at 08:00) and the forecast for noon 12:00 is available at 20:00. The grid resolution (i.e. longitude and latitude) is 0.125° by 0.125° for one cell. This results for Luxembourg in a North-South resolution of about 14 km and in an East-West resolution of 9 km. Considering that Luxembourg has maximum territorial dimensions in North-South direction of 82 km and in East-West of 57 km (with a total area of 2'586 km^2), this results in approximatively 20 cells. One could easily imagine that clouds have a much greater impact here compared to greater regions where the PV power forecasting gives already good results [7]. This smallness of Luxembourg is one of the challenges in the PV-Forecast project.
ECMWF uses a bi-linear interpolation technique. It uses a 2 x 2 grid points closest to the selected interpolation location and takes a weighted average to arrive to the interpolated value [17].
In the definition of the project was foreseen to make a regionalisation of the forecasting data. But it turned out quickly by considering the maximum available resolution of 0.125° that it makes no sense to try to interpolate this already interpolated forecasting data further for getting more areal resolution. Furthermore, in [7], this task was abolished.

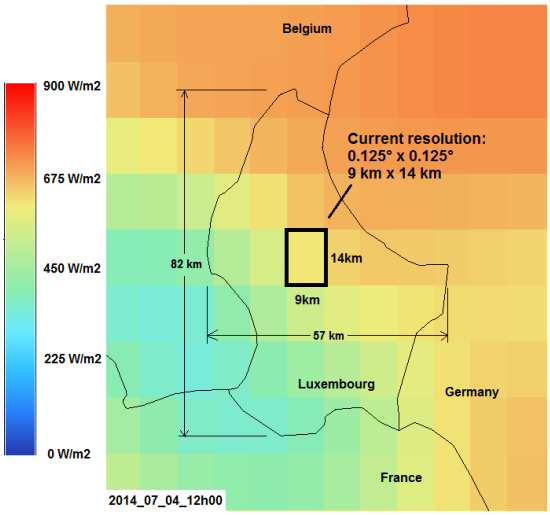


Fig. 2: Current available resolution for forecasting data from ECMWF compared to the size of Luxembourg

3.3. Conversion of the global horizontal irradiation forecast value into inclined plane irradiation

There are different methods for the calculation of the inclined global irradiation depending on the quantity and quality of available parameters. A rather simple method with a relatively good accuracy published in 1998 by Olmo et al. [6] is chosen. A recent study from Khalil et al. [8] points out the simplicity and accuracy of the method developed by Olmo et al. The advantage of this method is that it requires only the global horizontal irradiation GHI, the sun position (elevation to the horizon and incident angle to the inclined surface) and the albedo of the environment.

Calculation of the inclined global irradiation [6]:

$$G_\Psi = GHI \exp(-k_t*(\Psi^2 - \Psi_H^2))*F_c$$

Where:
G_Ψ is the inclined global irradiation [Wh/m²];
k_t is the clearness index [-];
Ψ is the incident angle [rad] of the sun to the inclined plane;
Ψ_H is the incident angle [rad] of the sun to the horizontal plane;
GHI is the global horizontal irradiation [Wh/m²];
Fc is the correction factor [-] that is taking the ground reflection (albedo) into account.

The first 3 reference PV systems and the location where the pyranometers are installed are on flat roofs with bitumen, that have a white mineral surface. The albedo factor of 0.35 seems to be appropriate for the calculations during the dry summer months.

A comparison of the calculated values based on GHI measurements with the measured inclined global irradiation (both measurements done with a Kipp & Zonen CM11 pyranometer) shows the relatively good fitting of the trends of both curves, measured and calculated (Fig. 3 and Fig. 4). However it can be seen that there are deviations of the calculated (blue) values above measured (red) values, that causes an overestimation of the calculated inclined irradiation for clear days. (Fig. 3) The total deviation for this specific clear day is + 9%.

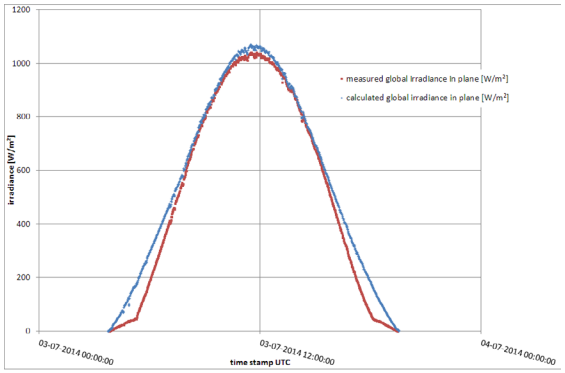


Fig. 3: On a clear day in July 2014, measured inclined global irradiance (elevation 30°, azimuth 180° in red) and the calculated inclined global irradiance (blue) using the method developed by Olmo et al. [6].

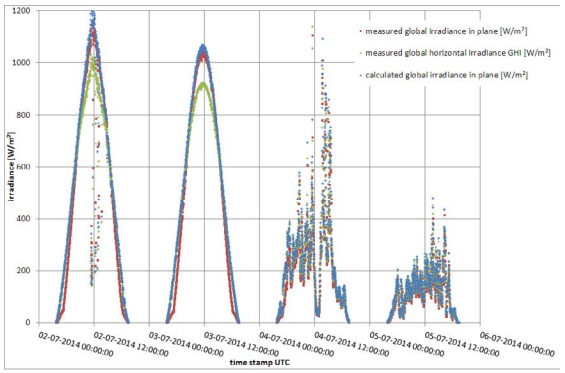


Fig. 4: overview on four days in July 2014 with two clear and two cloudy days.

Comparing the pyranometer measurement values at inclined orientation of 30° elevation and 180° azimuth with the calculated values based on GHI measurements, it can be seen, that the variance remains low even under different cloud conditions over a month. Fig. 5 shows the data of July 2014, that was mainly clear (six very cloudy days, five very clear days and twenty predominant clear days). It can be seen, that the calculated values have a low scatter but a systematic deviation in relation to the measured values probably due to the high daily irradiation during the measurement period.

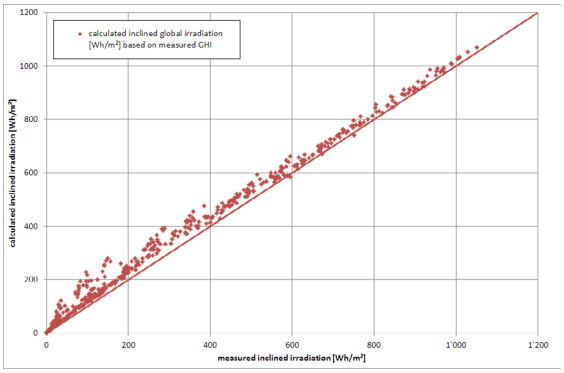


Fig. 5: July 2014, calculated inclined global irradiation based on GHI (y-axis) over measured inclined global irradiation (x-axis).

The described calculation method is now applied to the forecast value of the global horizontal irradiation GHI (ECMWF parameter identifier: 169; ssrd). This parameter represents the global horizontal irradiation with the unit [J/m²] (GHI) [9], [10]. There is already a large variance between the ground measured and the forecast data as it can be seen in the Fig. 6. The calculated values are scattered broadly around the measured values.

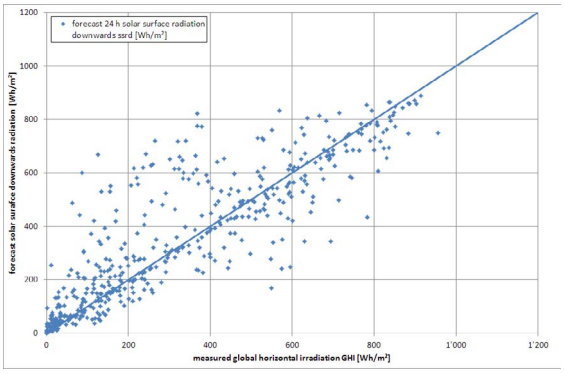


Fig. 6: 24 hours forecast ssrd values (y-axis) over ground measured GHI values (x-axis).

The main reasons for this broad scattering can be found in

a) forecast accuracy,
b) the same forecast area (9 km x 14 km) overlays different microclimate areas, that might have an effect on the ground irradiation,
c) inaccurate prediction of cloud distribution and movement in the forecast area (causing variance of the irradiation in both directions: enhancement due to reflection and cloud shading as well as an overlapping of both effects.

The points b) and c) are addressed in this project in order to improve the local forecast

3.4. PV-Power forecasts for the reference systems

It is foreseen to implement a number of reference PV systems into the model. The objective is to obtain a real time feedback from PV systems which are already measured in a high temporal resolution (smart metering) without additional effort, in our case every 15 minutes, in order to be able to adapt the short term forecasting based

on this information. At this moment in time, three PV systems are currently characterized and measured, but the aim is to include 30 – 40 systems of a good spatial distribution over Luxembourg.
Each reference system is characterized according to their technical parameters, such as type of PV modules, inverters and cable lengths, inclination and orientation of the PV arrays, age of the installation, location and array set up. After a site visit, the technical model is set up for each system using the following approach:
In the current version, on the level of reference systems, the only pre-photovoltaic losses which are being considered are the reflections on the module surface (no losses due to snow cover or dust). For simplicity reasons, the "ASHRAE incident modifier" was used to account for the reflection losses, although knowing for the limitations of the approach at high incident angles (angle between the normal on the PV plane and the sun rays). To attenuate this effect, the reflection losses at incident angles above 80° have been set to a fixed value of 30%. In a next development step, this approach might change.
The PV modules efficiency has been calculated based on their performance at standard test conditions (STC) and is temperature corrected by the temperature coefficient for the power at maximum power point (MPP). In order to estimate the modules temperature at operating conditions, the following formula was used [3]:

$$T_{module} = T_{amb} + \gamma\, G_{\psi}$$

T_{module} = module temperature [°C],
T_{amb} = ambient temperature [°C],
γ = 0.020 for free standing PV systems,
0.056 for building integrated PV systems,
G_{ψ} = inclined global irradiance [W/m²].

When calculating the performance based on nominal power, we need to consider the degradation of the modules. During the first year of operation, degradation losses of 2,5% were accounted for [11], [12], [13] and further 0.5% per year for each additional year of operation [14], [12], [11], [15], which is in good correlation with values found in the mentioned references. Also the mismatch losses of the PV arrays were taken into account by a simple lump sum reduction of 2.5% [18].

In a first version, the inverters performance was considered by applying the European inverter efficiency, as the inverters of the currently integrated three PV systems are rather stable over their operation range. In a next version, when integrating additional reference systems, this aspect will be integrated in dependence on the current part load operation point.

3.5. Statistical model

In order to reflect the circumstance that not all the PV installations have the same inclination and orientation, a so called statistical model was elaborated.
This model is based on statistical data of 6'308 subsidised PV installations out of the Luxembourgish area, representing app. 40% of the installed PV capacity. Only known parameters are considered: installed power and type of module, inclination and orientation angles. It is formed by a distribution matrix with percentages of installed PV-power for the elevation (vertical: 0° to 90° in steps of 5° => 19 classes) and the orientation (horizontal: -180° to 180° in steps of 5°, south = 0°, => 73 classes) for each of the three available module types (mono-Si, poly-Si, other). Each statistical technology matrix is formed by 73 x 19 = 1'387 values. As the mono and poly technologies represent more than 98% of the installed power in Luxembourg, only these two where used in the forecasting of the power. Fig. 7 shows a compressed example of this distribution matrix considering all PV module types together.
One can observe that most of the PV installation are south oriented with an elevation of about 30°.

elev/orient	-180°<-165°	-170°<-120°	-120°<-90°	-90°<-60°	-60°<-45°	-45°<-30°	-30°<-15°	-15°<15°	15°<30°	30°<45°	45°<60°	60°<90°	90°<120°	120°<150°	150°<180°
0°≤5°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
5°≤10°	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.6%	0.3%	0.1%	0.0%	0.1%	0.0%	0.0%	0.0%
10°≤15°	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%	1.6%	0.1%	0.4%	0.2%	0.0%	0.0%	0.0%	0.0%
15°≤20°	0.0%	0.0%	0.0%	0.4%	0.4%	0.9%	1.4%	3.3%	1.7%	1.9%	0.5%	0.4%	0.0%	0.0%	0.0%
20°≤25°	0.0%	0.0%	0.0%	0.3%	0.7%	1.4%	5.9%	8.1%	3.7%	2.9%	0.9%	0.5%	0.3%	0.0%	0.0%
25°≤30°	0.0%	0.0%	0.0%	0.1%	0.4%	1.4%	4.6%	7.7%	4.5%	2.3%	0.8%	0.6%	0.1%	0.0%	0.0%
30°≤35°	0.0%	0.0%	0.0%	0.3%	0.2%	1.4%	3.4%	9.7%	3.6%	2.1%	0.7%	0.5%	0.1%	0.0%	0.0%
35°≤40°	0.0%	0.0%	0.0%	0.2%	0.3%	0.7%	1.4%	4.6%	1.0%	0.7%	0.8%	0.6%	0.2%	0.0%	0.0%
40°≤45°	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	0.5%	1.5%	0.4%	0.3%	0.2%	0.1%	0.0%	0.0%	0.0%
45°≤50°	0.0%	0.0%	0.0%	0.0%	0.0%	0.1%	0.1%	0.6%	0.1%	0.2%	0.2%	0.0%	0.0%	0.0%	0.0%
50°≤55°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
55°≤60°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
60°≤70°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
70°≤80°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
80°≤90°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
90°≤180°	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Fig. 7: Statistical distribution of the orientation and slope of 6'308 PV systems all over the country

For each forecasted cell in Fig. 2 (defined by longitude, latitude, forecasted solar radiation and amb. temperature) and for each element out of the matrix of Fig. 7 (elevation, orientation, PV technology, percentage of installed PV power, solar incident angle, …) we calculate hourly the generated PV energy and sum them up at the end. As the statistical matrix is built up with percentages of a certain PV technology, the calculated sum has to be multiplied by the corresponding installed power of this technology in this cell. By this, we obtain a regionalised forecast for each cell of the grid. The sum of all the cells reflects the generated energy for the whole region of Luxembourg for a given hour based on forecasted meteorological data.
If it is desired to transfer the model to another region this input matrix needs to be adapted to the local situation and the regional installed PV capacities.

3.6. Adaptation of the regionalized forecast by feedback from the reference systems

The final aim of a combined approach of modelling the expected PV power based on irradiance predictions and measurements from reference systems is to increase accuracy and / or spatial resolution. So far, our work hasn't reached the point to give an answer to the question whether, or to which extend this is possible. But we would like to develop and test some of the following approaches and assumptions:
Results from other research groups show, that local temporal mismatches of predictions and real measured production data, balance out over larger areas [7]. This means in reverse, that predictions for smaller limited areas show higher inaccuracies, which is of particular importance to Luxembourg. One of our assumptions is, that on the very short time scale (intra-day) and with appropriate number and distribution of reference systems, it should be possible to identify temporal and spatial mismatches of the forecasts in one part of the region and

correct them partially for the coming time steps (3-4 hours).
Comparable to the identification of "cloud motion vectors", a simplified approach could be imagined in order to detect the movement of inaccuracies over the region that are due to temporal mismatches in the predictions. Errors between predictions and the most recent historical forecast can be derived to error maps, in 15 minutes time resolutions. Differences between error maps for the time n and n-1 can indicate direction and speed of such an "error movement".
A study about the operation performance of a large PV system has already used a similar approach to model variations in the incident irradiation due to cloud movement [16].

4. RESULTS

4.1 Comparison of historical irradiation forecast values with measurements of four different locations

The first location that is compared is our outdoor testfield in Esch/Alzette ("PV-Lab"; South-west of Luxembourg) with the ground measurements using high accuracy pyranometers and sun-tracked pyrheliometer. The other three locations are medium sized free-standing PV-systems installed on flat roofs as described in chapter 4.2. The three systems are monitored with reference solar cells (two per system) with an estimated overall accuracy of about +- 10%
PV-Lab outdoor test field:
The total error between forecast and measured inclined global irradiation can be very high because of the large variation between forecast and ground measured data in combination with the low variation that results from the conversion of the GHI into inclined global irradiation. However the total error will not have a big effect on the PV electricity production, if the irradiation is low. The forecast delivers hourly values within forecast periods from 12 hours to 72 hours. A 24 hour forecast period is applied and compared with the measured values in order to receive a "useful" forecast error. Additionally, the deviations between the hourly calculated forecast and measured value are normalized to 1'000 Wh/m².

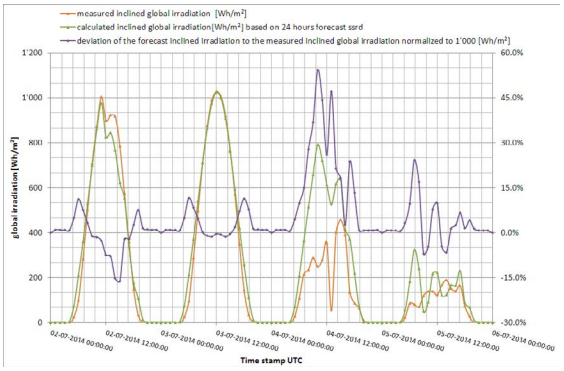


Fig. 8: four days in July 2014: measured inclined global irradiation (orange) and calculated inclined global irradiation (green) based on forecast ssrd. The violet curve shows the deviation between calculated forecast values and measured values normalized to 1'000 Wh/m².

The average of the hourly normalized deviation for the complete month of July 2014 is 9%. The total measured irradiation for July 2014 is 155 kWh/m² while the total forecast value is 186 kWh/m² which results in a difference of 32 kWh/m² or 17% between forecast and measurement.

4.2. Results of the power forecasts for three reference systems

Up to now, the first three reference systems have been chosen, characterized and their system parameter as well as the power metering data has been integrated into the modelling approach. Currently, the contacts to further approximately 100 systems is established out of which 30 to 40 systems will be chosen to serve as references.

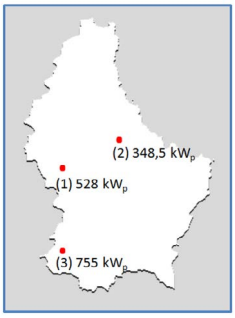


Fig. 9: position of the first three reference PV systems

The first three reference PV installations are free standing, crystalline silicon, large scale systems of the following characteristic:
1. nominal power: 528.00 kWp;
   slope: 15 ° (install. rack) +- 2 ° (roof inclination);
   orientation: 219 °
2. nominal power: 348.48 kWp;
   slope: 15 ° (install. rack) +- 5 ° (roof inclination)
   orientation: 158 °
3. nominal power: 755.04 kWp;
   slope: 20 ° (install. rack) +- 5 °(roof inclination);
   orientation: 158.5 °

4.2.1. Validating PV-model performance

Beyond the measured electricity, fed in at the injection point, we got access to the data acquisition server of the three installations and retrieved the measured irradiance data on site. Those values were measured by reference cells installed in plane with the PV modules.
In order to validate the quality of our PV system model, independent from the quality of the irradiance forecasts, we used the measured "irradiance in plane" from the reference cells as input for our model and compared the results for one month (July 2014) with the measured energy yield data from smart meters (see Fig. 10 as an example).
The four days show a comparably well correlation of the modelled PV yield with the measured data. We observe slight overestimations of the model at clear weather conditions and slight underestimations at cloudy conditions.

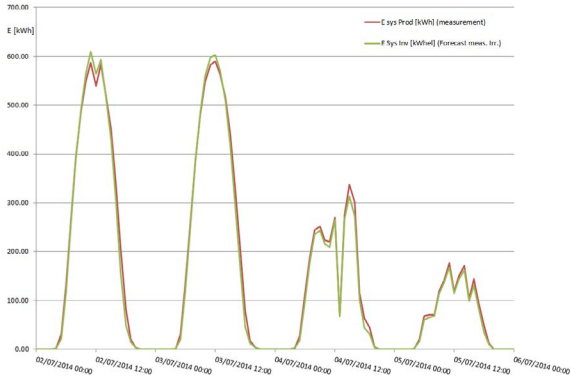


Fig. 10: comparison of measured energy yield data (red) of PV system (3) to modelled energy yield (green), using the irradiance measurements of reference cells as model input, for four days in July 2014

For reasons of data availability for the reference systems, the modelling results could until now only be evaluated for the time span of one month (July 2014). For this month, the deviation of the modelled yield (monthly sum) to the measured yield was -3.9 % of the measured value for system (1), -2.4 % for system (2) and -3.2 % for system (3), meaning an underestimation of the systems real performance.
The mean deviations of the modelled daily sums to the measured daily sums are -4.2 % for system (1), -2.4 % and -3.6 % for system (2) and (3).

Looking at hourly data, the picture becomes much more divers. The accuracy of the modelled hourly PV power, as compared to the measurement is being assessed by the relative error compared to the measurement and by the absolute error, normalized to the nominal power of the PV systems (both approaches are explained further down in 4.2.2.).

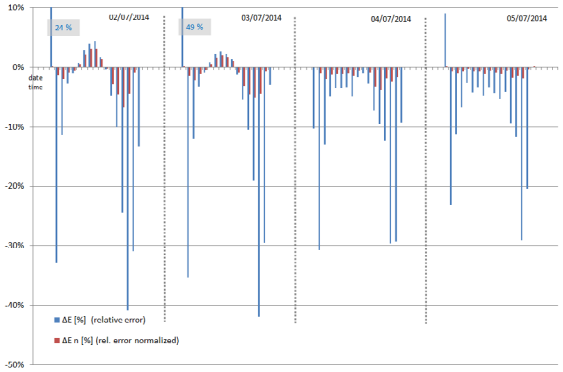


Fig. 11: comparison of the relative error of modelled hourly PV power (blue) of PV system (3) and the absolute error normalized to the nominal power of the PV system (red) (using the irradiance data from the reference cells as model input, for four days in July 2014)

Fig. 11 shows, that the PV model generally under-estimates the system performance, with higher deviations at high inclination angles. Especially the high inaccuracies in early and late hours over the day, at relatively low irradiations, result in a high relative error but being less important considering the low absolute energy values for those hours. This is illustrated by the red curve - the absolute error, normalized to the nominal power.

4.2.2. Validating the full forecast models performance

After having assessed the forecast accuracy of the irradiance forecasts incl. the processing of in plane data (4.1.) and the quality of the PV system model (4.2.1.), we assessed the full model, from irradiance forecasts to the PV power forecast for the reference systems by comparing it to the measured yields as well:

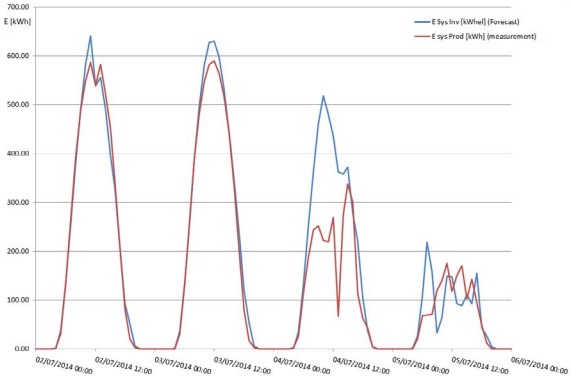


Fig. 12: comparison of measured energy yield data (red) of PV system (3) to modelled energy yield (blue), using the irradiance forecasts from ECMWF, for four days in July 2014

Fig. 12 illustrates a general tendency, which is clear from what we know about the irradiance forecasts already. The model seams to predict rather well the PV power at clear weather conditions (02/07/14 and 03/07/14), while at cloudy conditions (04/07/14 and 05/07/14) the predictions can deviate substantially from measurements.

Expressed in numbers, the picture is the following:
Also for July 2014, the deviation of the forecasted yield (monthly sum) based on solar irradiation forecasts to the measured yield was 10.2 % of the measured value for system (1), 13.7 % for system (2) and 11.0 % for system (3), meaning an overestimation of the real yield.
The mean deviations of the forecasted daily sums compared to the measured daily sums are 13.6 % for system (1), 17.6 % and 14.7 % for system (2) and (3). Those values are given just for reasons of completeness but should be taken cautiously, as the variance for daily sums is rather high. They mainly give an idea about the bias, which is (for July) an overestimation of the real system yield.

The final objective is obtaining reliable hourly forecast values, which is a much more difficult endeavour. Fig. 13 shows the relative error for the same four representative days as used before. The relative error has been calculated in relation to the real power of the system:

$$\Delta E\,[\%] = (E_{sys\ Inv} - E_{sys\ Prod}) / E_{sys\ Prod}$$

ΔE [%] = relative error of hourly PV power forecasts,
$E_{sys\ Inv}$ = hourly power forecast of the PV system behind inverter,
$E_{sys\ Prod}$ = hourly measured power of the PV system.

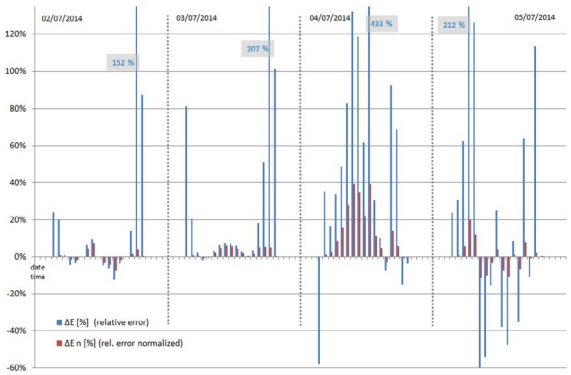


Fig. 13: comparison of the relative error of forecasted hourly PV power (blue) of PV system (3) and the absolute error normalized to the nominal power of the PV system (red) (using the irradiance forecasts from ECMWF, for four days in July 2014)

The Fig. 11 shows a rather good forecasting performance (relative error in blue) for days of clear weather conditions (2nd and 3rd of July), which is not surprising keeping in mind what we assessed before. On cloudy days, the results are more divers and worse in the overall general performance, what can be shown by the values for the 4th and 5th of July. Especially for the values at clear weather conditions, it became obvious that the relative error increases at high incident angles (in the morning and at evening) when the overall power is already rather low. Here, obviously, a relatively small absolute error leads to high relative deviations. Therefore, we assess the forecast values also by normalizing the absolute error of the forecasts to the nominal power of the respective PV system:

$$\Delta E_n\,[\%] = (E_{sys\,Inv} - E_{sys\,Prod}) / P_n$$

ΔE_n [%] = normalized absolute error of hourly PV power forecasts
P_n = nominal power of the PV system

Fig. 13 also shows this normalized absolute error (in red) and illustrates well that the performance of the forecast for days of good weather conditions is already quite promising (error generally below 10%). At cloudy conditions, the normalized absolute error can reach values of 50% at maximum in our sample of one month, but normally stays well below 35%.

5. CONCLUSION

Within this project, unlike most other existing models, the reference systems do not only serve as a basis to up-scale PV-production data for the forecast-area. The reference systems are being continuously monitored and deliver data for refinement of the forecasts.
The described approach contains a set of existing and proven models ranging from the processing of the irradiance forecasts, to the behaviour of the PV reference systems.
Using the model described by Olmo [6] to calculate the inclined surface irradiation out of the horizontal irradiation forecasts, performed well, but showed a slight positive bias, that we will further investigate as this behaviour doesn't correspond with the experiences of other research groups reported in the literature [8].
We assessed the performance of the PV model that calculates the hourly PV power of the reference systems and found that the model represents the actual performance rather well, although we had slight deviations which can be further reduced by calibration of the model.
When comparing the actual performance of the full forecasting model (using intra-day irradiance forecasts) for the three currently included reference systems, we find good correlation in time, at clear weather conditions. At cloudy conditions, PV power forecasts for the single site can deviate substantially from the real measured power, due to inaccuracies of the irradiance forecasts, mainly.
In order to correct our forecasts in the short term forecast horizons by the experience from the comparisons for the reference systems, we need a broader basis of reference systems, which we have not yet available. In the near future, the described adaptation approach can be further elaborated and evaluated for its performance.

6. OUTLOOK

The next step in the development of the described forecasting model is the choice, characterisation and inclusion of further PV reference systems (up to 40), that will allow for the calculation of the mentioned error maps. This will enable the evaluation of the technical part of the model for a higher number of systems. Also, the evaluation of the approach needs to be done in reference to monitoring data over a broader time period and for different forecasts horizons, 2- to 3-days ahead.
All the different reference systems will be analysed in order to identify systematic over- or underestimations by the technical part of the model, which we could minimize by calibration.
Once sufficient reference systems are included and error maps for a larger time period are available, the technique to identify error patterns, their direction and speed of propagation over the forecast area, will be further developed.

7. ACKNOWLEDGEMENTS:

Fondation Enovos for the financial support. CREOS, ENOVOS and their customers for providing measurement data of the reference PV systems, the Luxembourg Environment agency for providing statistical data, the ECMWF for giving access to the forecast data and our colleagues Christian Braun, Jonathan Hervieu and Markus Jonas (CRP Gabriel Lippmann) for the processing of the forecast data.

8. REFERENCES

[1] Mazhari E.; Zhao J.; Celik N.; Lee S.; Son Y-S; Head L., 2011, Hybrid simulation and optimization-based design and operation of integrated photovoltaic generation, storage units, and grid; Simulation Modelling Practice and Theory; Volume 19; Issue 1; pp. 463-481
[2] Braun M., Stetz T., Bründlinger R., Mayr C, Ogimoto K, Hatta H., Kobayashi H., Kroposki B., Mather B.,

Coddington M., Lynn K., Graditi G., Woyte A., MacGill I., 2011, Is the distribution grid ready to accept large-scale photovoltaic deployment? State of the art, progress, and future prospects; Progress in Photovoltaics: Research and Applications; Prog. Photovolt: Res. Appl.; John Wiley & Sons, Ltd

[3] E. Lorenz, T. Scheidsteger, J. Hurka, D. Heinemann, C. Kurz; 2010, Regional PV power prediction for improved grid integration; ISBN: 3-936338-26-4, 25th EUPVSEC 2010 - 5AO.8.1

[4] Lorenz, E.; Hurka, J.; Heinemann, D.; Beyer, H.G., 2009, Irradiance Forecasting for the Power Prediction of Grid-Connected Photovoltaic Systems; Selected Topics in Applied Earth Observations and Remote Sensing; IEEE Journal of , vol.2, no.1; pp. 2-10

[5] Dunlop, E. D.; Wald, L.; Súri, M., 2006, Solar energy resource management for electricity generation from local level to global scale; Nova Science Publishers; ISBN 1594549192

[6] F.J. Olmo, J. Vida, I. Foyo, Y. Castro-Diez, L. Alados-Arboledas, 1998, Prediction of global irradiance on inclined surfaces from horizontal global irradiance, Energy 24 (1999) 689–704

[7] Photovoltaic and Solar Forecasting: State of the Art, IEA PVPS Task 14, Subtask 3.1, Report IEA-PVPS T14-01: 2013, October 2013

[8] Samy A. Khalil, A. M. Shaffie, 2013, Performance of statistical comparison models of solar energy on horizontal and inclined surface , International Journal of Energy and Power (IJEP) Volume 2 Issue 1

[9] Alberto Troccoli, Jean-Jaqcues Morcrette, autumn 2013, ECMWF Newsletter 137

[10] Patrick Mathiesen and Jan Kleissl, 2011, Evaluation of numerical weather prediction for intra-day solar forecasting in the continental United States, technical report of the Department of Mechanical and Aerospace Engineering, University of California, San Diego

[11] Quintana, M.A. ; King, D.L. ; McMahon, T.J. ; Osterwald, C.R., Photovoltaic Specialists Conference, 2002. Conference Record of the Twenty-Ninth IEEE, Commonly observed degradation in field-aged photovoltaic modules

[12] Sakamoto, S., Oshiro, T., 3rd World Conference on Photovoltaic Energy Conversion, 2003, May 2003, Vol.2, pp.1888-1891, Field test results on the stability of crystalline silicon photovoltaic modules manufactured in the 1990s

[13] Dunlop, E.D., 3rd World Conference on Photovoltaic Energy Conversion, 2003, May 2003, Vol.3, pp.2927-2930, Lifetime performance of crystalline silicon PV modules

[14] Skoczek, A., Sample, T. , Dunlop, E.D., Progress In Photovoltaics, 2009(4), pp.227-240, The Results of Performance Measurements of Field-aged Crystalline Silicon Photovoltaic Modules

[15] De Lia, F.; Castello, S.; Abenante, L., Proceedings of the 3rd World Conference on Photovoltaic Energy Conversion, 2003, Vol.B, pp.2105-2108, Efficiency degradation of C-silicon photovoltaic modules after 22-year continuous field exposure

[16] P. Öchsner; M. Zehner; F. Lang; T. Rauscher; J. Weizenbeck; T. Weigl; G. Becker; G. Bettenwort; B. Giesler; T. Betts; R. Gottschalg, 28th European Photovoltaic Solar Energy Conference and Exhibition 2013, spatial modelling of grid connected PV plants with 3D irradiance values

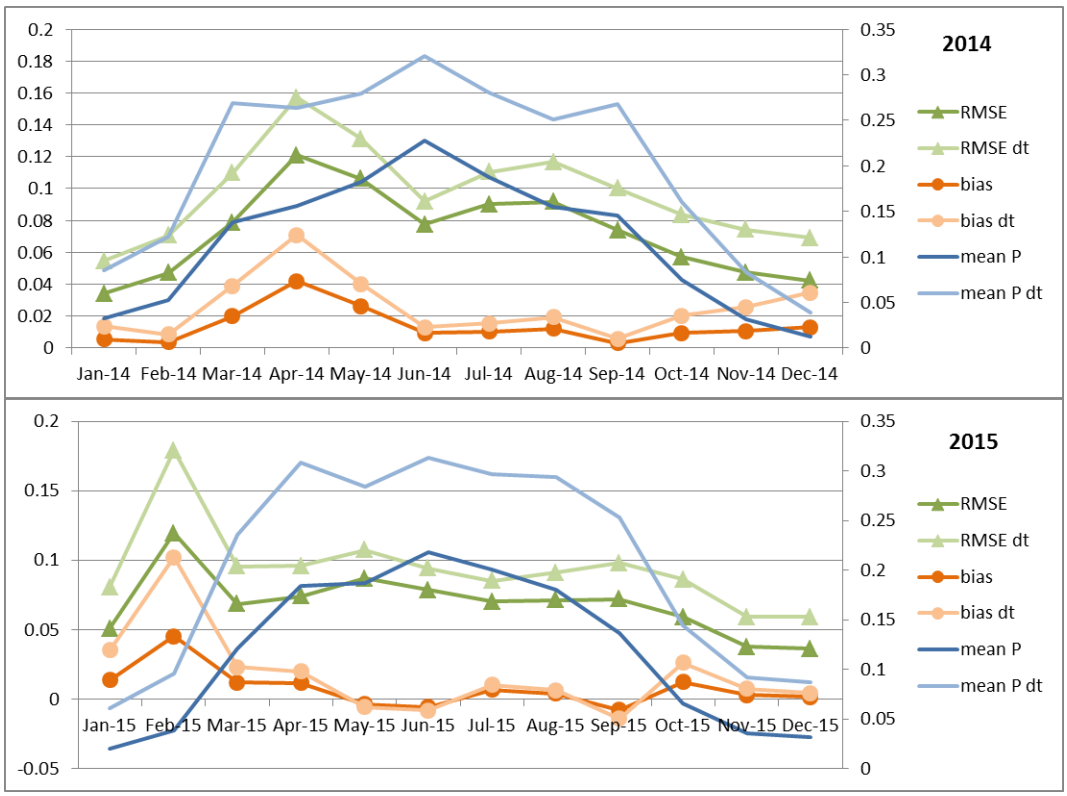
[17] Persson Anders, October 2011, User guide to ECMWF forecast products, 127 pages

[18] J. Webber and E. Riley, Mismatch Loss Reduction in Photovoltaic Arrays as a Result of Sorting Photovoltaic Modules by Max-Power Parameters, ISRN Renewable Energy Volume 2013, Article ID 327835, 9 pages

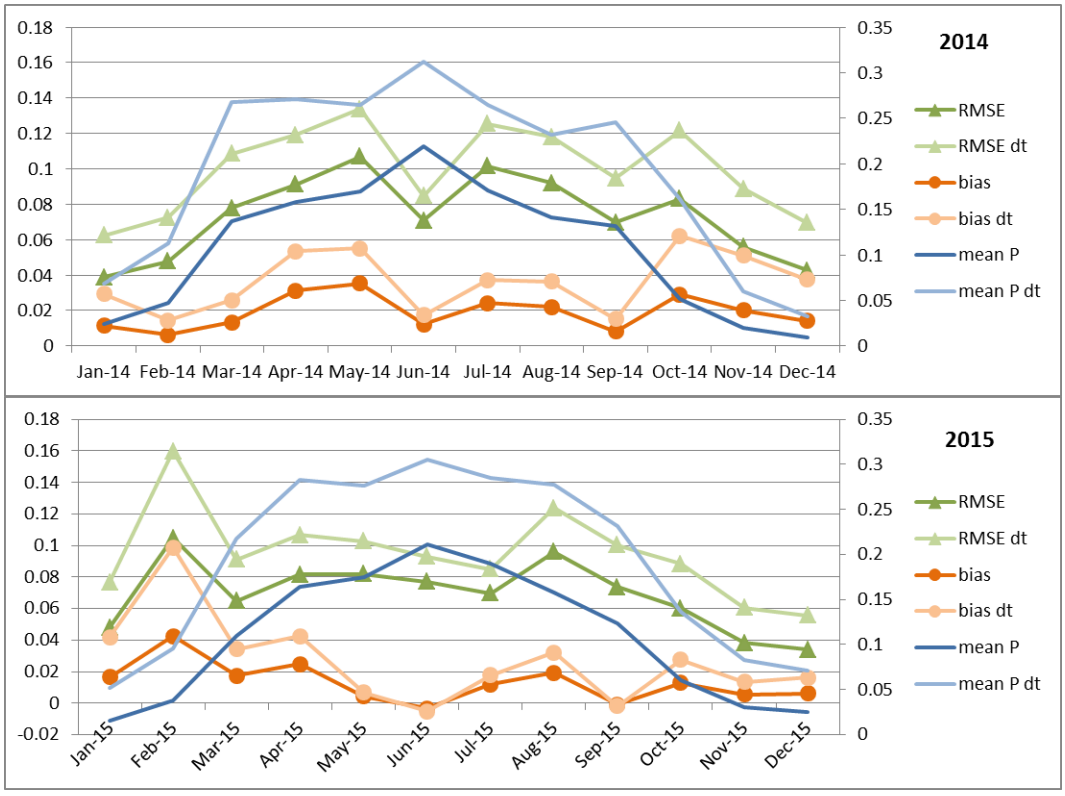
Annex

Graphs

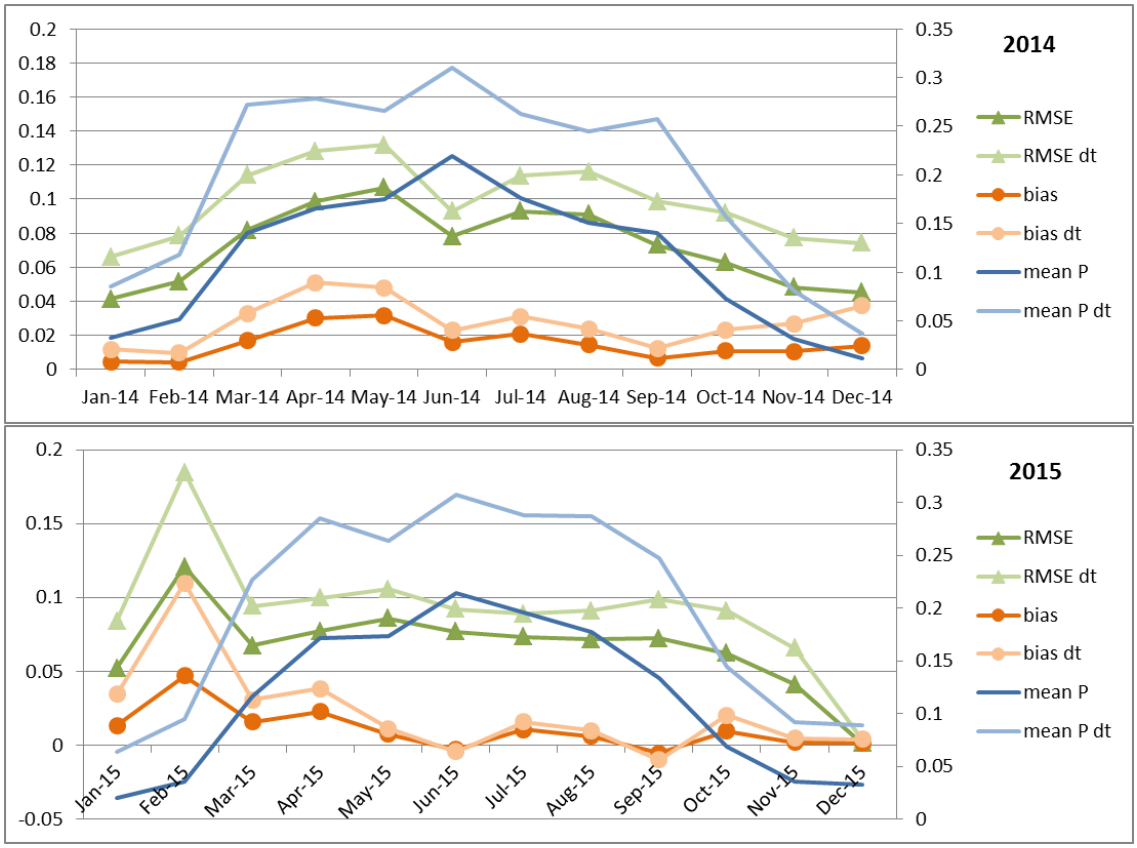
Reference System 0004 – monthly performance evaluation 2014 &2015 (24h forecast)



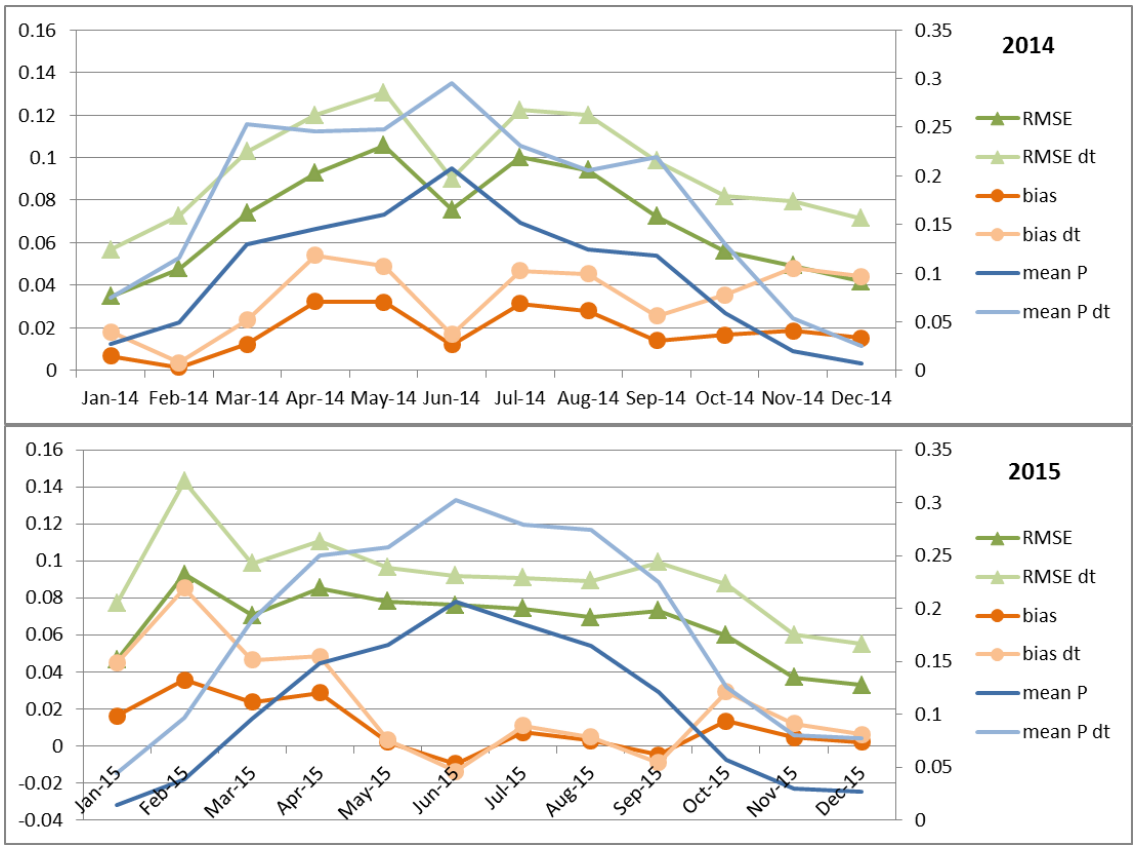
Reference System 0016 – monthly performance evaluation 2014 &2015 (24h forecast)



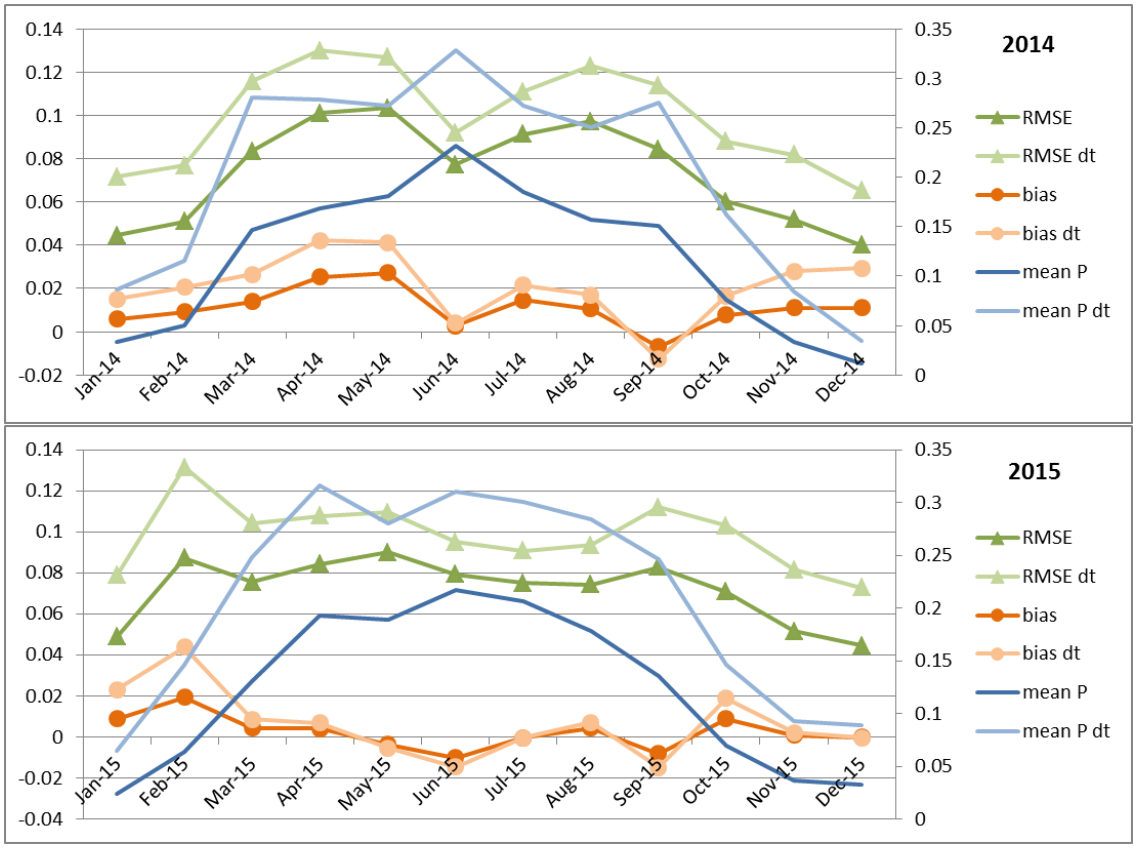
Reference System 0040 – monthly performance evaluation 2014 &2015 (24h forecast)



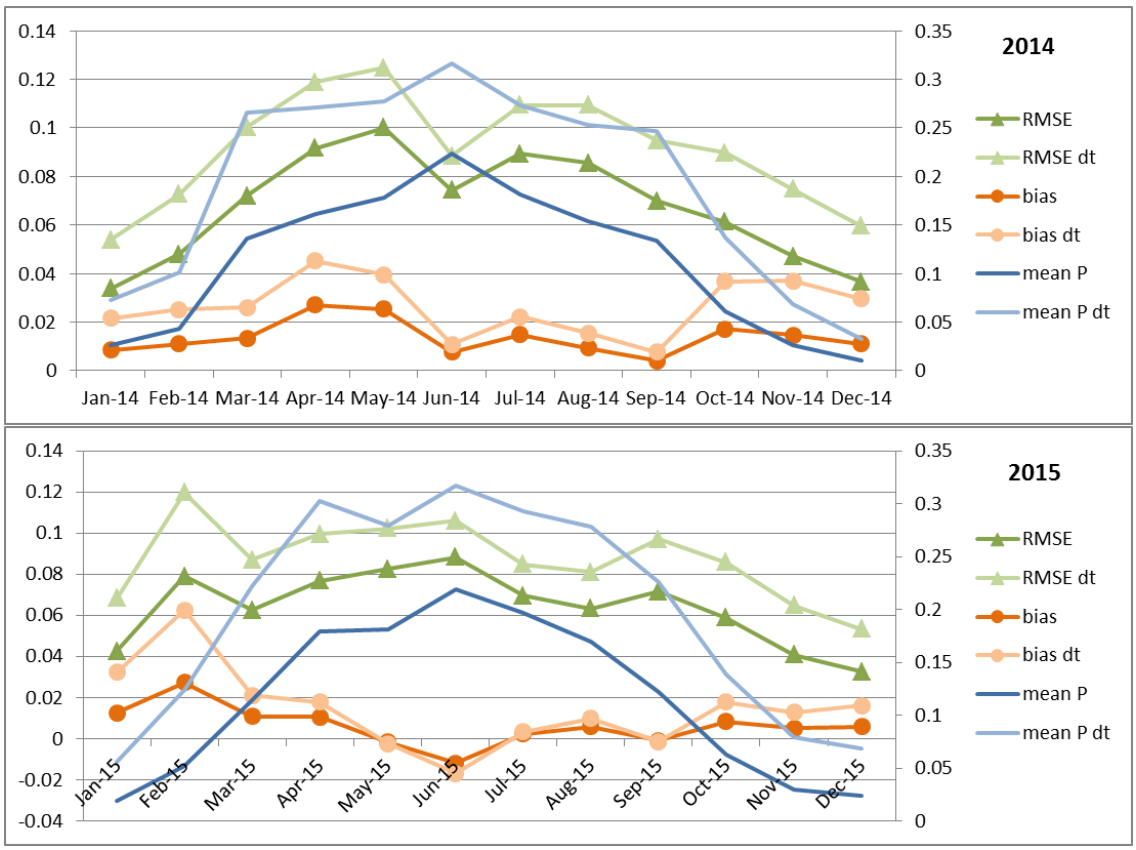
Reference System 0058 – monthly performance evaluation 2014 &2015 (24h forecast)



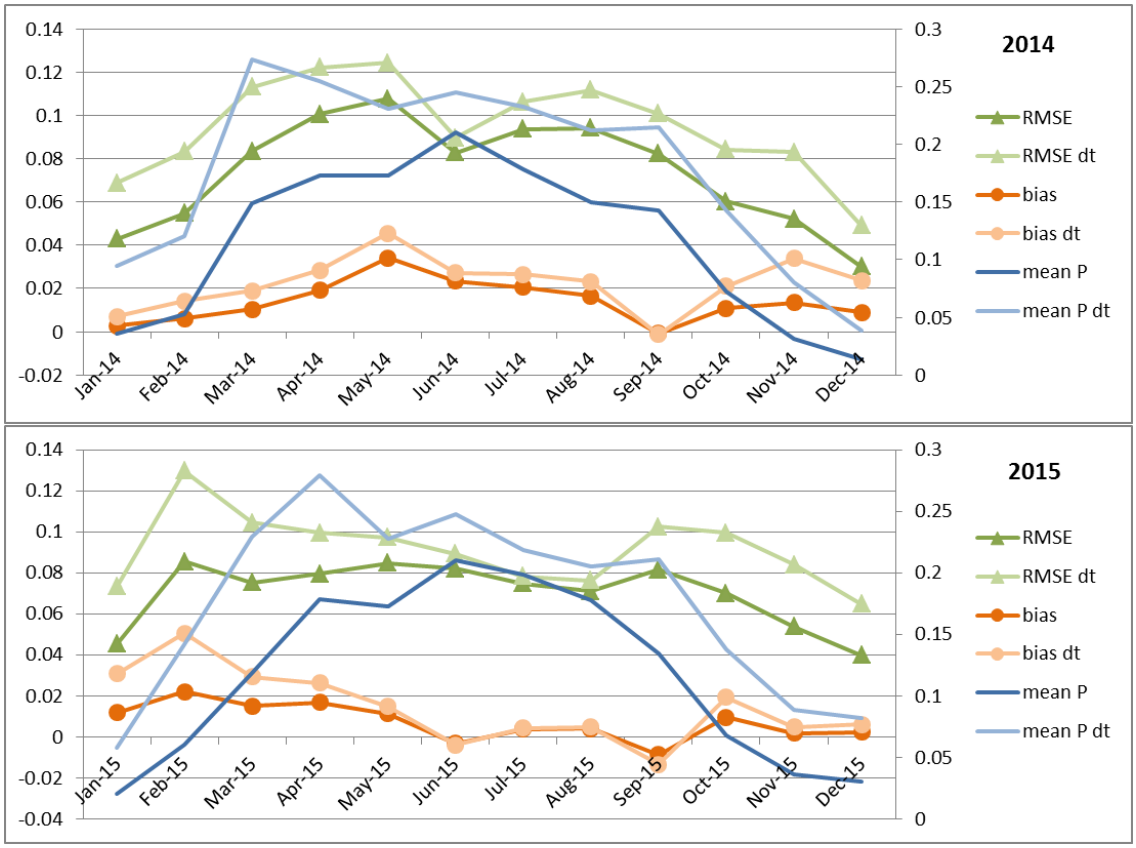
Reference System 0067 – monthly performance evaluation 2014 &2015 (24h forecast)



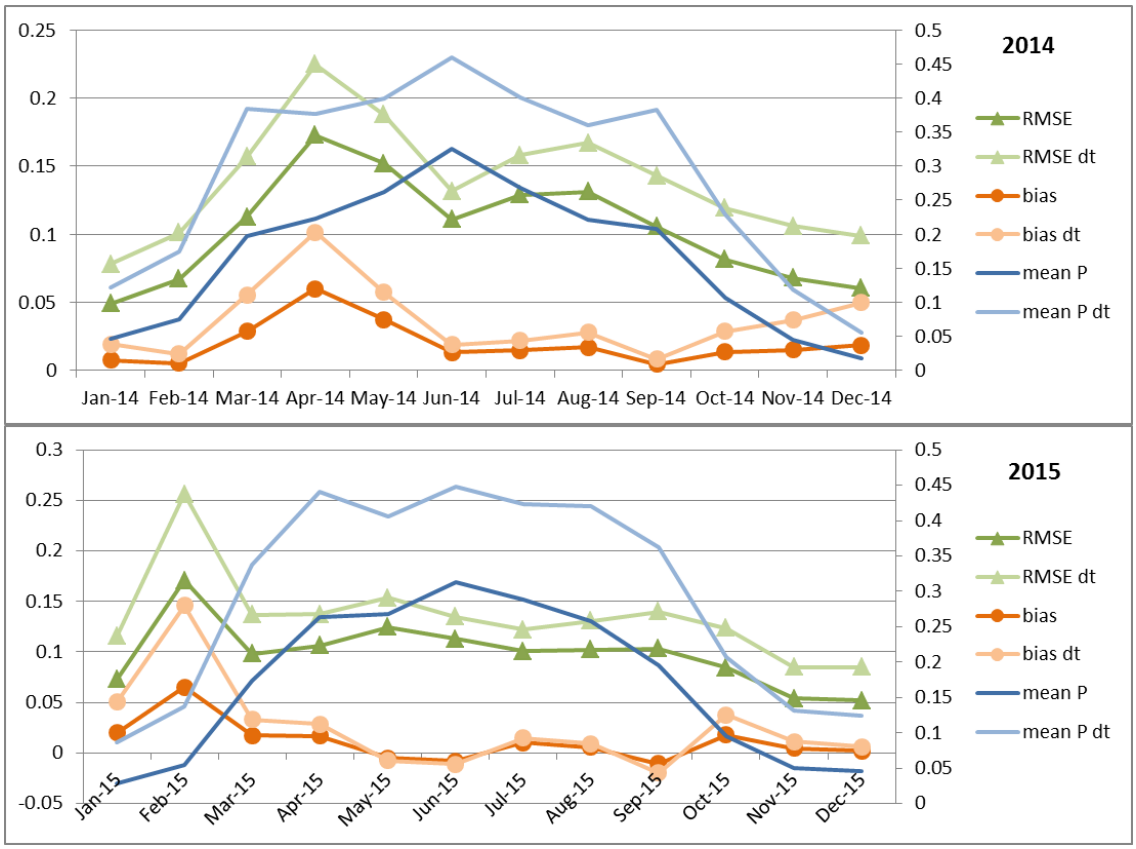
Reference System 0080 – monthly performance evaluation 2014 &2015 (24h forecast)



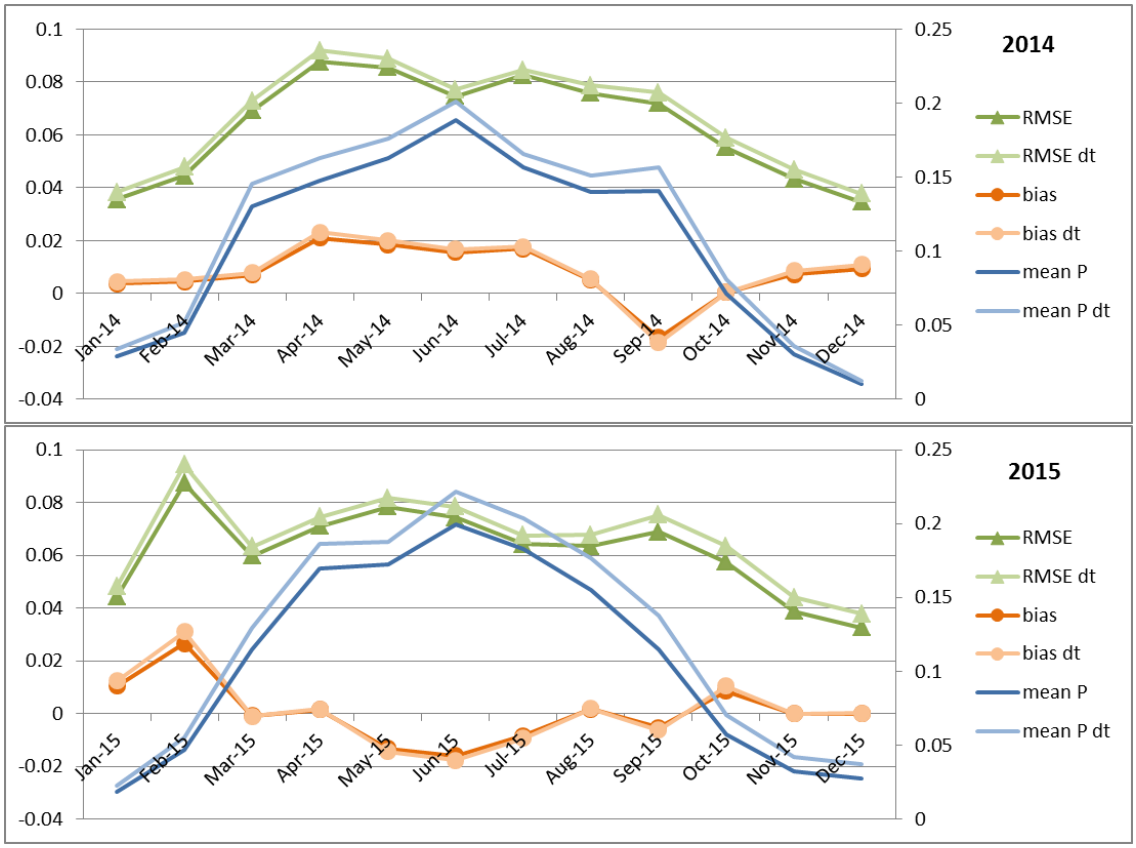
Reference System 0112 – monthly performance evaluation 2014 &2015 (24h forecast)



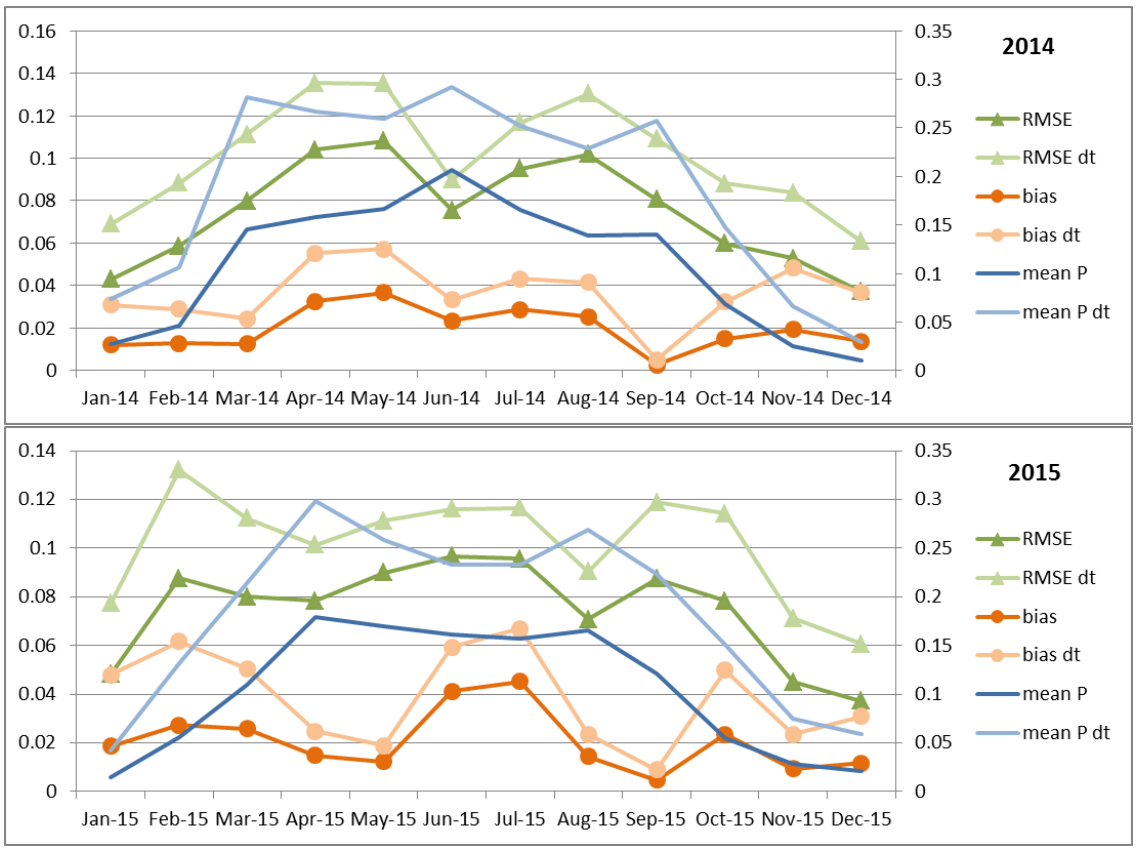
Reference System 0138 – monthly performance evaluation 2014 &2015 (24h forecast)



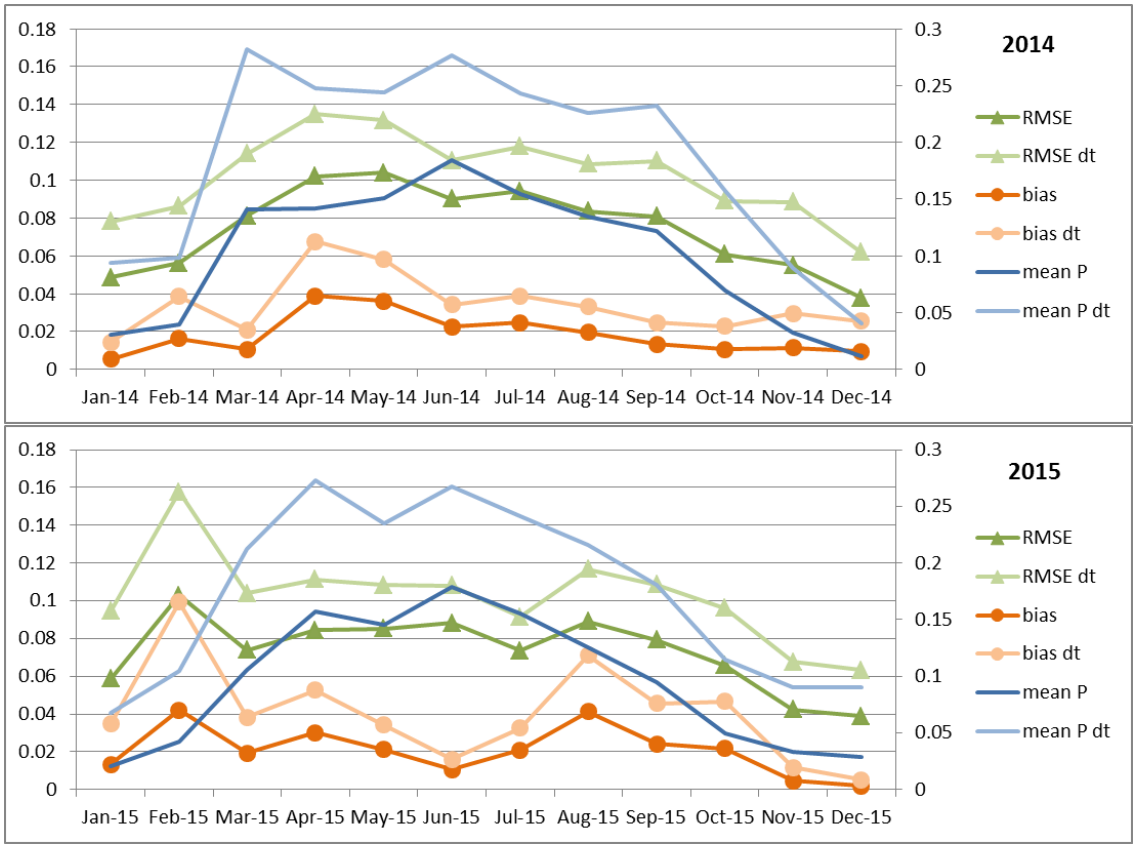
Reference System 1006 – monthly performance evaluation 2014 &2015 (24h forecast)



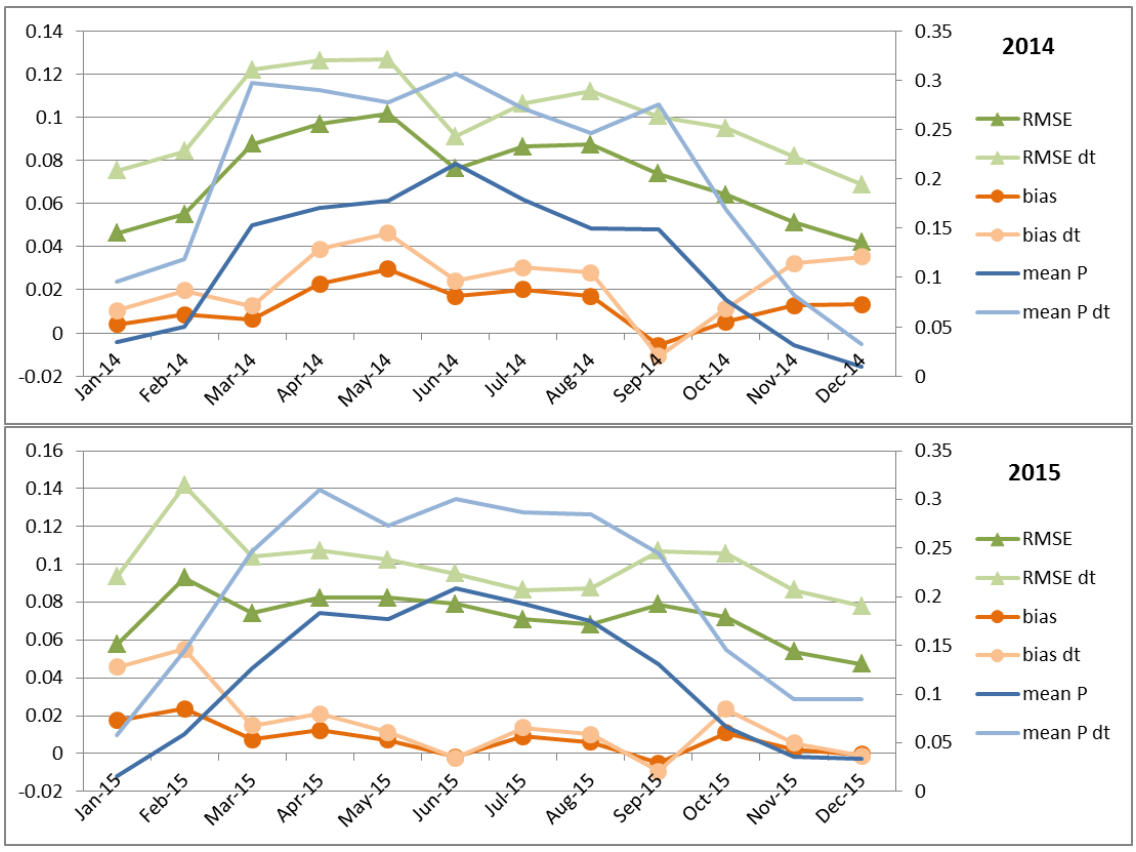
Reference System 1008 – monthly performance evaluation 2014 &2015 (24h forecast)



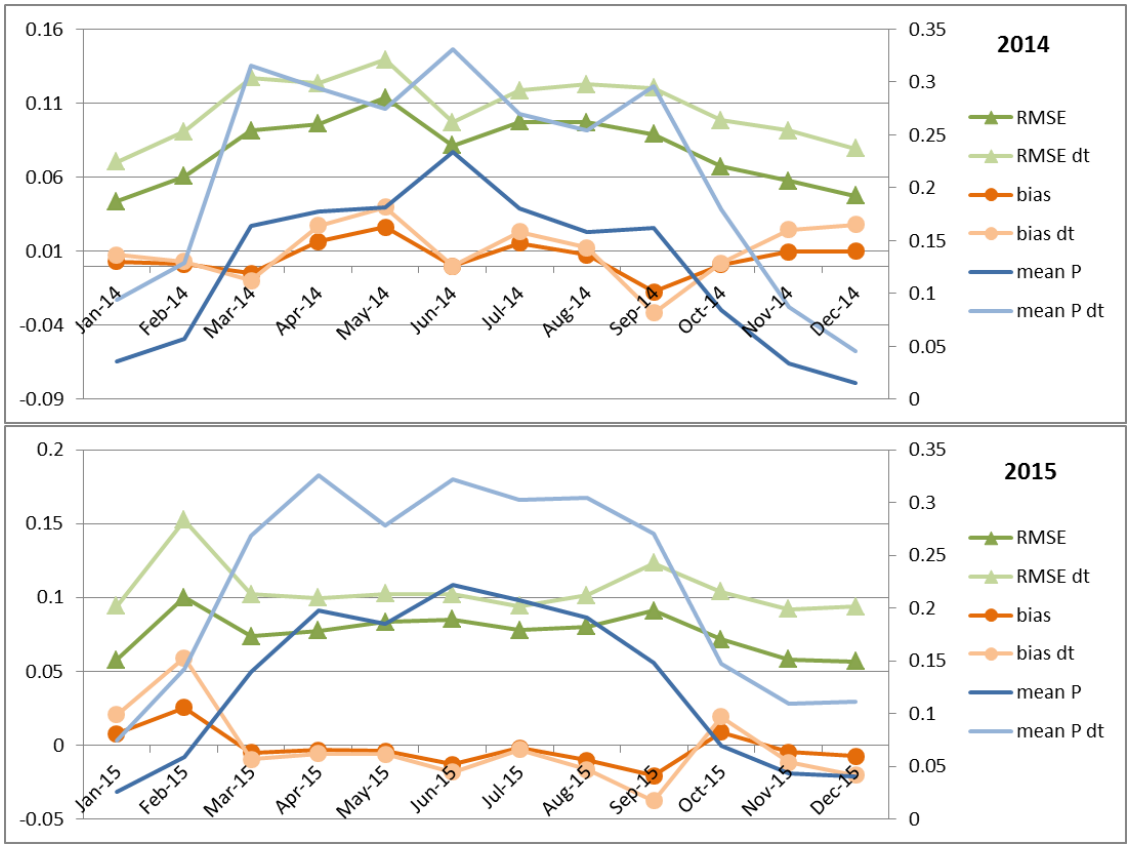
Reference System 1010 – monthly performance evaluation 2014 &2015 (24h forecast)



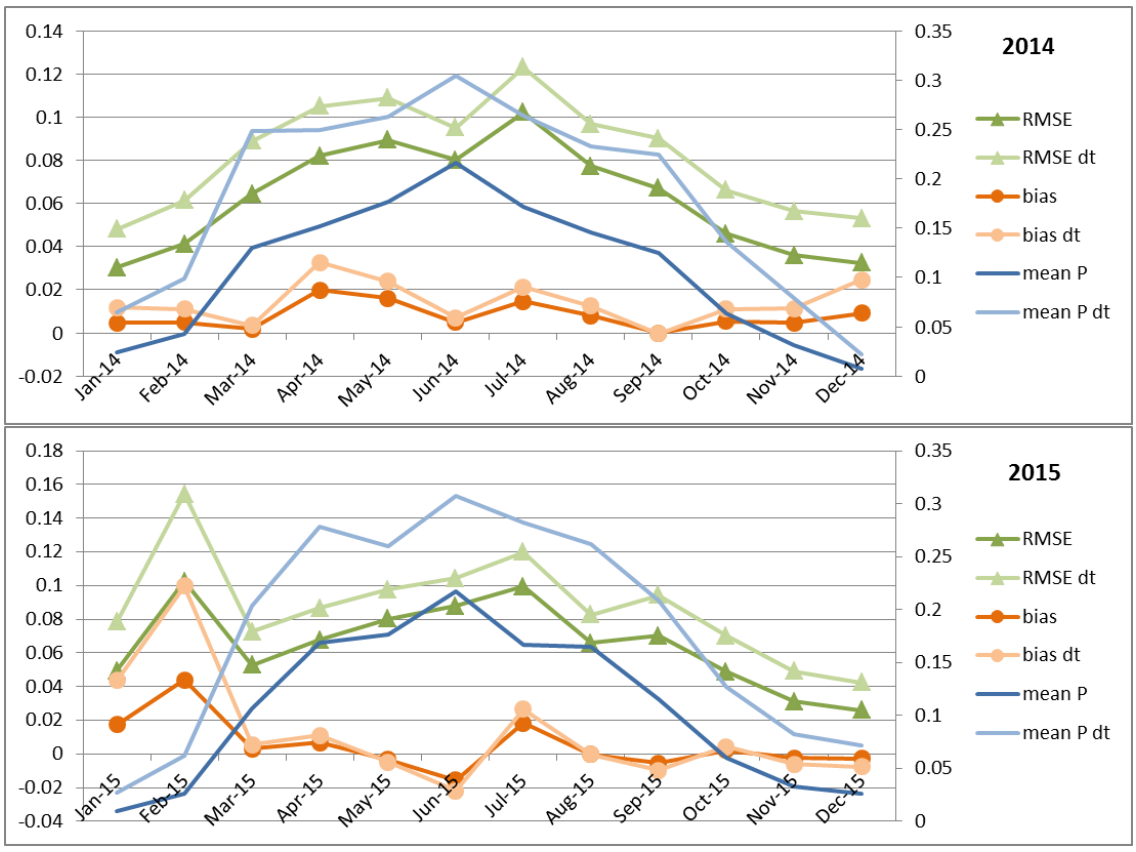
Reference System 1013 – monthly performance evaluation 2014 &2015 (24h forecast)



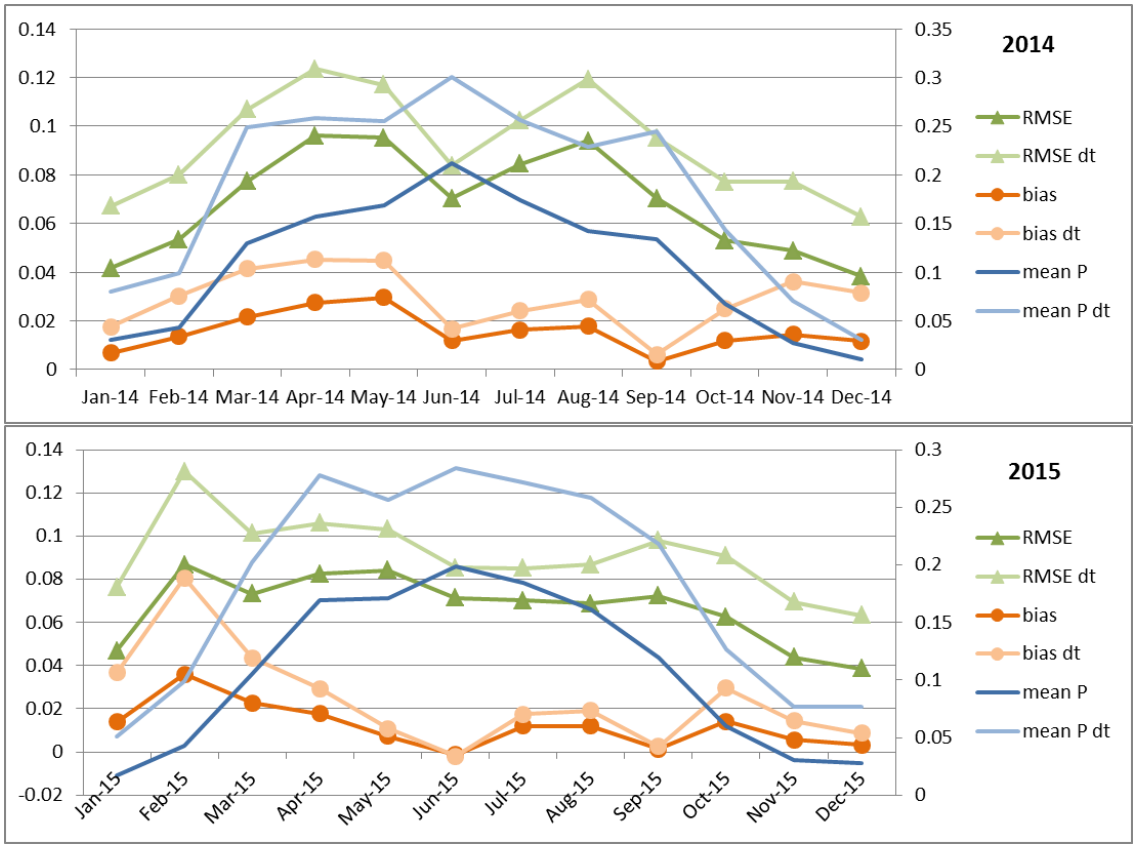
Reference System 1042 – monthly performance evaluation 2014 &2015 (24h forecast)



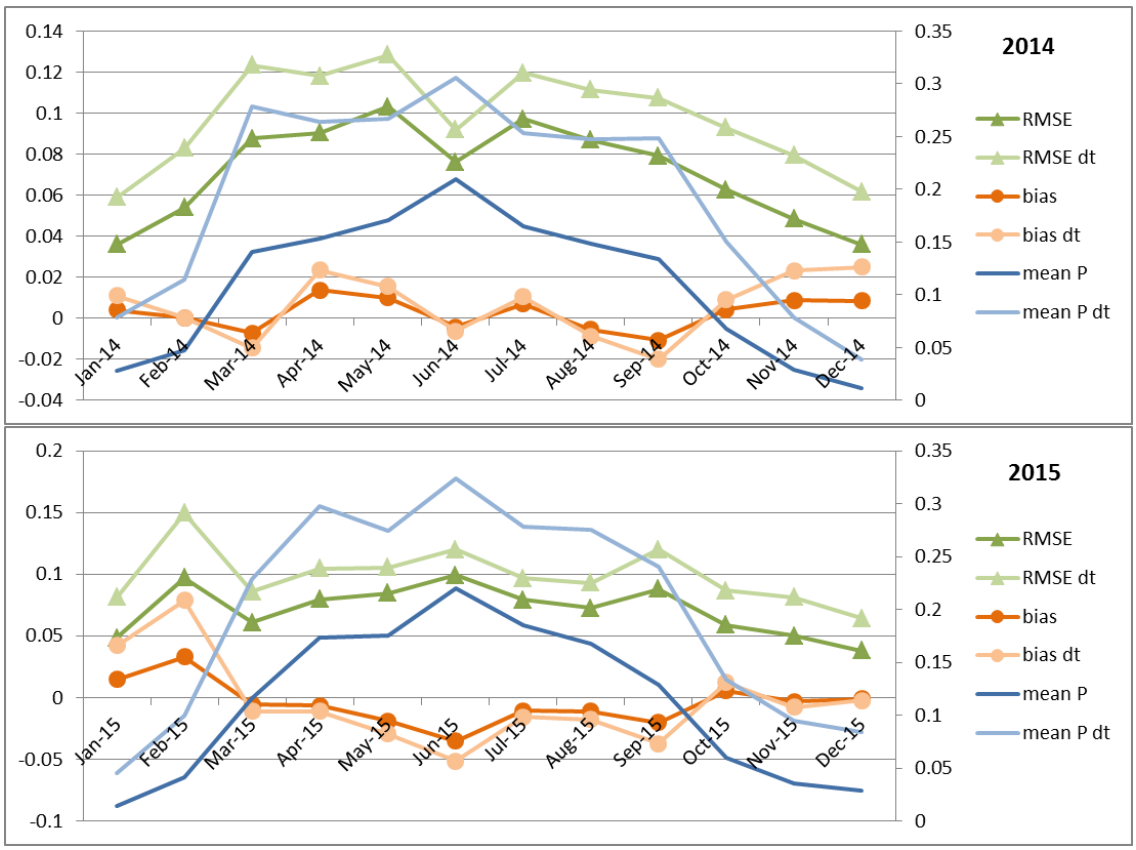
Reference System 1052 – monthly performance evaluation 2014 &2015 (24h forecast)



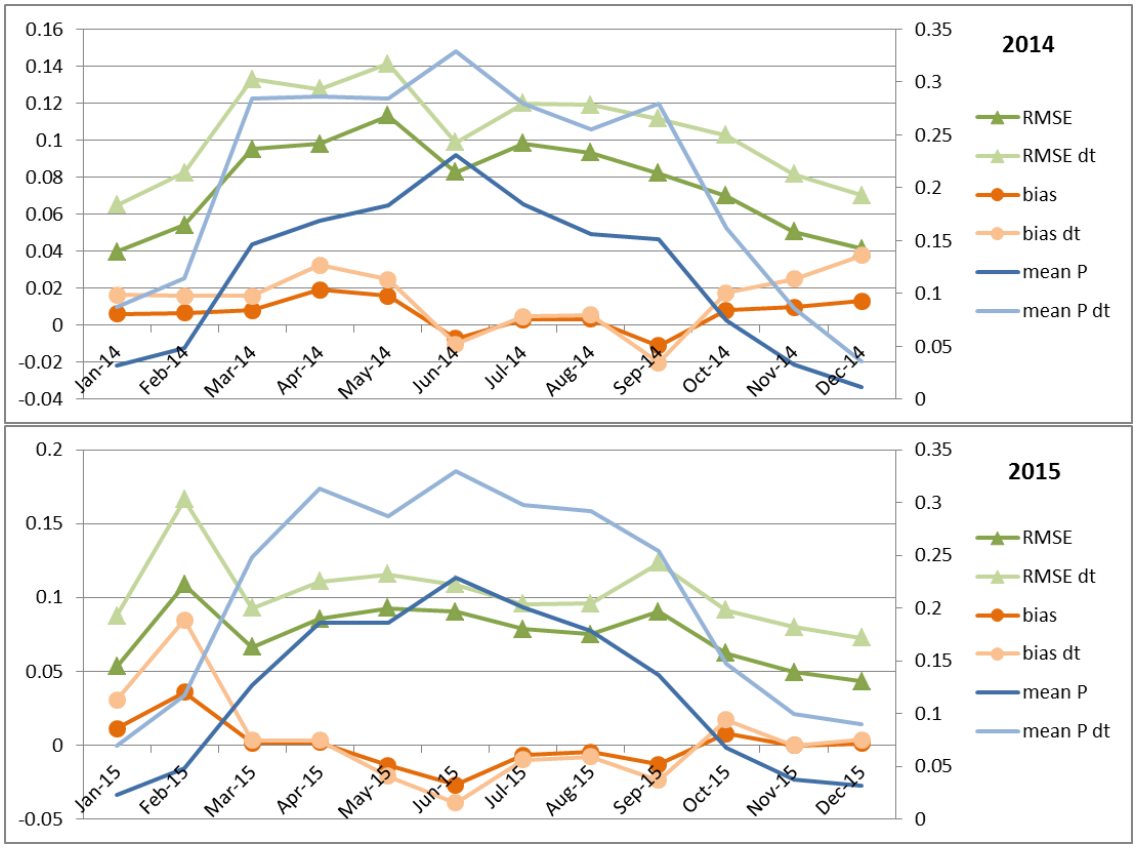
Reference System 1063 – monthly performance evaluation 2014 &2015 (24h forecast)



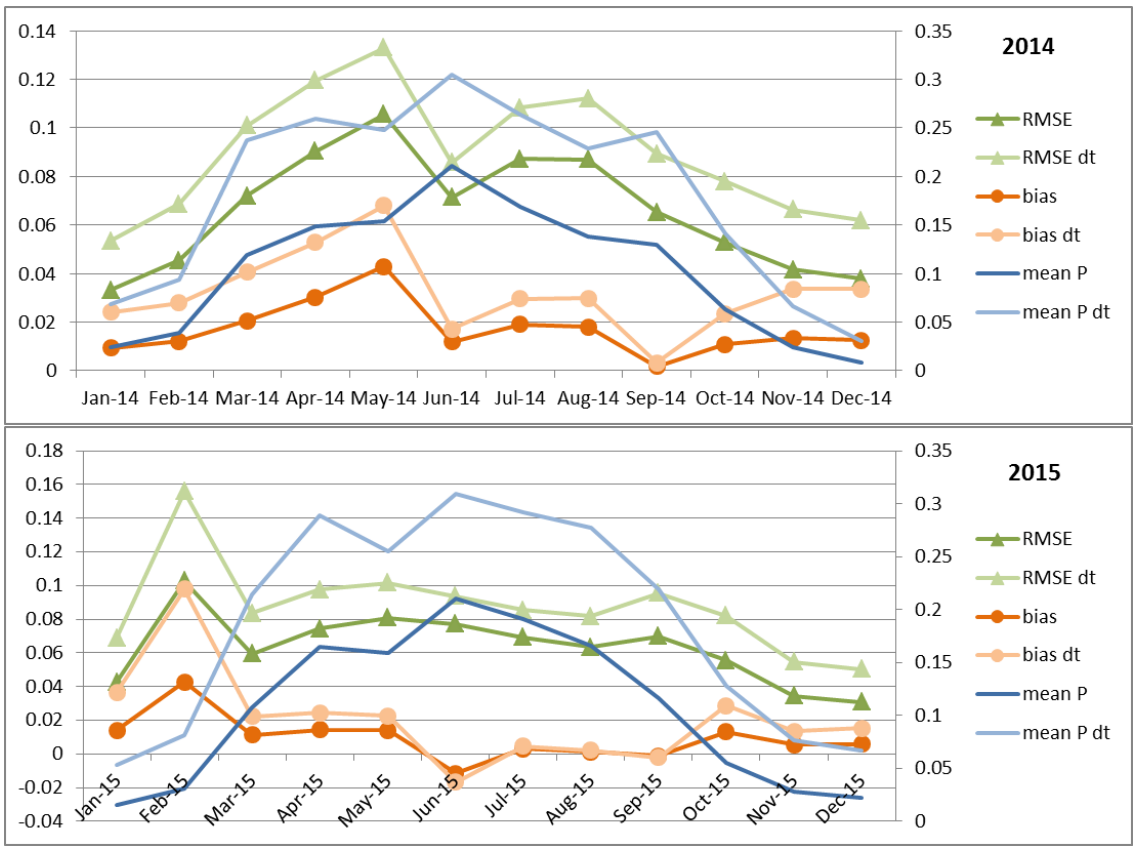
Reference System 1075 – monthly performance evaluation 2014 &2015 (24h forecast)



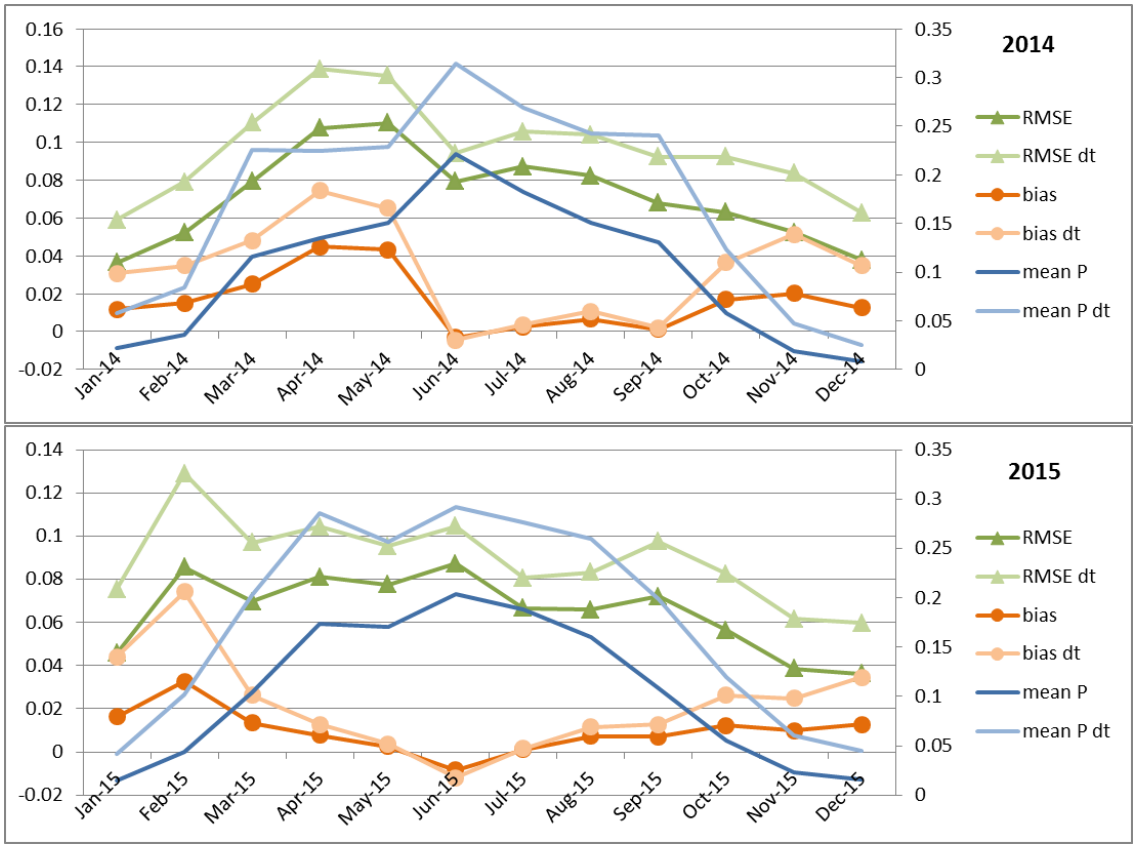
Reference System 1090 – monthly performance evaluation 2014 &2015 (24h forecast)



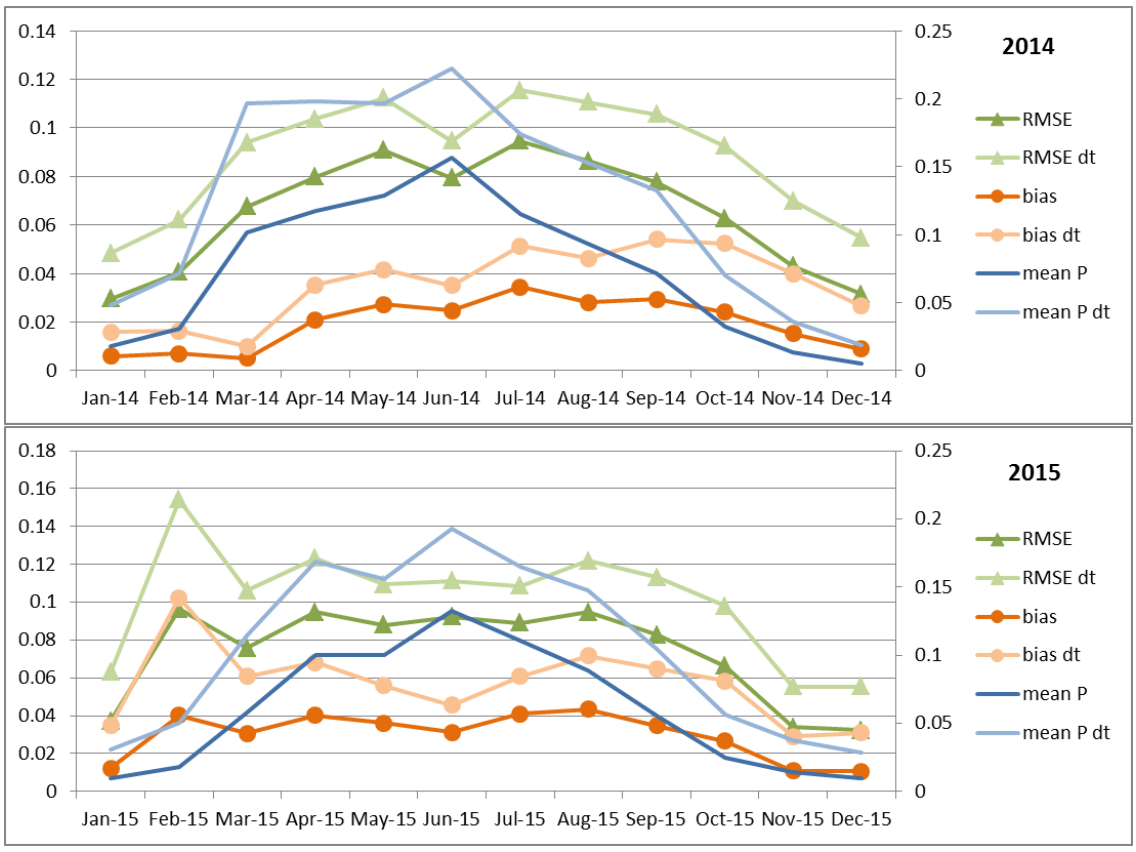
Reference System 1101 – monthly performance evaluation 2014 &2015 (24h forecast)



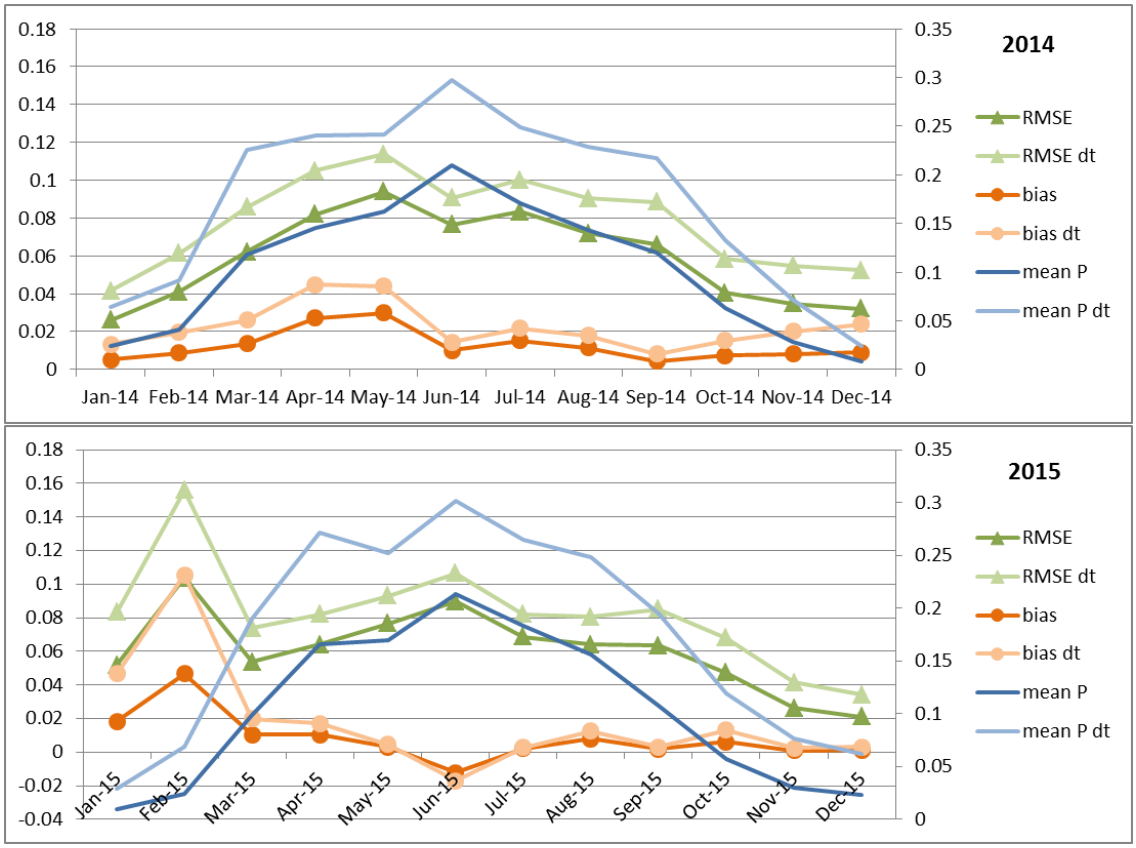
Reference System 1122 – monthly performance evaluation 2014 &2015 (24h forecast)



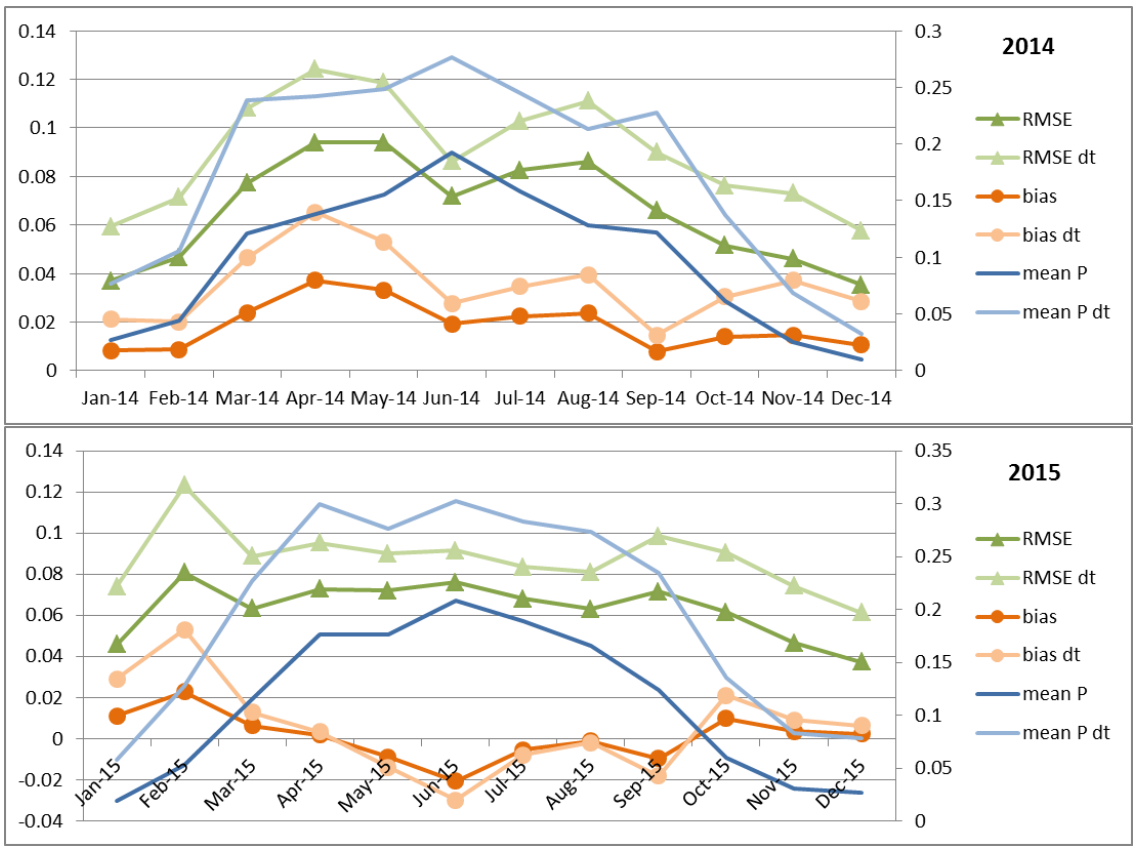
Reference System 1134 – monthly performance evaluation 2014 &2015 (24h forecast)



Reference System 1159 – monthly performance evaluation 2014 &2015 (24h forecast)



Reference System 1169 – monthly performance evaluation 2014 &2015 (24h forecast)



Reference System 1173 – monthly performance evaluation 2014 &2015 (24h forecast)