Master Thesis

on

Spatial and temporal variability of solar energy

24th of January, 2016 – Utrecht
Evan Willemsen, BSc F090755
Utrecht University

Universiteit Utrecht
Faculty of Geosciences

Supervisor: Boudewijn Elsinga. MSc

1. Table of Contents

2. Executive summary

This research was focused on the spatial dependence of irradiance variability with different time scales and includes the influence of the direction of movement of the clouds. The method mentioned in previous research from Perez et al. (2012) is used to measure the spatial- and temporal effects of the irradiance variability. Their research, uses the correlation of delta clearness index. Furthermore, the method mentioned in the research from Elsinga & van Sark (2014) is also used and compared with Perez et al. (2012). Both methods were investigated on a smaller timescale (5 seconds) than the original research (20 seconds& 1 minute) and with more PV systems (200 systems instead of 25).

Where Elsinga & van Sark (2014) used the standard deviation of power output of the PV systems, for this research power output was converted to clearness index and delta clearness index similar to Perez et al. (2012). Where both articles disregarded the wind direction, this research also investigated the effects of wind direction on decorrelation length for both methods. The fit models described in these articles were adjusted for this research with an extra fit parameter to handle the offset of the start of the fit models.

The results of this study shows that with delta clearness index the correlations are the highest on short distances and increase with higher time scales. In contrast, with clearness index, no significant results were observed. The decorrelation length (distance where the variation of irradiance becomes independent) for both methods with delta clearness index data is about 100 meter for 5 seconds data. These results are in accordance with the projected results of the literature. (Perez et al, 2012). Noted should be that, the decorrelation lengths under time intervals of 1 minute, have a decorrelation length which is much smaller than the average distance between the PV systems.

When separating the results in different wind direction, different correlation lengths for each direction are observed. In general, the decorrelation distance of the direction with the highest correlation increases significantly with separation. The direction with the highest correlation length is not always the same as the wind direction. This has multiple reasons: (1) the wind direction is measured on the ground and can therefore be different than the direction of the clouds. (2) The PV systems are non-evenly distributed and therefore the data points are non-evenly distributed. (3) The wind direction can vary over the day.

3. Acknowledgments

This study was made possible by Boudewijn Elsinga (Utrecht University) through providing valuable comments on the research. I would also like to thank Marinka Willemsen for her enormous effort, support and insightful comments.

4. Introduction

Utrecht University (UU) is collaborating in the Solar Forecasting & Smart Grids-project (SF&SG). The goal of the project is to develop a forecasting system based on information from a distributed set of photovoltaic (PV) systems. This research compares the variance of output of PV systems with different methods on a five second time scale.

The increasing penetration of PV has consequences for the energy grid. In particular, the increasing variability of the total power output will affect the grid mainly during daylight hours. Utilities need to adapt their scheduling, planning and operating to maintain the current reliability of the energy grid in the future. Even adaptions to the grid may be needed to not further limit the implementation of intermittent renewable energy systems (iRES) (David, 2013). Estimating the reserve requirements will be a challenge for grid operators: underestimating the requirements will result in power outages and overestimating the requirements will result in unnecessary costs (Hoff & Perez, 2010). For a correct estimation of the reserve requirements it is vital that the variance of the power output and input of the energy grid is known. Different variabilities can be addressed to iRES, but a main cause of short term variability PV is the passing of clouds (Hoff and Perez, 2011). This variability brings in challenges for the grid operators, so knowing what variability to expect on each time frame is essential if large scale solar polar is implemented

Clouds are as (un)predictable as the weather; they constantly change shape and size and are abundant some days, but there are also days without any clouds. Different weather will result in different output of individual PV systems. The spatial movement of clouds translates into correlation of the output of PV systems and the distance between the PV systems (Elsinga & Van Sark, 2014). This issue has been the subject of research in the last few years in France, Japan and Germany, where PV has a significant penetration in the energy market (David, 2013; Hoff & Perez, 2010). Elsinga and Van Sark (2014) have subjected a similar research on the power fluctuations of urban PV panels in the Netherlands.

In the Netherlands, the UU monitors about 200 different PV systems on a high resolution basis, distributed over the province of Utrecht, centred on Lombok. The data is stored constantly and analysed within the project SF&SG. One of the aims of this project is to analyse the overall variability and spatial relationships between PV systems in relation to changing cloud cover.

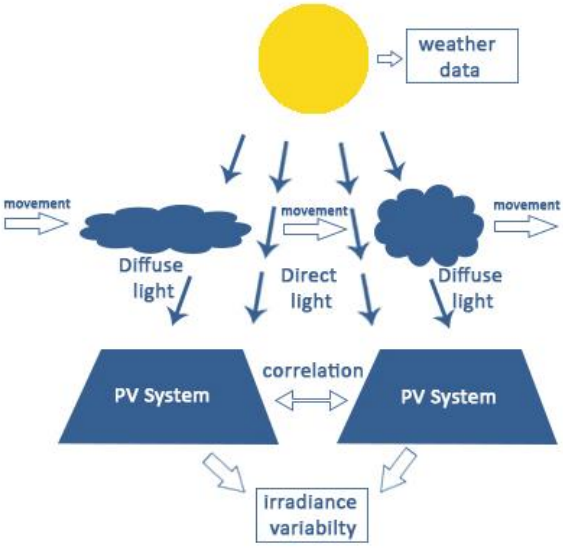


Figure 1. A graphical display of how the sun and clouds influence the correlation.

Research from Hoff and Perez (2011) and Mills and Wiser (2010) shows that PV systems decorrelate with increasing distance. Decorrelation is the distance where PV systems become independent of each other. In the above mentioned articles the time scales are one minute, whereas Elsinga and Van Sark (2014) state that researching on a as small as possible timescale will result in a better understanding of the variance of the power output. According to the article of Yordanov, Sastre and Midtgard (2013), a time resolution of 0.10 second is necessary to measure all changes in irradiance due to clouds passing. From the data available of the SF&SG project, it is possible to calculate the correlation of output and distance on a timescale of five seconds.

The correlation of output of PV panels is a predictable function of three factors according to Perez et al. (2012) and Lave and Kleissl (2015): (1) variance depending on the distance, (2) the timescale and (3) the speed of the clouds. Bosch, Zheng and Kleissl (2013) and Hinkelmann (2013) also include the vector in which the clouds move. In the current research the speed of clouds are omitted since no local data was available of the cloud speeds during the measurement period

The solar output of all PV systems combined varies less than the sum of all individual variances of the PV systems. This has to do with a soothing effect: the variance reduces on a larger scale base (Wiemken, Beyer, Heydenreich & Kiefer, 2001). This soothing effect has the result that the total fluctuation on the energy infrastructure will be less than the individual fluctuations. The soothing, the temporal and the spatial effect have been compared with different directions of cloud movement in the article of Hinkelmann (2013), although the amount of systems and distance between systems is limited in her research.

In the article of Perez et al. (2012) the irradiance variability of PV systems is compared with the correlation between the PV systems. The article of Elsinga and Van Sark (2014) uses the standard deviation of the variance to calculate the variability. In their article a comparison is made with the results of the article of Perez et al. (2012). They noticed that the chosen measure gives similar results as Perez et al (2012). Perez et al (2012) uses 24 PV systems and time scales from 20 seconds to 15 minutes and Elsinga and Van Sark (2014), 25 PV systems with a time scale of 1 minute. This research compared the two methods on a smaller timescale (5 seconds) and with more PV systems (200 systems). Both articles also disregard the wind direction, this research also investigated the effects of wind direction for both methods.

5. Project scope and objectives

In this research the solar intermittency was assessed by taking into account the temporal and spatial effects of the PV systems and after that a division was made in wind directions. To accomplish this the following question was derived as a guideline:

> *What are the relations of solar intermittency with inter-system correlation, spatial effect, temporal effect and the direction of movement of the clouds?*

The answers to the research question are obtained by answering the following sub questions:

> *-How does the 'Elsinga' [1] and 'Perez' [2] method compare with updated results on t the inter system correlation?*
>
> *-How does inter-system correlation depend on different time resolutions?*
>
> *-What is the effect of wind direction on solar intermittency?*

This research is focused on three different aspects: (1) temporal correlation, (2) spatial correlation and (3) on wind direction. Research has been done for the temporal effect to Perez et al. (2012), Hoff and Perez (2012), Lave and Kleissl (2015) but research is still needed on the variance of PV systems on a short timescale. Therefore this effect is researched thoroughly in this study.

[1] "Elsinga" is defined in this research as the method used to calculate inter sytem correlation in the article of Elsinga and van Sark (2014)

[2] "Perez" is defined in this research method used to calculate inter sytem correlation in the article of Perez et al(2012).

For the temporal and spatial effect on solar intermittency, the database of the SF&SG-project was used. From this database, data was gathered and analysed using MATLAB (version 14b, Mathworks, 2016). , software designed to calculate with data from large databases such as the SF&SG-database.

Most interesting for the Netherlands and this research are days with partly clouded skies. On days like these, power output of PV will fluctuate due to clouds passing over and because this weather is common in the Netherlands (KNMI, 2015). In the figure (2) below the irradiance on a day with a clear sky (left) and a clouded day (right) can be seen. On the clouded day the pattern is very unpredictable, this due to the fact that the clouds influence the amount of solar power which reaches the ground of the earth.

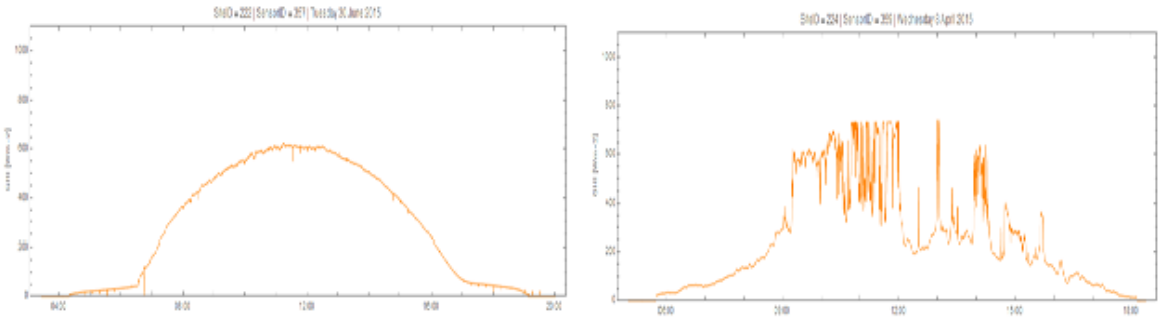


Figure 2. Irradiance on the left 20 June 2015: a day with clear sky, right 8 April 2015: a clouded day

This research was geographically limited to an area almost similar to the province of Utrecht. In this area the PV systems are non-randomly distributed. This is because more PV systems are installed in urban areas than in rural areas, urban areas have a higher representation of PV systems in the database. The locations of the PV systems can be found in figure 2.

Daily effects such as the sunset and sunrise were included in the calculations. The effect of sunset and sunrise on power fluctuations is researched in the paper of Pavanello et al (2015). They mention that the fluctuation of a PV installation can vary at these times due to objects in the skyline and their shadows. Yordanov et al, (2013) argues that a time resolution of 0.10 second will include all variances, while the dataset used in this research gives data up to every 5 seconds and has given the chance to research the temporal effect up to 5 seconds.

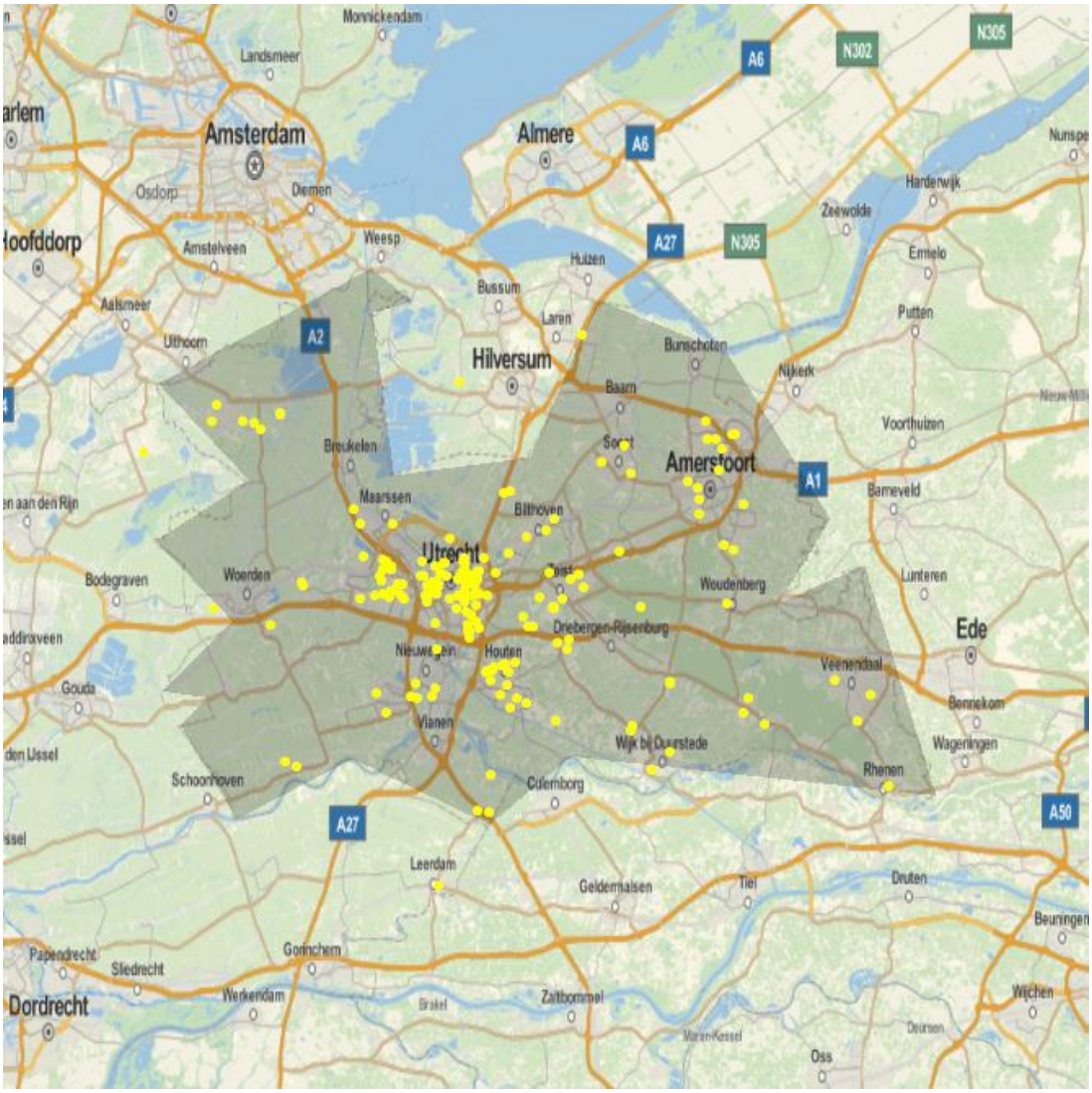


Figure 3: Location of 202 constantly measured solar panels from the SF&SG project.

6. Method

6.1. Research Outline

To answer the research question of what the temporal, spatial and wind direction characteristics of short term solar intermittency are, the following steps were taken: firstly, the scope and boundaries of the research were defined. Secondly, the current state of research was found with a literature study. Data was gathered from 200 PV systems and non-useable data was discarded (see figure 4).

Thirdly, intersystem correlation was compared for 202 PV systems with (1) the temporal effect on correlation, (2) spatial effect on correlation. This research builds upon research of Elsinga and Van Sark (2014), with more PV systems and with the use clearness index (K_t^*) instead of power output. Clearness index is defined as the ratio of the measured global horizontal irradiance to the modeled clear sky irradiance at a horizontal level at a given time. An updated database of the SF&SG project was used for correlation of variance described in the paper of Elsinga and Van Sark (2014) and with the method of Perez et al. (2012). These results have been compared on the temporal and spatial effects.

Fourthly, to answer the research question the temporal and spatial effect was calculated and the correlations were specified with wind direction and compared with the temporal and spatial effects of non-specified correlations.

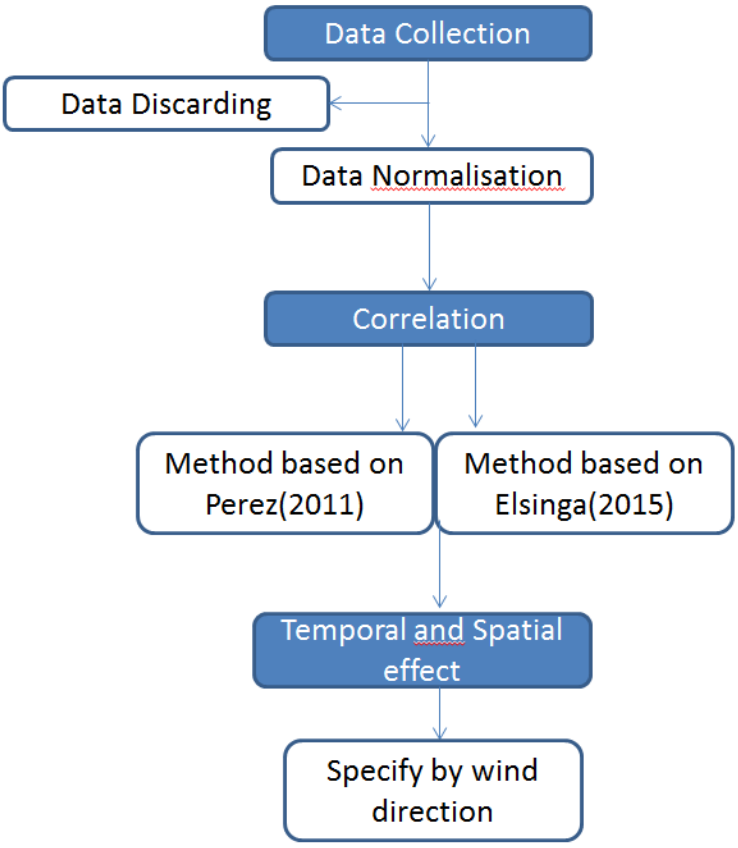


Figure 4. Steps taken to answer the research question

6.2. Data

In this chapter the collection of data, handling of data and discarding of data will be discussed. In the same order as figure 4.

6.2.1. Data collection

The data of the PV systems for this research is the same as Elsinga and van Sark (2015) and has been collected at a 5 second rate for 24 hours a day for the years 2014 and 2015 in the city region of Utrecht (NL) for 202 different PV systems. The energy output of these PV systems is measured with online UPP energy power measurement data loggers (Upp Energy, 2015). The location of these 202 PV systems can be found in figure 3. The output data of the data loggers is converted to Global Horizontal Irradiance (GHI) using a model of Ecofys within the SF&SG project. The GHI is modelled by a model from Elsinga and Van Sark (2016) to K_t^* .this model is based upon Perez et al. (1987). The conversion from GHI to K_t^* in short was done by dividing GHI with the Modelled clear sky Global Horizontal Irradiance (MGHI) with a known location, tilt and orientation.

Where the paper of Elsinga and Van Sark (2015) uses time bins of 10 seconds, this research uses time bins of 5 seconds. This is done so the temporal effect can be examined on an even smaller scale. The GHI andK_t^* data was provided for this research.

Not only irradiance values were used in the calculations but also weather data, publicly available data from the weather station in De Bilt- situated in the city region of Utrecht- from The Royal Netherlands Meteorological Institute (KMNI). The KNMI uses an octane scale for the amount of clouds covering the sky. In this research table 1 was used for the classification of cloud covering of the sky.

Table 1.
Weather classifications used in this research

KNMI octane scale of cloudiness	Day class
<0.3	Cloudy
0.3<scale <0.5	Partly Cloudy
>0.6	Clear sky

The average of the KNMI octane scale of cloudiness during the measured periods was calculated with the accompanying median wind direction. A selection of days was made based on these results, and the weather data of these chosen days can be found in Attachment 1.

A different way to classify the weather is by looking at the irradiance values themselves. This method is the same as Elsinga and van Sark (2016). The classification is done by a so called "arrowhead" diagram: A diagram which plots the clear sky index (Cd) against the relative variability (Vd).

To calculate the clear sky index Cd, the GHI of an entire day was divided by the MGHI of that day (see equation 1) and to calculate the relative variability for each day the irradiance variability compared to the clear sky variability. With the assumption that for each day a system situated in the middle of Utrecht the Cd and Vd for all systems are comparable with the chosen system. In other words, the weather is the same with 25 km radius from the chosen system on each day.

$$Cd = \frac{\sum GHI(t)}{\sum MGHI(t)} \tag{1}$$

$$Vd = \frac{\sqrt{\sum (GHI_t - GHI_{t-1})^2}}{\sqrt{\sum (MGHI_t - MGHI_{t-1})^2}} \tag{2}$$

When the relative variability is plotted against the clear sky index, as can be seen in figure 5, the separation becomes visible between overcast days, clear days and broken cloud days. These results is also visible in the article of Elsinga and van Sark (2016).

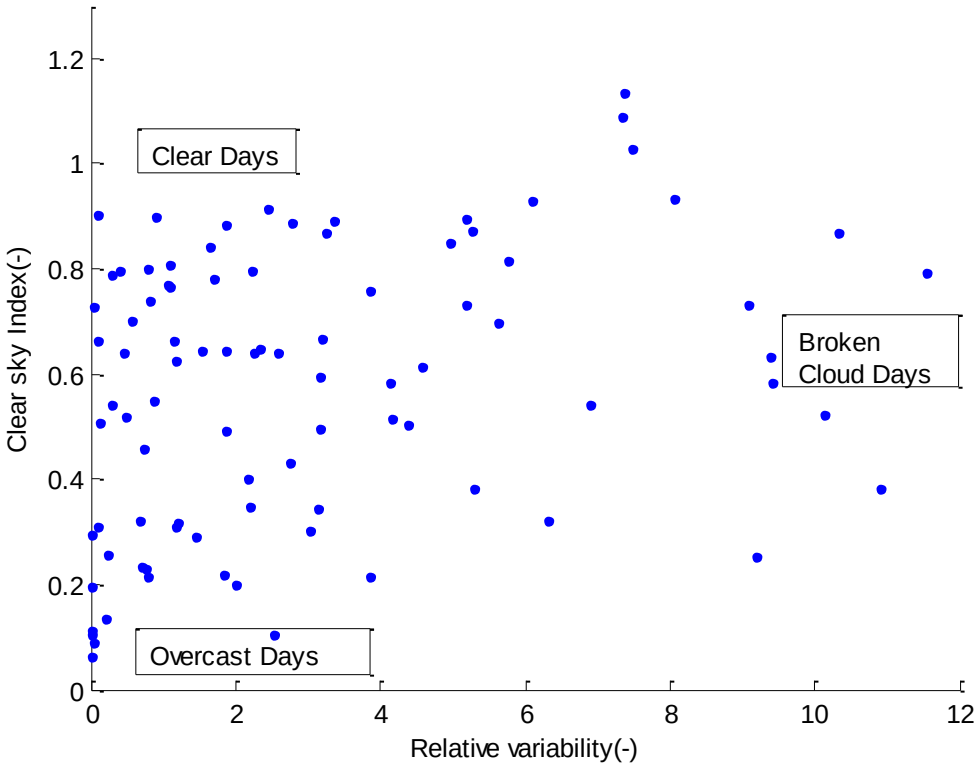


Figure 5. Characterisation of the variability and clear sky index of the days in the measurement], in an arrowhead diagram similar to Elsinga and van Sark(2016).

6.2.2. Data discarding

Measurements from two hours after sunrise and before two hours before sunset will be used, to avoid shading from nearby buildings and trees due to a high solar zenith angles at this times. Data of clearness index with data gaps or more than 140% of the site specific installed power is discarded. This is the same approach as Elsinga and Van Sark (2014) used. When increasing the resolution of the irradiance index the variation of K_t^* changes as described in the introduction.

To research the temporal effect as described in the main question, the change of K_t^* with time increment variables varying from 5 seconds to 20 minutes were investigated, where the variation decreases with increasing time increment as described in Elsinga and Van Sark (2015). For the current research the input data of K_t^* in 5 seconds was averaged in time bins to the desired time series. The averaging time increment values were: 10, 20, 25, 60, 120, 180, 240, 300, 450, 600, 750, 900, 1050 and 1200 seconds. The 20 minutes timescale is chosen as an upper limit due to grid operators predicting the needs and demands of energy at this timescale (see attachment 3).

6.3. Correlation

The method used for calculating the standard deviation of the input data in the papers of Elsinga and Van Sark (2014) and Hoff and Perez (2012), is done by calculating the distance between each pair of the PV systems and comparing respectively the variance of the power fluctuation and the correlation of delta clearness index ($\Delta K_{\Delta t}$). Furthermore, a comparison between the two different methods of the papers are described and a comparison between using $\Delta K^*_{\Delta t}$ and K^*_t for both methods is given.

6.3.1. Distance & azimuth

To simplify the calculation of the distance between the PV systems, the assumption was made that earth is a sphere where in reality earth is an oblate spheroid. This simplification does not have a large impact on the resulting distances since the maximum distance between two PV systems is only 57 km. Firstly, to determine the distance, with the location of each PV system given in longitude and latitude of each PV system, the delta longitude Δlon and delta latitude Δlat for all i,j pairs are calculated:

$$\Delta lon_{i,j} = lon_j - lon_i \quad (1)$$

$$\Delta lat_{i,j} = lat_j - lat_i \quad (2)$$

Secondly, the great circle distance of the i,j pairs ($d_{i,j}$) in kilometres was calculated, see Equation 3. This equation is from Sinnot (1984), where R is the radius of the earth, which is stated in this article as 6371 kilometre.

$$d_{i,j} = R * 2 * \arcsin\left(\min, \left(1, \sqrt{\left\{\sin\left(\frac{\Delta lat_{i,j}}{2}\right)\right\} + \left[\cos(lat_i) * \cos(lat_j) * \left\{\sin\left(\frac{\Delta lon_{i,j}}{2}\right)\right\}^2\right]}\right)\right) \quad (3)$$

To compare the correlation of all i,j pairs with the wind direction it was necessary to calculate the azimuth for each possible pair of the 200 PV systems. Azimuth is described in this research as the horizontal component of coordinates on the earth[3]. The east of the earth for example is on a 90 degrees angle with the North Pole. In this research for the wind direction and the azimuth the north is defined as 0 degrees, west as 90 and south is 180 and east as 270. The azimuth angle α can be calculated similar to Sinnot (1984) by:

$$x = \operatorname{acos}\left(\frac{\sin(\mathrm{lat}_j) - \sin(\mathrm{lat}_\mathrm{i}) * \cos(\mathrm{d}_{i,j})}{\sin(\mathrm{d}_{i,j}) * \cos(\mathrm{lat}_\mathrm{i})}\right) \tag{4}$$

If $\sin(\Delta\mathrm{lon}_{\mathrm{i,j}}) < 0$, $\alpha = x$

If $\sin(\Delta\mathrm{lon}_{\mathrm{i,j}}) > 0$, $\alpha = 2 * \pi - x$

6.3.2. Standard deviation: Elsinga's method (2014)

Where in the paper of Elsinga and Van Sark (2014) the comparison of the PV systems is done with power fluctuations, in this research the power output of the PV systems is converted to clearness index. There are two reasons for this conversion: (1) it is easier to compare the methods of Perez (2012) and Elsinga and Van Sark (2014) and (2) clearness index includes correction for the slope and orientation.

As described in section 6.3.1., the clear sky index (K_t^*) is defined as the ratio of the measured GHI to the modeled GHI (see equation 5). In this research delta clear sky in Index ($\Delta\mathrm{K}_{\Delta\mathrm{t}}^*$), is defined as K_t^* minus K_t^* of the previous time step similar to Hoff and Perez (2012), see Equation 6.

$$\mathrm{K}_{\mathrm{t,i}}^* = \frac{\mathrm{GHI}_\mathrm{i}}{\mathrm{MGHI}_\mathrm{i}} \tag{5}$$

$$\Delta\mathrm{K}_{\Delta\mathrm{t,i}}^* = \mathrm{K}_{\mathrm{t,i}}^* - \mathrm{K}_{\mathrm{t-1,i}}^* \tag{6}$$

[3] With the simplification that earth is a sphere where in reality earth is an oblate spheroid similar to the calculation of distance

The research of Elsinga and Van Sark (2014) uses the standard deviation of the variability as a way to calculate the decorrelation between PV systems. This research also includes the variance, this was calculated for each possible pair (i,j) of the 200 PV systems with K_t^* and with $\Delta K_{\Delta t}^*$.

$$\mathrm{Var}_{ij,d}^{K_t^*} = \mathrm{variance}(K_{t,i}^* , \Delta K_{t,j}^*) \quad (7)$$

$$\mathrm{Var}_{ij,d}^{\Delta K_{\Delta t}^*} = \mathrm{variance}(\Delta K_{\Delta t,i}^* , \Delta K_{\Delta t,j}^*) \quad (8)$$

Then the standard deviation is calculated for $Var_{ij,d}^{I}$ and $Var_{ij,d}^{\Delta I}$ with the following formulas.

$$\sigma_{ij,d}^{\overline{(K_t^*)}} \equiv \sqrt{\mathrm{Var}_{ij,d}^{K_t^*}} \quad (9)$$

$$\sigma_{ij,d}^{\overline{\Delta K_{\Delta t}^*}} \equiv \sqrt{Var_{ij,d}^{\Delta K_{\Delta t}^*}} \quad (10)$$

Elsinga and Van Sark (2014) uses an exponential fit model for the standard deviation of the intersystem power output variance over inter system distance. In this article a fit model is used to quantify the relation between the standard deviation $\sigma_{ij,d}^{\overline{\Delta K_{\Delta t}^* / K_t^*}}$ with the distance between PV systems ($d_{ij)}$ (see equation 11). In this equation the a and b are fit parameters, where b is the decorrelation length and a is the distance independent variable.

$$\sigma_{ij,d}^{\overline{\Delta K_{\Delta t}^* / K_t^*}} d_{ij} = a\left(1 - e^{-\left(\frac{d_{ij}}{b}\right)}\right) \quad (11)$$

In this research an extra fit parameter (c) is added to better fit the results. This c is an offset for the start of the equation (since the variation does not always need to start at 0, when the distance between PV systems is small) .

$$\sigma_{ij,d}^{\overline{\Delta K_{\Delta t}^* / K_t^*}} d_{ij} = a\left(1 - e^{-\left(\frac{d_{ij}}{b}\right)}\right) + c \quad (12)$$

6.3.3. Correlation: Perez's method (2012)

As mentioned above, Perez et al. (2012) use $\Delta K^*_{\Delta t}$ to show fluctuations of the PV systems. The $\Delta K^*_{\Delta t}$ is correlated with the Pearson correlation. Pearson correlation coefficient ρ with i and j time series of equal length and nonzero standard deviation σ is defined as:

$$\rho_{i,j} = \frac{\text{Covariance(i,j)}}{\sigma(i)\sigma(j)} \tag{13}$$

.

Similar to Elsinga and Van Sark (2014), the relation of correlation $\rho_{i,j}$ with distance $d_{i,j}$ can be fitted by an exponential model. Hoff and Perez (2012) use as a fit model:

$$\rho_{i,j} = \frac{1}{1+\frac{d_{i,j}}{(\Delta t)(relative\ speed)}} \tag{14}$$

The relative speed of the clouds is an unknown factor in this research so a the fit mode from Lave and Kleiss (2013) is used for the correlation. This fit model is slightly adjusted similar to Equation 12 an extra parameter is added. The new equation can be seen in Equation 15. This definition is similar to Equation 11, but subtly different, the fit does not increase with distance but decreases with distance. The reason for this[4], is that a correlation of 1 states that $\Delta K^*_{\Delta t,i}$ is the same as $\Delta K^*_{\Delta t,j}$, where zero variance states that $\Delta K^*_{\Delta t,i}$ is the same as $\Delta K^*_{\Delta t,j}$. As can be derived from equation 12, Perez et al. (2012) also uses covariance of the standard deviation of σ where Elsinga and van Sark (2013) compare σ_i with σ_j.

$$\rho_{i,j} d_{i,j} = a\left(e^{-\left(\frac{d_{ij}}{b}\right)}\right) + c \tag{15}$$

[4] This is also the same for $K^*_{t,i}$ and $K^*_{t,j}$

6.3.4. Comparison between Elsinga & Perez

A comparison was made of the two methods for multiple reasons: (1) to see if there is a difference of results with the updated database of Elsinga (2014) and (2) to measure the effect of both methods on the temporal correlation, spatial correlation and (3) the effects of both with a separation in wind directions.

For the spatial effect, b in formula 11 and formula 13 will be used to measure the decorrelation length. Elsinga and van Sark (2014) uses 3b as the decorrelation length where Perez uses a single b as decorrelation. Elsinga and van Sark (2014) mention that 3b is the range over the distance independent constant a is achieved with a 95% confident interval. But for easy comparison, single b will be used to compare the decorrelation length of both methods.

For the calculation of the temporal and spatial effect separated by the wind direction the clear sky index data was separated in different directions bins. This separation was based upon the azimuth and eight different direction bins were constructed. Eight direction bins of each bin results in 360/8= 45 degrees.

In **Fout! Verwijzingsbron niet gevonden.**, the PV systems are shown as they are situated in the city region of Utrecht. From a system selected in the city center of Utrecht (the yellow dot in **Fout! Verwijzingsbron niet gevonden.**) a selection is made based upon a chosen direction. For this figure, this direction bin is from 0 degrees to 45 degrees, so from the North Pole towards the northwest. Subsequently, this selection is limited for a certain distance. In this case, the limited distance is 15 km, and this selection can been seen in **Fout! Verwijzingsbron niet gevonden.** with the red dots.

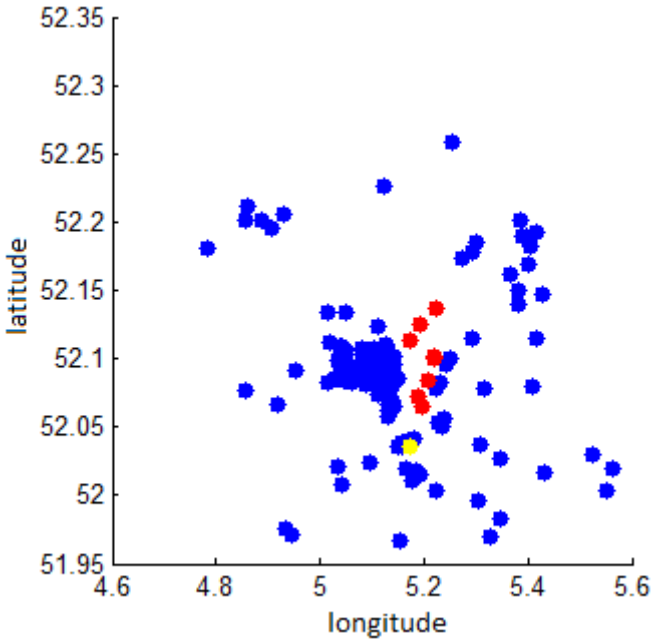


Figure 6. A geographical layout of the PV systems (of figure 2) with a highlighted selection of a chosen direction bin and maximum distance.

With the clear sky index input data, sorted by direction- and distance bins, it becomes possible to compare the correlation separated in wind direction for the spatial effect. The correlation is done with K_t^* and $\Delta K_{\Delta t}^*$ data. When also averaging the clearness index over different time resolutions, the effect of wind direction can be calculated with different time series.

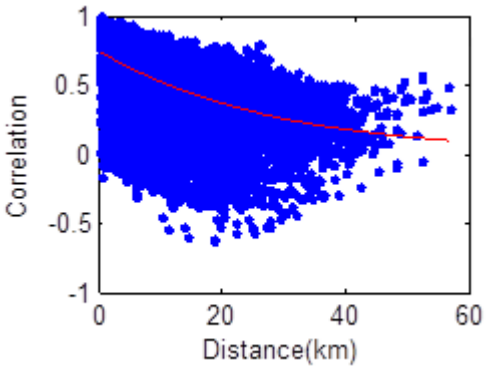


Figure 7. Correlation of all systems combined for a single day (23-04-2014) versus the distance between panels .

This method only works for correlation of an individual system. As can be seen in figure 3 and 6, the distribution of the PV systems is nonrandom; this causes that some directions are over presented, as for example the 315-360 angle in figure 6. While other directions may be under presented, e.g. a system at the North West corner will not have PV systems measurements north or west. Idem, some PV systems might not have other PV systems in the neighborhood before the decorrelation length.

To compensate for this fact, a method similar to Hoff and Perez (2012) and Hinkelmann (2013) is used. This system combines all correlation coefficients for all i,j pairs and this increases the data points from N to $N^2 - N$. Then the correlation of all i,j pairs versus the distances between all i,j pairs are fitted with the corresponding equation 12 or equation 15. See figure 7 for an example.

The article of Hinkelmann (2013) also splits the correlation of all i,j pairs in direction bins. In this article this results in cross wind and against wind pairs (see figure 8). Since the correlation between pair (i,j) is the same correlation as the correlation of inverse pair (j,i). This results, in that the correlations of the chosen wind direction bin are the exact same correlations as that of the inverse wind direction bin. This research specifies the direction bins even more and uses eight direction bins, as mentioned in the top of this paragraph.

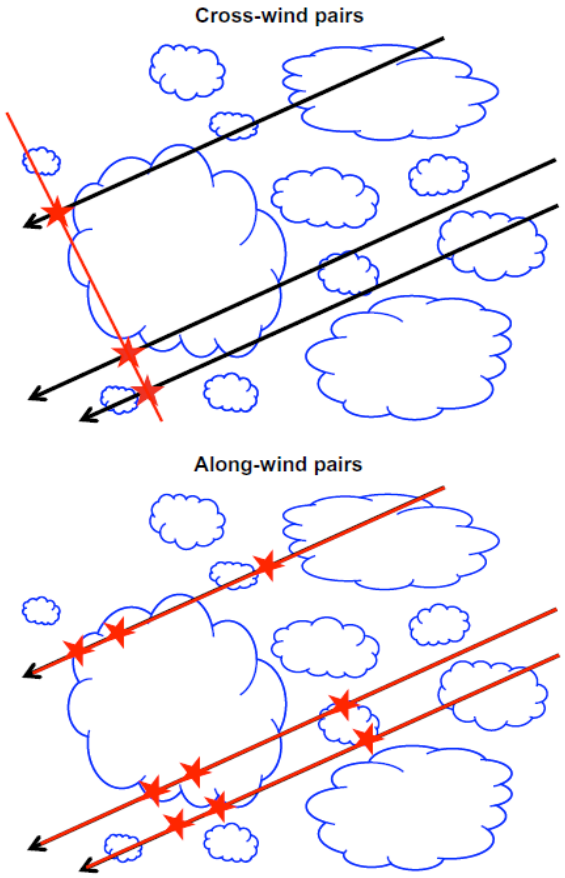


figure 8. Schematic view of cross wind and along wind pairs. From "Differences between along wind and crosswind solar irradiance variability on small spatial scales" by Hinkelmann, Solar Energy, 88, 194, 2013.

7. Results

In this section the results of the different methods will be given. First, an analysis of the data provided is given. Second, the temporal and spatial effect will be examined with the method of Elsinga and van Sark (2014) and the method of Perez et al. (2013). This will be done on three different levels: (1) Individual systems (2) all systems and (3) all systems for multiple days. Third, the results of the separation of wind data will be given for a single day as an example and for multiple days.

7.1. Data

Data was gathered during two years from power output measurements. Next, the data was converted to 5 second K_t^* and used to calculate $\Delta K_{\Delta t}^*$. Subsequently, the variance and correlation among different PV systems was calculated. In figure 7 two typical semi clouded days with different clearness indexes can be seen. The upper left graph is the K_t^* in February 2014, with a lower Clearness Index in the city region of Utrecht than that of August 2015, which can be seen in the upper right graph.

In the lower left graph the $\Delta K^*_{\Delta t}$ of 23 February 2014 is found and on the lower right, the $\Delta K^*_{\Delta t}$ of 20 August 2015. The variation in August is entirely different than in February: the maximum variation of Clearness Index is 2 times bigger on August than on February, while there is less variation in the morning and late afternoon. This phenomena are caused by different weather situations (see Attachment 1).

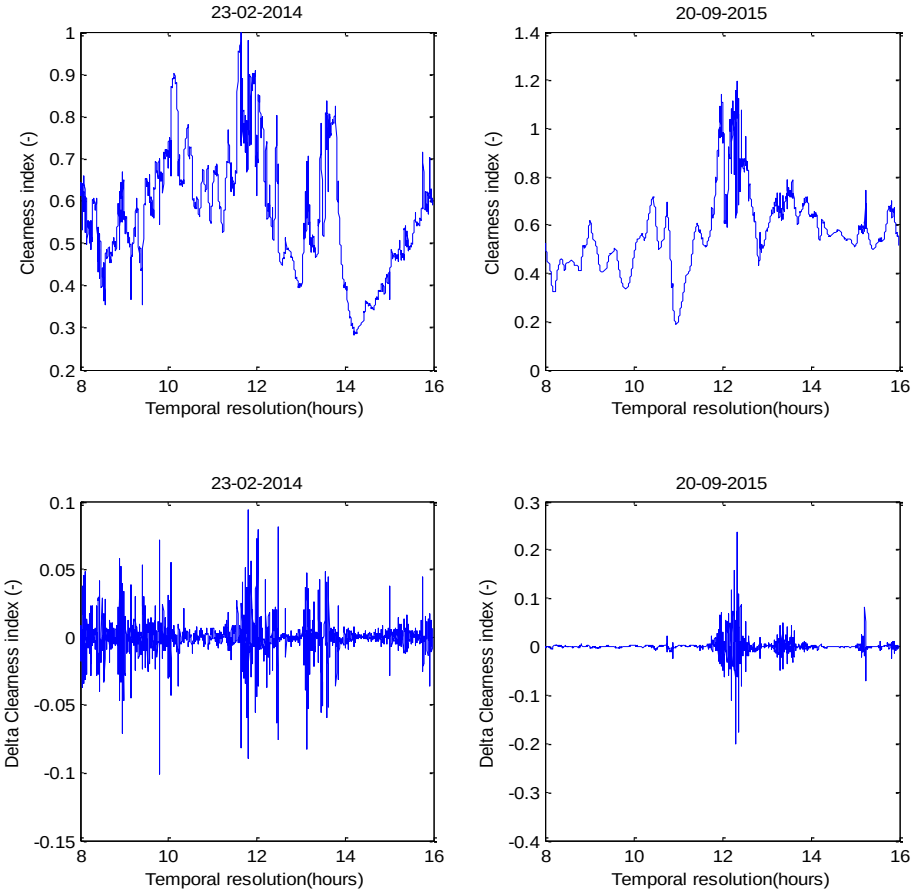


Figure 9. The clearness index and delta clearness index 23-02-2014 and of 20-09-2015.

7.2. Temporal and Spatial effect

In this chapter the results of the analysis of the temporal effect and the spatial effect on solar energy are depicted. In the results a separation is made with four different methods: (1) 'Elsinga', which is short for standard deviation calculated with $K^*_{t,i,j}$, (2) 'Perez' which correlates $K^*_{t,i,j}$ with the Pearson correlation. (3) 'DeltaElsinga', which shows standard deviation calculated with $\Delta K^*_{\Delta t,i,j}$ and (4) 'DeltaPerez' correlates $\Delta K^*_{\Delta t,i,j}$ with the Pearson correlation. Firstly, the results for an individual system for a single day will be depicted, then for multiple systems and lastly the combination of multiple systems on multiple days.

7.2.1. Individual systems

For the calculation of the spatial effect of power output between different PV systems on five second based data, it was needed to calculate the distance between each system (i) and all other systems (j). After the calculation of the distance, the variance and correlation for each i,j pair was calculated. In figure 10, the correlation on a semi clouded day can be seen, as the PV systems get lighter (yellow) color when moving closer to the location of the system where all other systems are correlated with (the X mark). Likewise visible is the decorrelation, as the PV systems in the outer circle/space get closer to a blue color, indicating less mutual connection with the X mark.

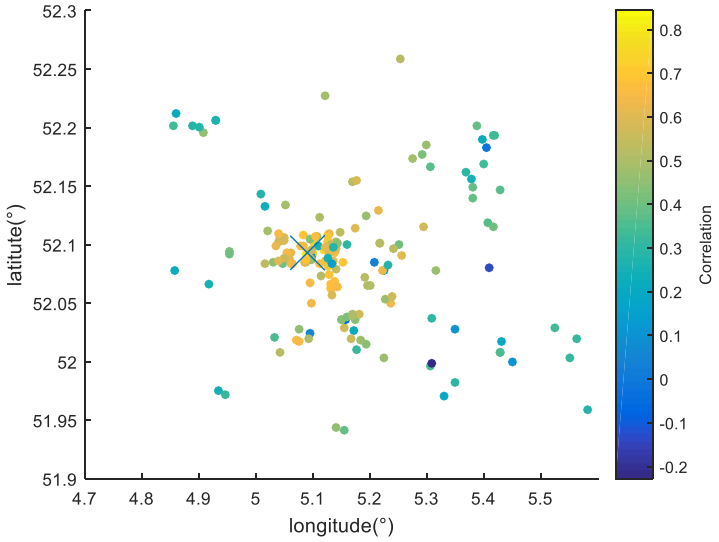


Figure 10: Correlations of different system specified per location for 23-02-2014. X marks the chosen system and the correlation is higher or lower depending on the colour.

The system used for comparison (the X mark in figure 10, here after referred to as system X) is picked as a PV system situated in the middle of the city of Utrecht with no shading effects of nearby buildings of trees for two hours after sunrise and two hours before sunset. The time resolution in figure 10 is set up to 20 minutes, since as mentioned in section 6.3.3, the correlation increases with a higher time resolution. In attachment 7, the correlation of different systems of multiple days with this time resolution can be found.

When the correlation is plotted versus the distance between station (figure 11) it becomes possible with a fit (the red line), based upon the equation 12 and 15, to calculate the decorrelation length (b in the fit model) and the R-squared of the fit. In figure 12, the same PV system is chosen as in figure 10, with the same day but with a much lower time resolution, to show the effect that the time resolution has on the different methods.

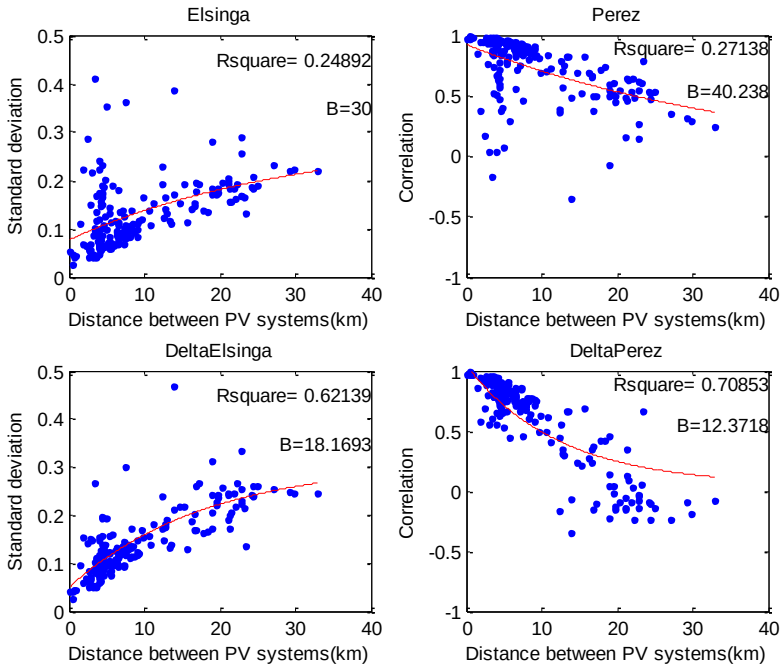


Figure 11. Comparison of methods for system X on 23-02-2014. With time resolution of 20 minutes. Inserted in the graphs is the decorrelation length (B) and the R-squared

For ‘Perez’ and ‘DeltaPerez’ (see figure 11 and 12), negative correlation can be observed. Negative correlation indicates according to Hinkelmann (2013) that “Clouds tend to reach one site as a cloud passes off the second site”. This phenomena depends on the size and speed of the clouds. Perez et al. (2012) give the partial cancellation of fluctuations by the movement of clouds as reason for the negative correlation.

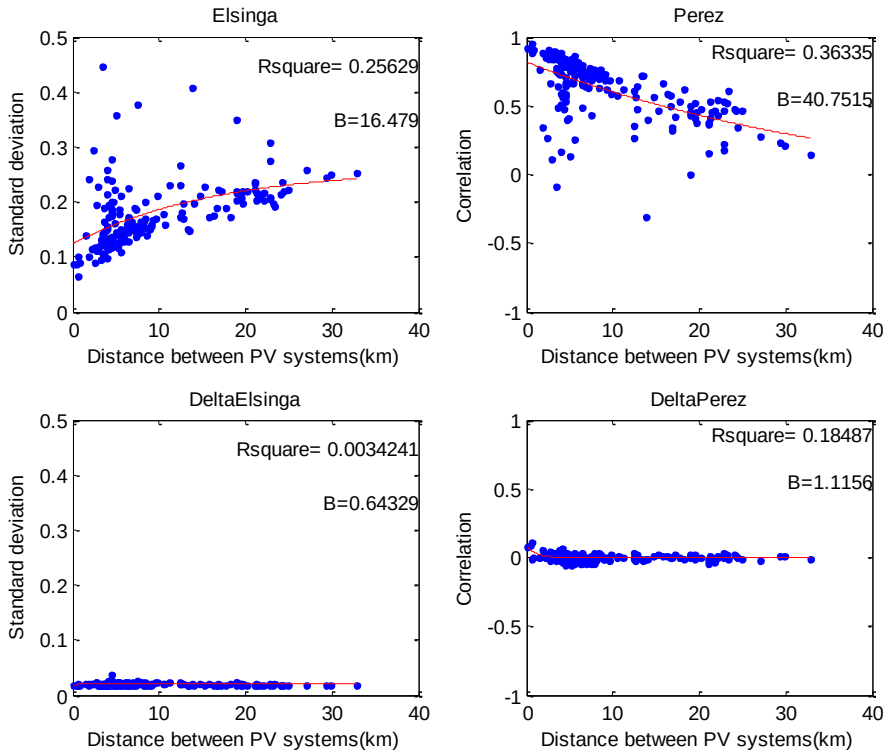


Figure 12. Comparison of methods for system X on 23-02-2014. With time resolution of 5 seconds. Inserted in the graphs is the decorrelation length (B) and the R-squared

In figure 11, the distance between PV system (X) and PV system (j) is plotted against the correlation and then fitted (see the red line). For a time resolution of 5 seconds the decorrelation length of the methods of “DeltaElsinga and ‘DeltaPerez’ is low. Even lower than the distance between PV system (X) and the closest PV system (which is 400 meter). This indicates that the decorrelation length is smaller than 400 meter for 5 second data. This result is comparable with the methods observed in Elsinga and van Sark (2014), where the smallest decorrelation lengths of 100 meter is observed.

'Perez' and 'Elsinga' have a decorrelation length of respectively 16 and 40 kilometer in figure 11 and 30 and 40 kilometer in figure 11. There is no significant decorrelation observed in both methods with different time resolutions. This is caused by the principle that averaging over different time resolutions does not influence the relative variance.

If figure 11 and figure 12 are compared it becomes visible that, similar to Perez (2011) and Elsinga (2015), the decorrelation length increases significantly with different time resolutions for 'DeltaPerez' and 'DeltaElsinga'. This is further explained in figure 13.

In figure 13 below, the decorrelation length of the four methods on a minute time resolution scale is given. Inserted in the graph is the decorrelation length of the four methods on a seconds time resolution scale. In the upper figure in figure 13, the accompanying R-squared of the 4 methods is depicted with also an insertion on a smaller time resolution scale.

The decorrelation length of 'Elsinga' and 'Perez' for the smallest time resolution is respectively 5 and 15 kilometer. An value higher –with comparable time resolutions- than seen with in the articles of Perez et al. (2012), Elsinga and Van Sark (2014) and Hinkelmann (2013). The R-squared does not increase with higher time resolutions, with the same data as is similar to figure 11 and 12. 'DeltaElsinga' and 'DeltaPerez' show a mayor improvement in the first three minutes of the R-squared, where 'DeltaPerez' does not improve after a 15 minute time resolution while 'DeltaPerez' still increases with this time resolution. This is most likely due to random errors of the system. The increase of decorrelation length 'DeltaElsinga' and 'DeltaPerez' with higher time resolutions is similar to the increases observed in Perez et al. (2012) and Elsinga and Van Sark (2014)

When looking at the inserted graphs in figure 13, the R-squared of 'DeltaElsinga' is near 0, this indicates that no fit could be found by the system. The accompanying decorrelation length is therefore most likely to be smaller than 400 meter (as observed in figure 12). 'DeltaPerez' shows doubling of the R-squared with the time resolution of 5 to 25 second and a small increase in decorrelation length.

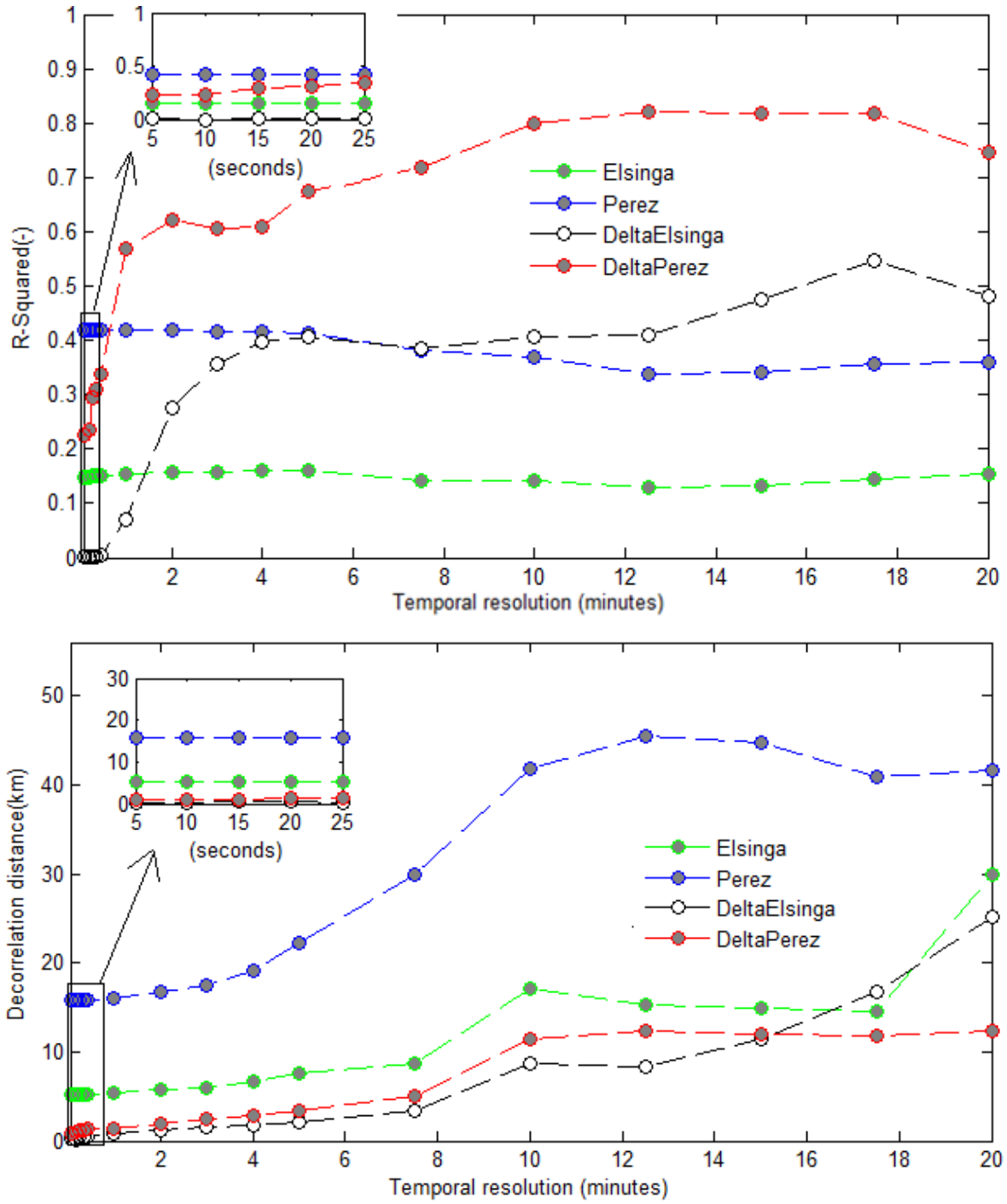


Figure 13. Above: the R-squared of the four methods for system X on 23-02-2014 on a temporal resolution of minutes. Inserted is the temporal resolution on seconds base. Below is the corresponding decorrelation lengths of the R-squared above. Inserted in the figures is also the temporal resolution on a seconds time resolution.

A selection of multiple days was made based upon the methods mentioned in 6.2.1. The weather data of this days can be found in attachment 1 and the accompanying clearness indexes in attachment 2. When comparing the decorrelation length of 'DeltaPerez' for the chosen days, it shows that the decorrelation lengths are not similar for every day, see figure 14. The difference between the days are caused by the differences in weather circumstances. Although a careful selection has been made, the weather is never similar. Figure 14 shows that the decorrelation length increases with a higher time resolutions for every day. The R-squared for most days increases significantly within 5 minutes time resolution. Only for the 8th of March 2015 the R-squared does not increase, although the decorrelation length does increase with a higher timer resolution on this day.

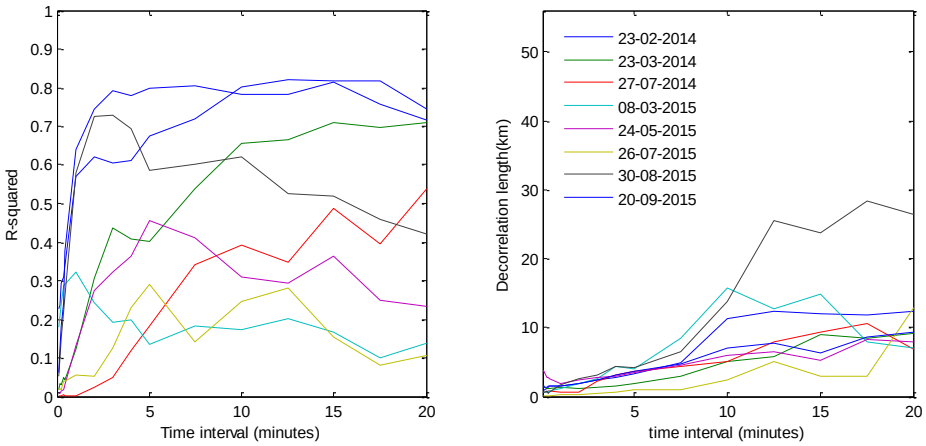


Figure 14. On the left R-squared of DeltaPerez with different time intervals on multiple days with similar weather for system x. on the right the corresponding decorrelation lengths.

In table 2, the decorrelation distance of system x for 5 second data is shown. The 'Perez' method gives decorrelation distances at the maximum of the model or near maximum, while decorrelation distances of 0-3000 meter should be expected according to Elsinga and Van Sark (2015). 'Elsinga' has decorrelation distances which are for most days significantly higher than the 3000 meter mentioned above. 'DeltaPerez' and 'DeltaElsinga' vary each day where most of the days 'DeltaPerez' is higher while the decorrelation length of "DeltaElsinga is lower. The choice of method to calculate the decorrelation length is therefore very important, since results vary for each method.

Table 2. Decorrelation distances (in m) of system X for 5 seconds for days with similar weather.

	23-02-2014	23-03-2014	30-03-2014	27-07-2014	09-11-2104	09-08-2015	30-08-2015	18-10-2015
Elsinga	5242	12572	8870	6073	11391	4163	2262	4342
Perez	15804	31154	56000	9548	15787	6136	3286	19859
DeltaElsinga	242	193	173	223	232	357	3604	196
DeltaPerez	882	1330	2147	547	3683	1800	683	1829

7.2.2. All Systems

With the four methods described above there are only a maximum of 200 data points available each day. This is because there are 200 systems compared with 1 chosen system. When all systems are compared with each other there are 201^2-201 = 40200 data points available. This increase of data points is expected to give better results and is similar to the method used in Hoff and Perez (2012). This method gives far more data for fitting than for an individual system. Some PV systems are less than 200 meters apart, where for system X the nearest neighboring PV system is 400 meter. The inclusion of data points before the 400 meter might include better results.

In figure 15, the correlation and variance of K_t^* and $\Delta K_{\Delta t}^*$ of the second day of February 2014 is taken and plotted against the decorrelation distance for a time resolution of 5 seconds. The effects shown in figure 16 are similar to the effects in figure 12, but the spread of data is bigger for 'Perez' and 'Elsinga', while 'DeltaPerez' has a higher R-squared than in figure 9. This has to do with the fact that data points are available before the decorrelation length and therefore a better fit can be found. With 'DeltaElsinga' some points have a slightly higher variance than the other points, which is caused by the fact that one system delivered faulty data for this day, so the variance if higher.

In figure 16, for every distance bin of 100 meter the average correlation and variance of K_t^* and $\Delta K_{\Delta t}^*$ is used, similar to Hinkelmann (2013). The standard deviation of 'Elsinga' shows no significant increase over time, as seen before in figure 13. The correlation of Clearness index shows a decrease with distance, but no decorrelation length can be observed. With 'DeltaElsinga' and 'DeltaPerez' a sudden increase can be seen at the beginning similar to figure 14.

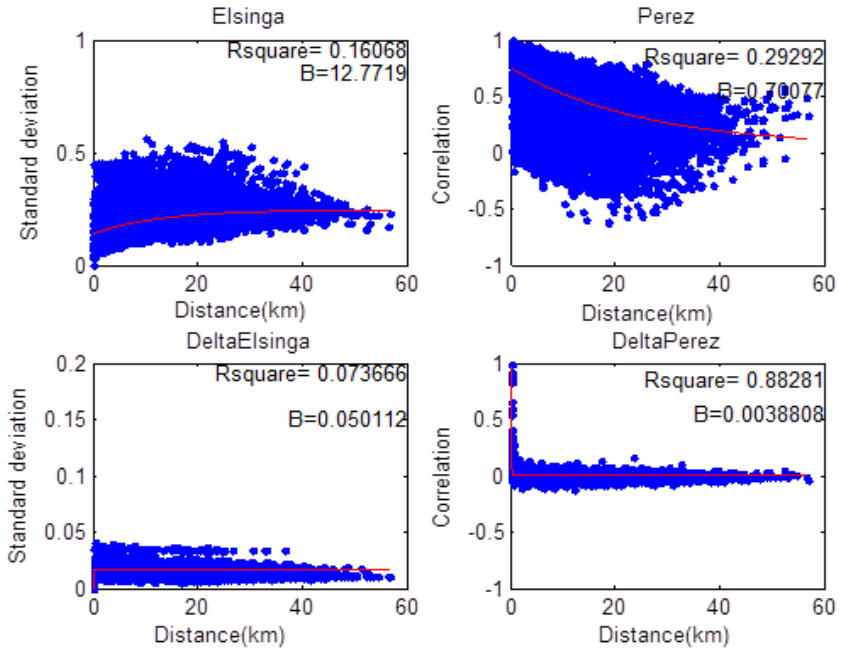


Figure 15. Decorrelation length of all systems combined of 23-02-2014. With methods clockwise: 1. Elsinga variance of K_t^* 2. Perez correlation with K_t^*. 3. Perez correlation with $\Delta K_{\Delta t}^*$ and 4. Elsinga variance with $\Delta K_{\Delta t}^*$. Inserted in every figure is the R-squared and the decorrelation length (b).

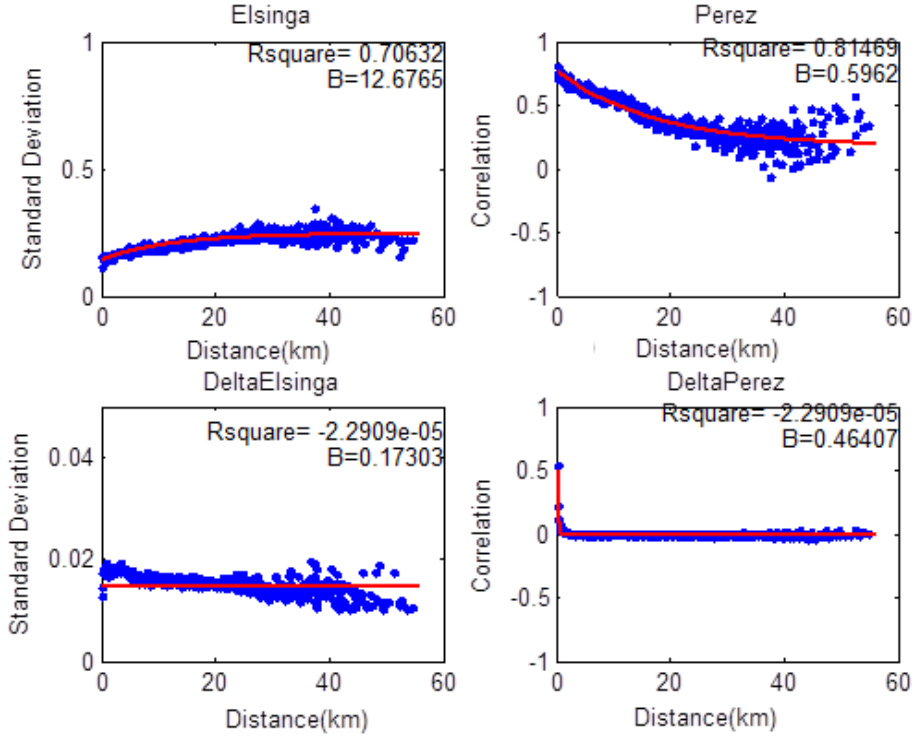


Figure 16. Decorrelation length of all systems combined averaged for 100 meter of 2014-02-23 with the standard deviation of the averages. With methods clockwise: 1. Elsinga variance of K_t^*. 2. Perez correlation with K_t^* . 3. Perez correlation with $\Delta K_{\Delta t}^*$ and 4. Elsinga variance with$\Delta K_{\Delta t}^*$. Inserted in every figure is the R-squared and the decorrelation length (b).

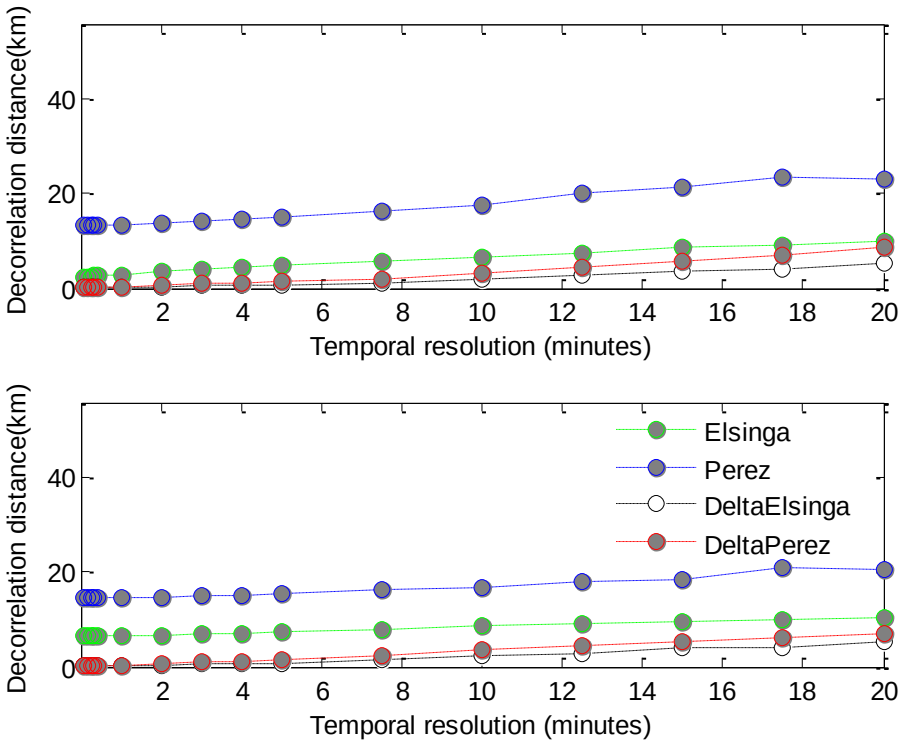


Figure 17: Decorrelation length of all methods for all systems on 23-02-2014 with different temporal resolutions. Above: the different methods with non-averaged distance bins. Below: the different methods with distance bins of 100 meter.

When comparing the decorrelation length for different temporal resolutions on a single day as in figure 16, it becomes visible that there are only small differences between using the average over distance bins and using non averaged distance bins. The decorrelation length of 'DeltaPerez' and 'DeltaElsinga' stays constant (and near 0) for the first minute and gradually start to rise to about 10 km with a temporal resolution of 20 minutes. 'Perez' and 'Elsinga' start at a high decorrelation length and only slightly increase with time. These results are comparable with what is observed in figurer 15 and 16. The R-squared of all these results are higher than observed in section 7.2.1. This is, as explained before, because of the higher amount of data points that are available.

A comparison has been made of the decorrelation length with a selection of days with high variability of clearness index and semi clouded according to the KNMI, (see attachment 1). This selection is the same selection of days as in figure 14.

In figure 18, as has been seen before the correlation length does in general increase with distance, but the slope is volatile and for some days even decreasing with a larger time scale. Especially after the 15 minutes time interval with non-averaged data, the slope increases with multiple days. This is most likely caused by the fact that the spread of data points becomes larger with every time interval increase and too large for some days after the 15 minutes time interval. With 'DeltaPerez' for the 24th of May 2015 the decorrelation length at 17 minutes is near 0. The model could not find a suitable fit with specific settings.

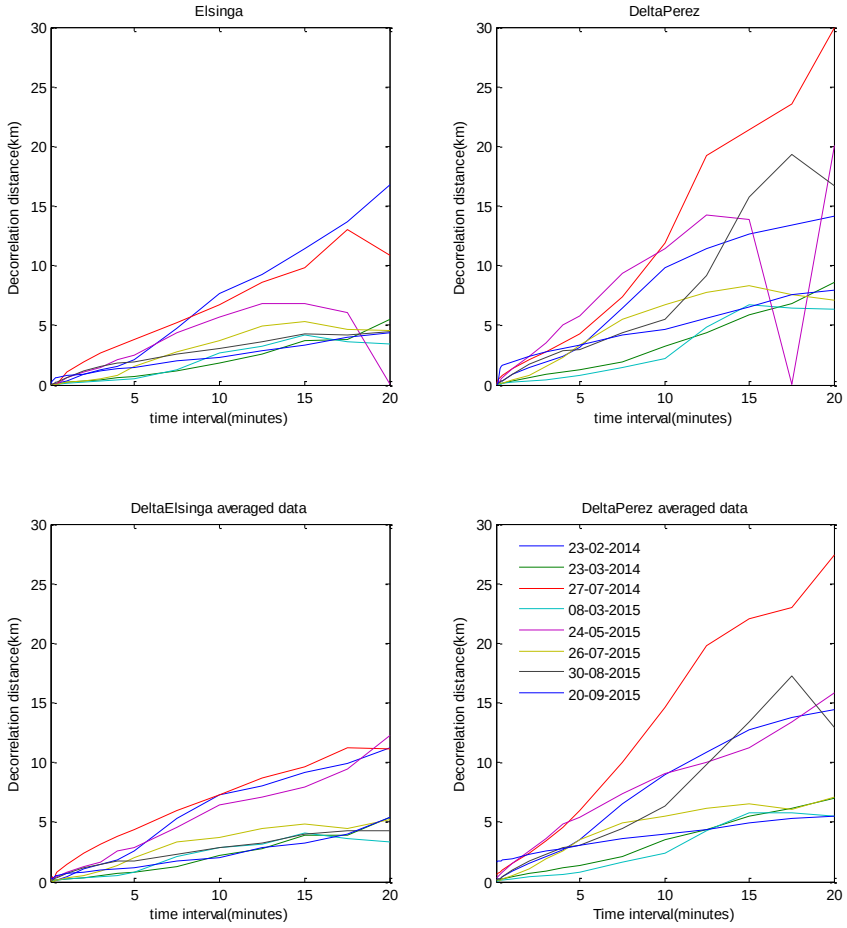


Figure 18. comparison the decorrelation length of DeltaElsinga (upper left) and DeltaPerez (upper right) for multiple days with similar weather. Below are the results averaged per 100 meter distance bin.

Table 3.
Decorrelation distances (in m) of all systems combined for 5 seconds with different methods for days with similar weather

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2014	9-8-2015	30-8-2015	18-10-2015
Elsinga	12772	2470	7900	2712	4661	302	2403	4375
Perez	25295	13190	49264	9661	9	14	3005	14718
DeltaElsinga	4	21	4	12	4	16	10	183
DeltaPerez	119	63	11	30	205	62	105	6

Table 4.
Decorrelation distances (in m) of all systems combined for 5 seconds with different methods for days with similar weather with averaged data over a distance bin of 100m.

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
Elsinga	12676	6716	9073	5783	5888	3398	2504	4827
Perez	17474	14573	29757	19386	12603	3961	3055	9521
DeltaElsinga	464	57	0	99	217	134	136	375
DeltaPerez	151	121	632	83	368	106	143	1665

When comparing the decorrelation length for multiple days on a 5 seconds scale in table 3, the spread of the results is imminent. The method 'Perez' constantly has a decorrelation length that is too high to be true, as can be seen in figure 15 and 17. For every day the different methods give different decorrelation lengths, which is caused by the fact that days are chosen with broken clouds with never entirely similar weather (see Attachment 1). Table 4 depicts very similar results as table 3, but the decorrelation length is higher for most days. This is most likely caused by taking only the average of all correlations in the first distance bins (the first 100 meter).

When comparing general trends, 'DeltaElsinga' has a lower decorrelation length compared to the other methods and non-averaged data works better on a smaller time resolution while average data in distance bins works better with larger time resolutions. 'DeltaPerez' shows slightly higher decorrelations lengths than "DeltaElsinga' but the general trend lines are similar. 'Perez' and 'Elsinga' are non-viable methods to calculate the decorrelation length.

7.3. Wind Direction

For this chapter the correlations and standard deviations of clearness index and delta clearness index was separated in different wind directions. First an analysis of decorrelation lengths for different time resolutions on a single day will be given and second an analysis of the decorrelations lengths of multiple days.

7.3.1. Single Day

When separating the K_t^* and $\Delta K_{\Delta t}^*$ of the 23th of February 2014 in eight different directions and calculating the accompanying decorrelation distances for different time intervals, it becomes clear that - as expected- there is a difference between decorrelation length for each direction. The result can be seen in figure 19. As explained in section 6.3.4, only four directions from the eight directions are visible. To recapitulate: the reason for this is that for each system the chosen direction is exactly the same as the inverse of the chosen direction (e.g. 45°- 90° is the same as 215°- 270°).

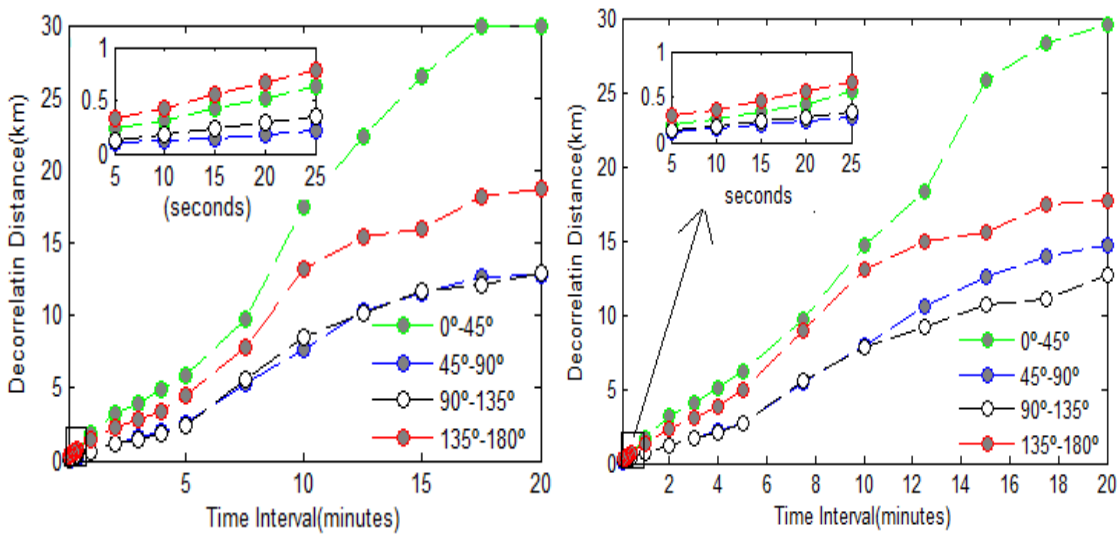


Figure 19: Comparison of decorrelation length of 'DeltaPerez' for 23-02-2014 separated in 8 different angles.

The decorrelation in figure 17 of 'DeltaPerez' is about the average decorrelation of the lines in figure 19, although the direction of 0°- 45° degrees has a two and a half times higher decorrelation distance at a time resolution of 20 minutes. The wind direction of 135°- 180° has on average a one and a half times higher correlation than the lowest two direction bins. To compare the separated directions with the wind direction, the median wind direction measured on the ground was calculated as: 180 ° at the 23th of February 2014 (see attachment 1).

The highest correlation direction is not similar to the wind direction on any time resolutions. This could be explained by multiple things: (1) the wind direction is exactly 180° , which is the difference between 135°-180° and 0°-45° (remember that 180° is also 0° with this method), (2) The wind direction on the ground is slightly different than the direction of the clouds and (3) -as can be seen in figure 3 - the PV systems are non-evenly distributed and therefore there are more data points before the decorrelation length with the 0°-45° direction than on 135°-180°.

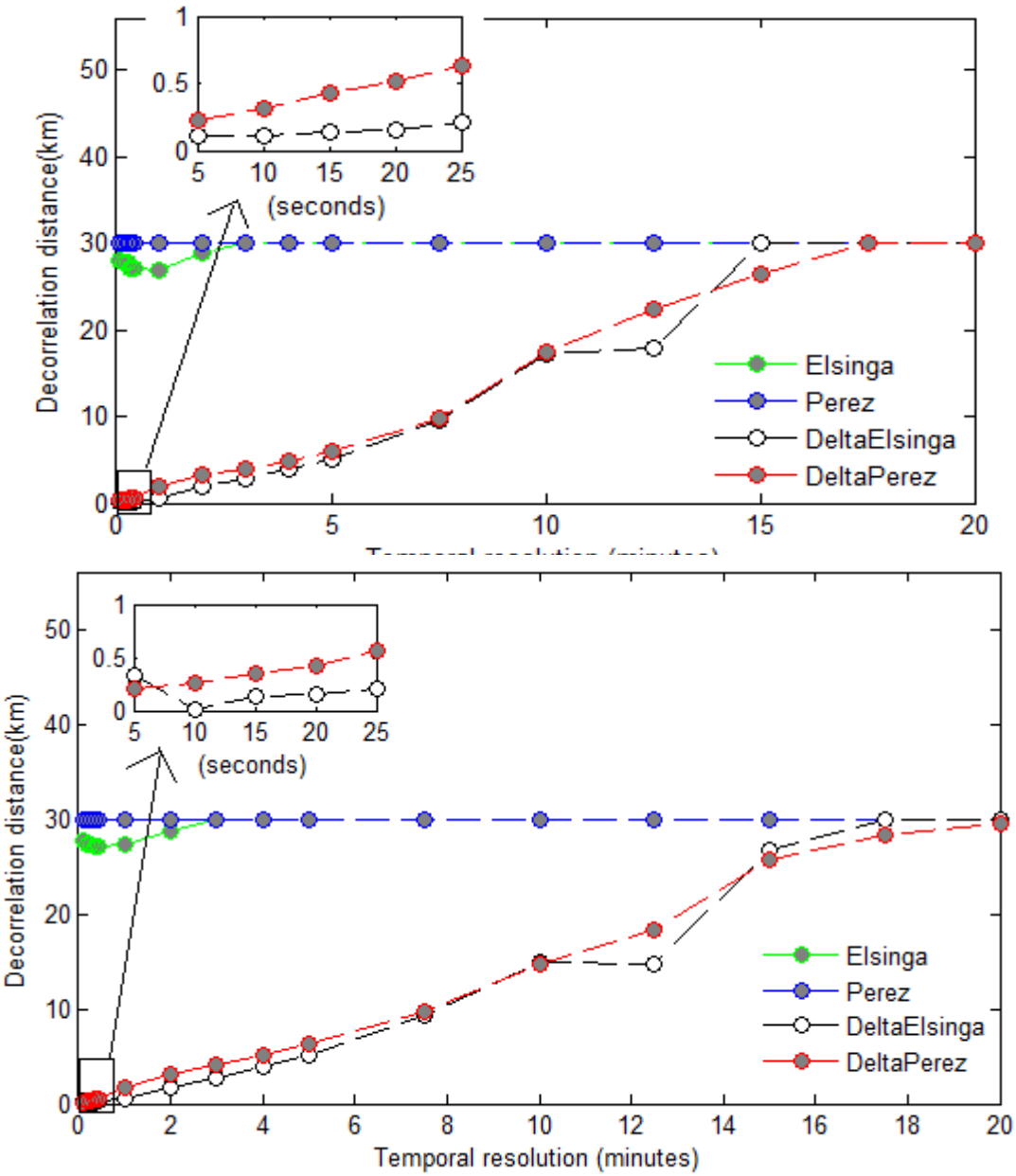


Figure 20. The decorrelation length of different methods over different temporal resolutions of direction bin of 0°-45° on 23-02-2014. Above with non averaged data and below with averaged data.

Where in figure 18 the decorrelation lengths of all direction bins are depicted, in figure 20 only the 0°-45° direction bin is shown to compare the different methods. Where in figure 18, before the decorrelation length of 'Perez' and 'Elsinga' slightly increased with time, they do not increase anymore in this picture. The decorrelation length of 30 kilometre is the maximum where the model correlates with, this means that the decorrelation length is even higher than 30 kilometres for both methods even on a small time resolution. The temporal resolution of 'DeltaElsinga' and "DeltaPerez' are similar in this picture. When comparing these methods with previous results (figure 17) it shows that the decorrelation length is much higher for some direction bins than none separated data.

7.3.2. Multiple Days

When plotting the decorrelation distance of the same eight days as used in the previous calculations, it becomes clear that the maximum decorrelation distance can be different for each day when looking at 5 seconds data or at 20 minutes data. The plots shown in figure 22 are so called "angle histogram plots": graphs where the eight wind direction bins are plotted in eight pie chart on a circle, where every pie chart is presented at its corresponding direction.

The prevalent wind direction on these eight days according to the KNMI is shown as the red triangle in figure 22. This figure shows that the prevalent wind direction not always the same as the highest decorrelation length. The reasons for that are already mentioned before: (1) The wind direction is measured on the ground and can therefore be different than the direction of the clouds, (2) the PV systems are non-evenly distributed and therefore the data points are non-evenly distributed and (3) the wind direction can be between direction bins.

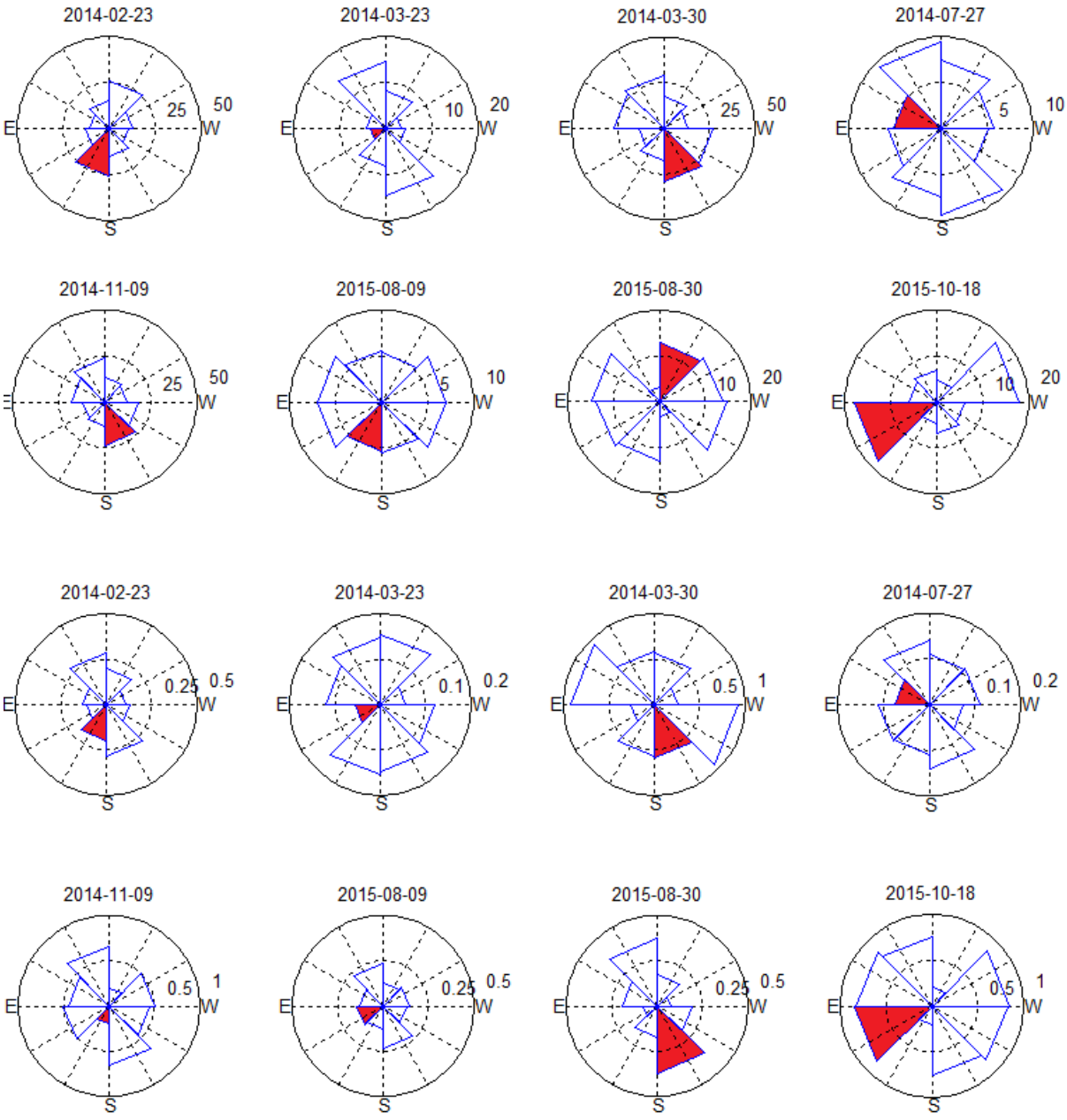


Figure 21: Angle histogram plots of eight decorrelation distances of 'DeltaPerez' of eight different days, with the first eight on 5 second data and the latter eight with 20 minutes data.

Table 5.
Maximum decorrelation distances (in m) of all systems split in direction s combined for 5 seconds with different methods for days with similar weather

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
Elsinga	29999	9747	29720	10782	8079	6558	9002	17414
Perez	-	23422	-	17651	12825	7283	11372	30000
DeltaElsinga	246	182	157	23235	496	200	7328	1270
DeltaPerez	331	180	982	189	1147	2525	1052	1131

Table 6.
maximum decorrelation distances (in m) of all systems split in direction s combined for 5 seconds with different methods for days with similar weather with averaged data over distance bins of 100 meter.

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
Elsinga	27678	6731	15062	9147	7348	5225	-	18998
Perez	-	-	-	17511	11437	8934	-	-
DeltaElsinga	628	391	142	12019	415	464	4358	295
DeltaPerez	288	156	928	140	664	233	371	858

The maximum decorrelation distance of the same eight days as used in the previous calculations are shown in table 5 and 6. When comparing the results with table 3 and 4 the same effects can be seen, although for 'DeltaElsinga' and 'DeltaPerez' for every single day table 5 and 6 have a higher maximum decorrelation length. Even splitting the data into different directions is to no avail for 'Perez' and 'Elsinga', there is still no realistic decorrelation length found. 'DeltaElsinga' has a decorrelation length of near 20 km for 5 second data on the 27th July of 2014. Which is too high when compared to other literature (Hoff & Perez, 2012; Elsinga & Van Sark, 2014; Hinkelmann, 2013). When looking at the specific details it shows that the decorrelation length is smaller than the first data points.

In table 7 the increase of decorrelation length is shown for 'DeltaPerez' and 'DeltaElsinga'. This is done by subtracting the results of table 5 from table 3 and table 6 from table 4. For the 18th November 2015 there is a decrease in correlation length instead of an increase. When looking at the weather data in attachment 1, it shows that this day is very variable, but does have a clouded sky. Which is very different than the other days. This result is seen as an outlier. The 27th July 2014 has an increase of more than 10 kilometers, because the maximum wind directions decorrelation length of this date is a fault in the calculation.

Table 7.
The increase of decorrelation length when comparing all systems with the maximum decorrelation length when separated in wind direction

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
DeltaElsinga	243	161	153	23223	492	185	7318	1087
Perez	212	116	971	159	942	2463	947	1125
DeltaElsinga with averaged data	164	334	142	11920	198	329	4222	-80
DeltaPerez with averaged data	137	34	295	57	295	127	229	-807

In table 7 and 8, the along wind correlation and against wind decorrelation lengths are shown. The along wind correlation is the decorrelation length of the direction bin where the wind direction is assigned to (see the red direction bin in figure 22). The against wind direction is the perpendicular direction of the along wind direction.

Table 8.
Along wind decorrelation distances (in m) of all systems split in direction s combined for 5 seconds with different methods for days with similar weather with averaged data.

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
Elsinga	6296	6731	3066	4573	6369	4206	2531	15766
Perez	23883	7580	6321	17511	11437	8934	-	28062
DeltaElsinga	537	391	142	2455	415	46	589	138
DeltaPerez	288	57	585	76	664	130	173	858

Table 9.
Against the wind decorrelation distances (in m) of all systems split in direction s combined for 5 seconds with different methods for days with similar weather with averaged data.

	23-2-2014	23-3-2014	30-3-2014	27-7-2014	9-11-2104	9-8-2015	30-8-2015	18-10-2015
Elsinga	17759	2904	15062	619	6648	4761	2531	17414
Perez	22429	-	-	10513	9936	4837	-	-
DeltaElsinga	628	140	65	101	126	238	589	95
DeltaPerez	106	147	256	110	508	131	173	1081

When looking at the specific along wind correlation in table 8 and comparing these results with table 4, an increase of decorrelation distance is visible for most days for the methods of 'DeltaPerez' and 'DeltaElsinga'. As seen before: 'Elsinga' and 'Perez' have a nonrealistic decorrelation distance for along wind correlation. If table 8 is compared to table 6, the correlation length of the along wind direction bin is not always the highest correlation as figure 22 also showed.

If the against wind correlation in table 8 is compared to the against wind correlation length in table 4 the decorrelation distance decreases for five out of the selected eight days for 'DeltaPerez' and 'DeltaElsinga'. Figure 22 shows that the highest correlation, when looking at different time resolutions, the against wind correlation is never the highest correlation for the selected days. In Attachment 3 the decorrelation distances with 1 minute data and with distance bin of 500 meter is shown.

In table 9 and 10 the direction with the maximum wind correlation are compared with the median wind direction of the day. None of the methods or time resolutions have the highest correlation consistently as the wind direction. 'DeltaPerez' with a time resolution of 5 seconds is the method where most direction bins are the same as the wind direction. These results indicate that these methods are not a consistent way to predict the wind direction with the maximum wind correlation distance. In Attachment 5 the direction with the maximum wind correlation is even specified further, but there similar effects are witnessed.

Table 10.
Direction with maximum wind decorrelation distances with different methods for days with similar weather with non-averaged data

	Median wind direction	5 second Delta Elsinga	5 second Delta Perez	20 minutes Delta Elsinga	20 minutes Delta Perez
02-23-2014	180°	45°- 90°	135°- 0°	0°- 45°	0°- 45°
23-3-2014	230°	45°- 90°	0°- 45°	0°- 45°	135°- 0°
03-30-2014	150°	135°- 0°	90°- 135°	45°- 90°	135°- 0°
27-7-2014	100°	45°- 90°	135°-0°	45°- 90°	135°- 0°
11-09-2014	160°	135°- 0°	135°- 0°	0°- 45°	135°-0°
9-8-2015	10°	45°- 90°	135°- 0°	45°- 90°	45°- 90°
30-8-2015	30°	45°- 90°	135°- 0°	45°- 90°	90°-135°
18-10-2015	70°	0°- 45°	45°- 90°	45°- 90°	45°- 90°

Table 11.
Direction with maximum wind decorrelation distances with different methods for days with similar weather with averaged data

	Median wind dir	5 second Delta Elsinga	5 second Delta Perez	15 minutes Delta Elsinga	15 minutes Delta Perez
02-23-2014	180°	90°- 135°	45°- 90°	45°- 90°	90°- 135°
23-3-2014	230°	135°- 0°	135°- 0°	0°- 45°	135°-0°
03-30-2014	150°	90°- 135°	45°- 90°	0°- 45°	135°-0°
27-7-2014	100°	135°- 0°	135°- 0°	45°- 90°	45°- 90°
11-09-2014	160°	90°- 135°	45°- 90°	45°- 90°	45°- 90°
9-8-2015	10°	135°- 0°	135°- 0°	45°- 90°	45°- 90°
30-8-2015	30°	90°- 135°	45°- 90°	45°- 90°	90°- 135°
18-10-2015	70°	135°- 0°	135°- 0°	0°- 45°	135°- 0°

8. Discussion

This study and the articles of Perez et al. (2012) found that decorrelation of solar intermittency decreases over distances less than 500 meter for 20 seconds time intervals. This analysis confirms these results. Although, decreasing the time resolution to 5 seconds has a very small effect on the decorrelation length compared to 20 seconds time intervals, when comparing the standard deviation and correlation of $\Delta K^*_{\Delta t,i,j}$. Elsinga and Van Sark (2014) argue that the decorrelation length of a time resolution of 1 minute is on average 340 meter, whereas Perez et al. (2012) found 850 meters as decorrelation length with that time resolution. The difference in decorrelation length is caused by the different circumstances in which both researches operated.

With the time resolution of 1 minute, the average decorrelation length is found in this study to be 300 meters with the updated database of Elsinga and Van Sark (2014). Thus, the results of the current study are very similar to what is observed in literature. The addition of 175 systems to the database only slightly changed the decorrelation length.

When comparing the maximum decorrelation distance while separating the data in eight different directions, multiple phenomena can be observed. The maximum decorrelation distance significantly increased (with +-200 meter) when not including the outliers. However, the maximum decorrelation distance is not always the same as the measured wind direction. This is in contradiction with Hinkelmann (2013). In her study, she found that the correlation between PV systems depends on their alignment relative to the wind on Oahu Island.

For this dissimilarity there are multiple reasons: (1) the wind direction is measured on the ground and can therefore be different than the direction of the clouds, (2) the PV systems are non-evenly distributed and therefore the data points are non-evenly distributed (some direction bins have more data points than others especially in the distance <1000 meter), (3) although careful selection of the days occurred, the wind direction can vary over the day and (4) clouds can have a size of multiple kilometers and can cover multiple PV systems at the same time. This last point causes correlation of PV systems which are not along the wind direction.

With an average distance of 9 kilometers between the PV systems in this measurements and only minimal data points of distances between PV systems smaller than 1000 kilometer (Elsinga and Van Sark, 2015). The locations of the PV systems used to measure the irradiance variability in the current study are not applicable for small time scales (and thereby small decorrelation lengths).

The current study shows that for short timescales (10 - 60 seconds) an opportunity is given to separate the dataset to distances between PV systems under 1 kilometer and to compare these results with a small scale. The exclusion of the large distances might give more specific details about why the decorrelation length is not as predicted in the 5 seconds time resolution (for this prediction see Perez et al. (2012). For an accurate prediction of variability of solar intermittency, it could be beneficial to compare the results with local cloud data gain from the same period as the measurements. This way the influence of cloud direction, size and speed on the variability can be calculated.

The exact influence of adding the extra fit parameter c in the fit models for both methods can be the subject of future research. Noticed is that fit parameter c does increase the R-squared when implemented in this research, but unknown is the quantitative amount. Also a comparison with the results of the distance independent value a, with the distance independent value research of Elsinga and Van Sark (2015) is an interesting subject to research.

The essence of calculating the variation of irradiance on PV systems is to know the exact variation of the electrical output of the PV systems when large scale PV is implemented. This study helps in the understanding of this variation, but does not assist in the understanding how to reduce this variation. Suggested therefore, are future studies on how to make a cost effective energy infrastructure which reduces the variation of multiple PV systems as much as possible.

9. Conclusion

The analysis undertaken to determine the relations of solar intermittency with inter-system correlation, spatial effect and temporal effect shows that the solar intermittency becomes uncorrelated within a distance less than 200 meter for broken cloud days with a temporal resolution of 5 seconds. The decorrelation distances increase to more than 5000 meter when looking at a 15 minute time resolution.

The linear increase of decorrelation distance with temporal resolutions seen in Perez et al. (2012) and Elsinga and Van Sark (2014) is also found in this research for time resolutions bigger than 1 minute. Smaller time resolutions were influenced by the lack of data points of distances between PV systems smaller than 1000 kilometer. The network used to measure the irradiance variability is not perfectly suited for small time resolutions (and with that small decorrelation lengths).

In this research, two methods to calculate the inter system correlation of $\Delta K^*_{\Delta t,i,j}$ are compared: (1) a method based on the article of Elsinga and van Sark (2014) that calculates the standard deviation of $\Delta K^*_{\Delta t,i,j}$ and (2) a method based on the article of Perez et al. (2012) to correlate $\Delta K^*_{\Delta t,i,j}$ with the Pearson correlation, which is adjusted with a slightly different fit model. Moreover, these two methods are also compared with inter system correlation of$K^*_{t,i,j}$.

The comparison of methods shows that comparing $K^*_{t,i,j}$ will not result in realistic decorrelation distances, even when different time resolutions are used. Direct comparison of the clearness index was not seen as a viable method, which is logical in hindsight with the variation of the spread of the clearness index. When looking at the delta clearness index, the methods both show decorrelation lengths which are comparable with the articles in which they are described. Both methods are suited for calculating the decorrelation distance.

To study the effect of wind direction on solar intermittency, the data was separated in different wind directions. Thereafter, with this separation different correlation lengths for each direction were observed. In general, the decorrelation distance of the direction with the highest correlation increases significantly with separation. Nevertheless, the direction with the highest correlation length is not always the same as the wind direction.

To put these results in a social perspective: for a significant increase of the PV systems in the Netherlands, it is advised to install multiple smaller systems instead of a few large conglomerates of PV systems, so the variation of the total PV system is far less than the sum of the variations. For forecasting the PV output in smart grid situations the decorrelation length should become a predictable function of known input parameters. It seems that the measurements of wind directions on the ground are not suited for prediction of the highest variance, although significant differences are observed while looking at different wind directions. This research confirms the need of understanding the local weather patterns for installing large scale iRES.

10. References

Bosch, J.L., Zheng, Y., Kleissl, J. (2013). Deriving cloud velocity from an array of solar radiation measurements. *Sol. Energy,* 87, 196–203.

David, M., Ramahatana, F,H. Liandrat, O. (2013). Spatial and temporal variability of PV output in an insular grid: Case of Reunion Island. *ISES Solar World Congres*, Reunion Island

Elsinga, B., & van Sark, W. (2014). Spatial power fluctuations correlations in urban rooftop photovoltaic systems. *Progress in Photovoltaics: Research and Applications*, Utrecht,

Elsinga, B., & van Sark, W. (2016). Short-term peer-to-peer solar forecasting in a network of photovoltaic systems. *Solar Energy* (submitted)"

Hoff,T.,Perez,R. (2010). Quantifying PV power output Variability: Correlation Coefficients. Technical report to the California Solar Initiative, Grant Agreement for Advanced Modeling and verification for High Penetration PV

Hoff, T.E., Perez, R. (2012). Modeling PV fleet output variability. *Sol. Energy* 86, 2177–2189.

Hinkelmann, L.M. (2013). Differences between along-wind and cross-wind solar irradiance variability on small spatial scales. *Solar Energy 88,* 192–203

Lave, M. and Kleissl, J. (2013). Cloud speed impact on solar variability scaling Application to the wavelet variability model. *Solar Energy 91*. 11-21.

Lemmens, J., (2014). PV Potentieel van zonnestroom in de gebouwde omgeving van Nederland –PV potentieelstudie. *Internal research*. PBL &DNV GL.

Mathworks (2016). *Curve fitting toolbox* (2014). Retrieved: January, 23, 2016 from www.mathworks.com/Curvefitting

Tarroja,B., Mueller, F., Samuelsen, S. (2012). Solar Power Variability and spatial diversification *international journal of energy research*. 37, 1002-1016.

Perez, R., Kvalov,S., Hoff, T. (2011). Spatial& temporal characteristics of solar radiation variability. *Proceeding of international Solar Energy World Congres*.

The Royal Netherlands Meteorological Institute, 2014. NubiScope scanning IR radiometer cloud cover of the Netherlands. Retrieved from: htttp://data.knmi.nl

Pavanello, D., Zaaiman, W., Colli, A., Heiser, J., & Smith, S. (2015). Statistical functions and relevant correlation coefficients of clearness index. *Journal of Atmospheric and Solar-Terrestrial Physics*, 130, 142-150.

Sinnott, R.W., (1984) . Virtues of the Aversine. *Sky and Telescope* , 68, 158

Wiemken, E., Beyer, H. G., Heydenreich, W., & Kiefer, K. (2001). Power characteristics of PV ensembles: Experiences from the combined power production of 100 grid connected PV systems distributed over the area of Germany. *Solar Energy,* 70(6), 513-518.

Walling, RA., Banunarayanan, V., Chahal, A, Freemen,L., Martinez,J., Miller, N., van Zandt, D., Walling, M., (2008). Analysis of wind generation impact on ERCOT ancillary services Requirements for Electric Reliability council of Texax, GE, New York

Yordanov, GH., Sastre, TO.,Midtgard O., (2013). Optimal temporal resolution for detailed studies of cloud enhanced sunlight (overirradiances). 39th IEEE PVSC Proceedings, Tampa, Florida

11. Attachments

Attachment 1. KNMI weather data of weather station De Bilt

Day	Hour	Wind dir (degrees)	Cloud covering (1/8)	Day	Hour	Wind dir (degrees)	Cloud covering (1/8)
02-23-2014	8	200	0	11-09-2014	8	140	1
	9	190	0		9	150	5
	10	190	1		10	160	7
	11	200	1		11	160	8
	12	200	4		12	160	3
	13	200	7		13	150	1
	14	210	7		14	160	3
	15	200	8		15	160	5
	16	200	8		16	150	4
02-23-2014	8	180	4	9-8-2015	8	-	0
	9	180	2		9	-	0
	10	180	2		10	-	0
	11	180	3		11	180	1
	12	180	6		12	150	5
	13	190	7		13	180	8
	14	190	8		14	170	8
	15	180	8		15	200	8
	16	180	5		16	190	7
03-30-2014	8	140	2	30-8-2015	8	70	7
	9	130	2		9	80	2
	10	160	4		10	30	7
	11	140	2		11	60	2
	12	160	4		12	30	3
	13	150	6		13	40	1
	14	130	6		14	20	4
	15	150	7		15	30	4
	16	150	7		16	30	7
27-7-2014	8	280	7	18-10-2015	8	270	8
	9	280	7		9	160	8
	10	250	7		10	230	8
	11	280	3		11	270	8
	12	280	3		12	230	8
	13	300	2		13	250	8
	14	300	1		14	250	8
	15	320	3		15	260	8
	16	330	5		16	-	8

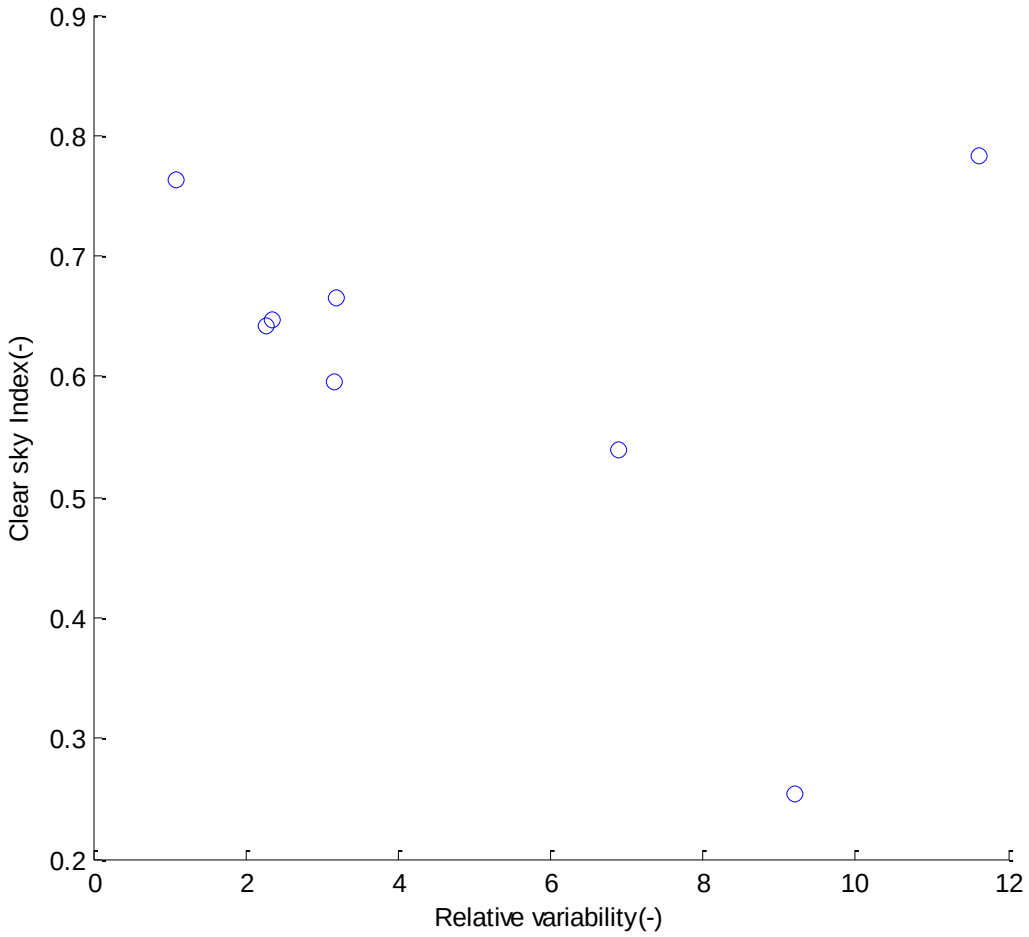


Figure 22. The relative variability versus the clear sky index of the selection of days

Attachment 2. The Clearness Index and Delta Clearness Index of multiple days

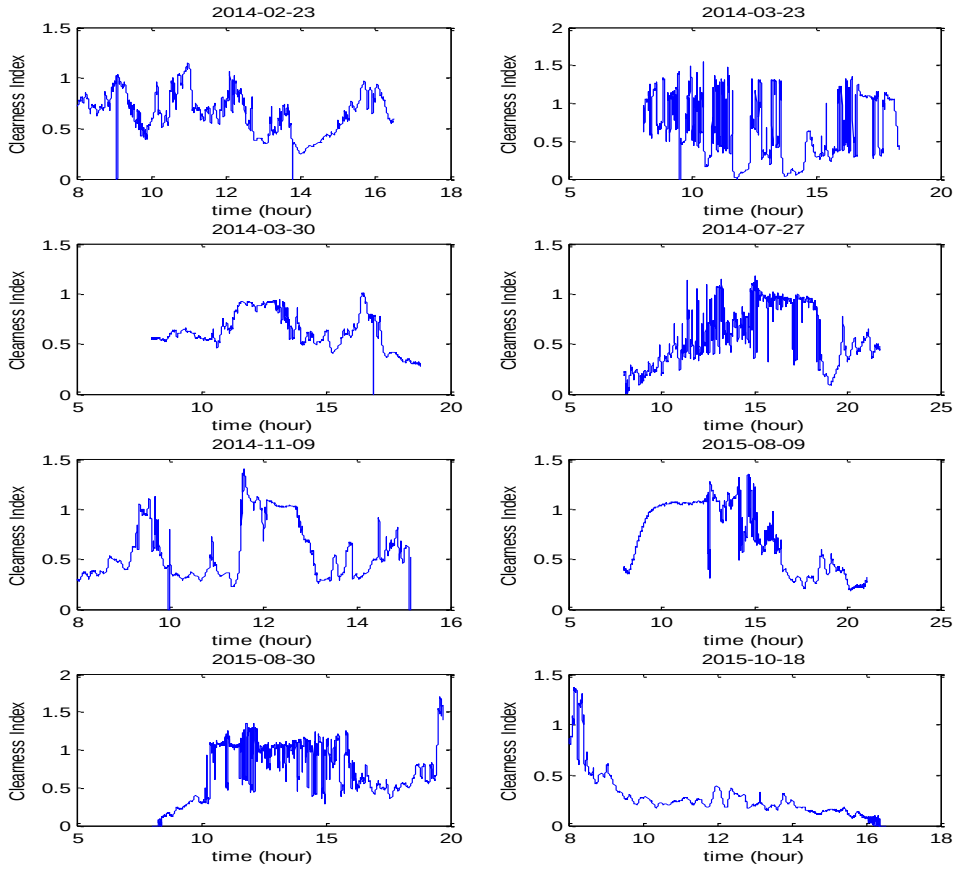


Figure 23. Clearness index of the selected days.

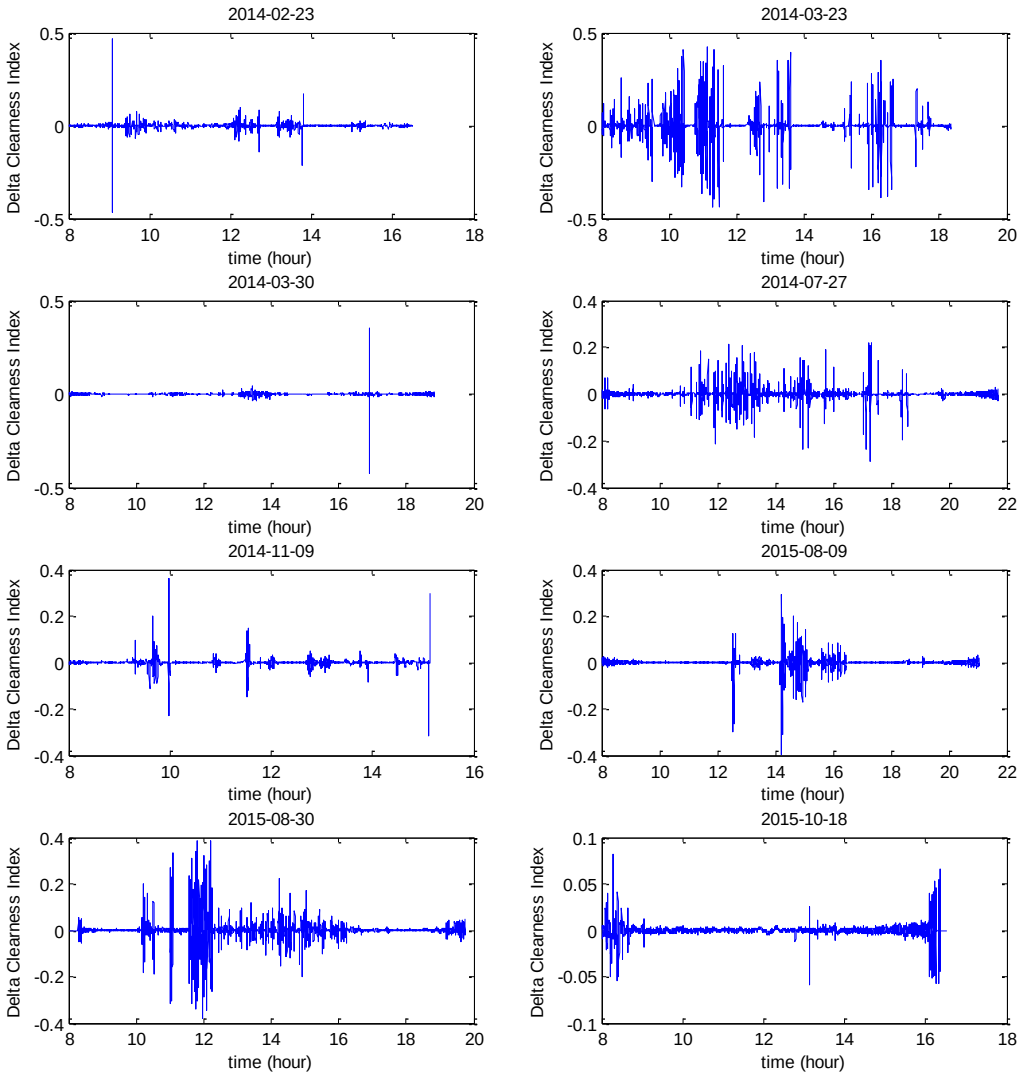


Figure 24. Delta Clearness Index of the selected days

Attachment 3. Grid operator planning.

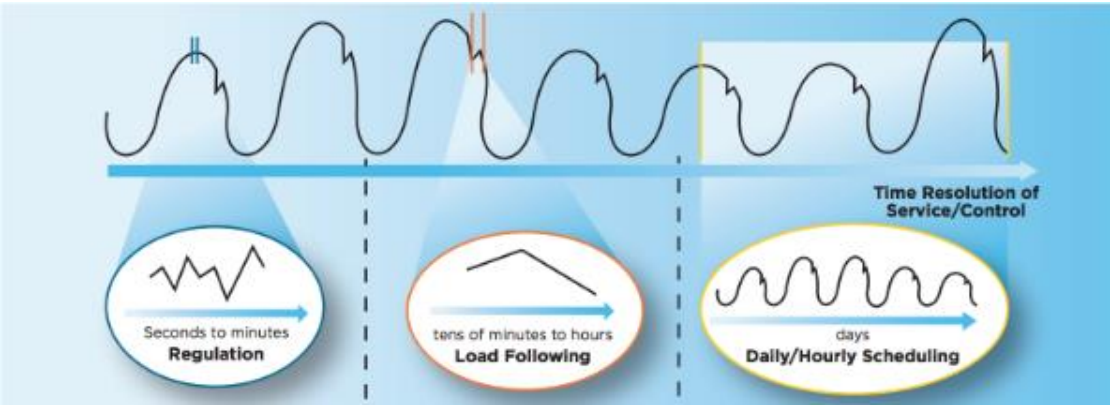
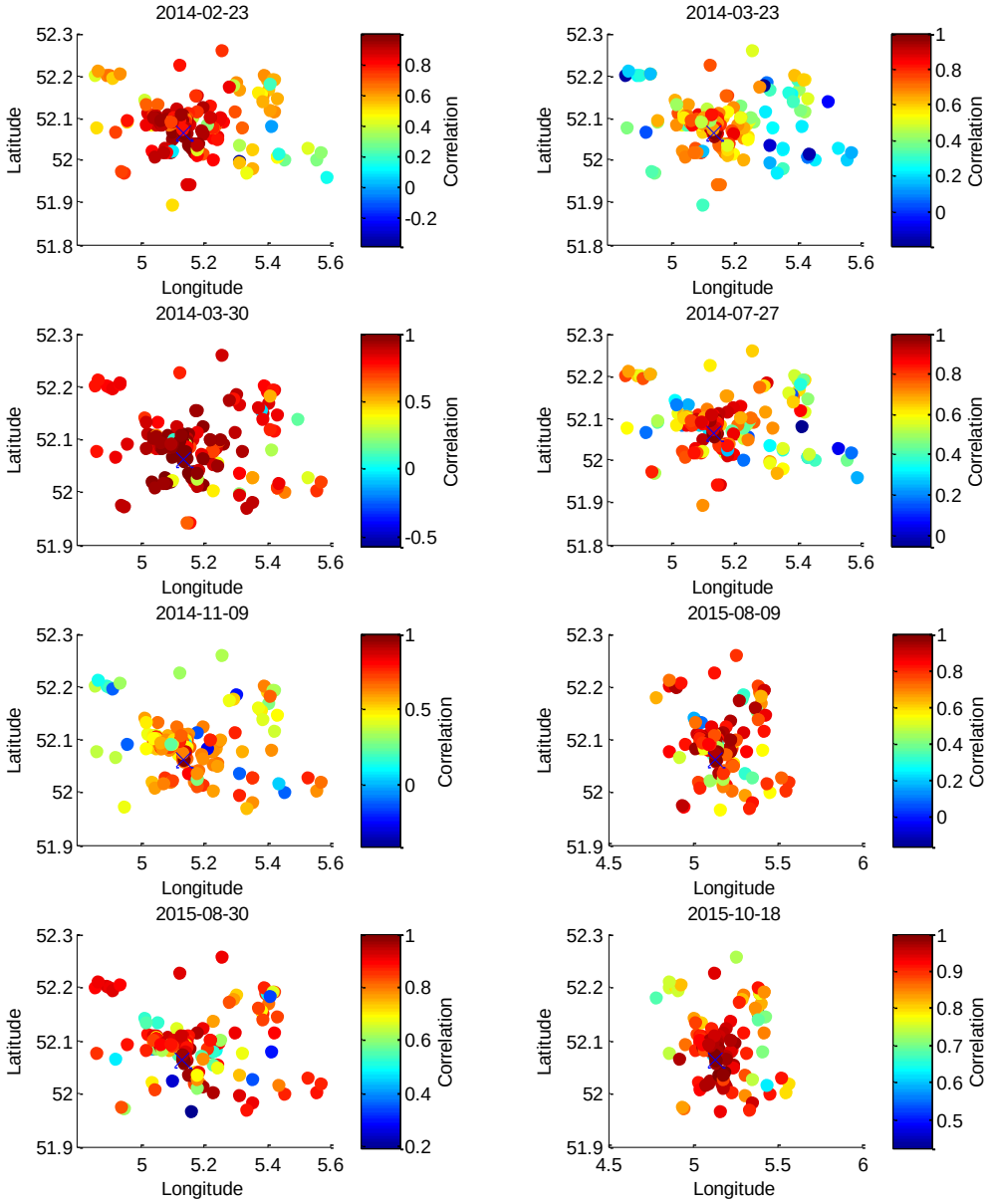


Figure 25: The grid operator plans the operation of the grid over three different time scales. The amount of electricity supply available is scheduled one day ahead, and then balanced with demand on an hour-by-hour and second-by-second basis. (Source: EERE[http://www1.eere.energy.gov/solar/pdfs/50060.pdf])

Attachment 4. The correlation of multiple days

In the figures below, the correlation of Delta clearness index with a time resolution of 20 minutes is plotted for all locations. X marks the chosen location. The days



Attachment 5. Decorrelation distances of 1 minute of all systems seperated by direction .

Table 12: Decorrelation distances (in m) of all systems combined for 1 minute with different methods for days with similar weather

	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	225	1707	2860	204	908	310	217	273
Perez	max	28279	max	42865	32251	max	max	max
DeltaElsinga	199	356	843	111	540	186	193	208
DeltaPerez	259	880	1566	185	1598	241	292	417

Table 13: Decorrelation distances (in m) of all systems combined for 1 minute with different methods for days with similar weather with averaged data.

	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	378	2689	3086	307	969	407	327	399
Perez	max	38671	max	max	max	max	max	max
DeltaElsinga	294	464	915	204	607	282	335	320
DeltaPerez	452	1052	1711	350	1791	413	487	595

Table 14: maximum decorrelation distances (in m) of all systems combined for 1 minute with different methods for days with similar weather

	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	333	2452	3909	338	1035	374	240	528
Perez	max	max	max	max	max	max	max	max
DeltaElsinga	266	546	1034	123	864	230	235	255
DeltaPerez	420	2509	3456	237	2470	277	410	1156

Table 15: maximum decorrelation distances (in m) of all systems combined for 5 seconds with different methods for days with similar weather with averaged data.

	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	469	2340	5294	410	1093	508	382	588
Perez	max	max	max	max	max	max	max	max
DeltaElsinga	430	622	1004	275	882	403	391	345
DeltaPerez	581	2265	3354	449	2443	451	611	972

Table 16: maximum decorrelation distances (in m) of the along wind direction for 1 minute with different methods for days with similar weather with averaged data.

Along wind average	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	352	1374	1538	347	924	351	351	402
Perez	max	max	max	max	max	max	max	max
DeltaElsinga	430	622	840	201	882	316	326	344
DeltaPerez	464	1596	3354	292	2443	432	428	545

Table 17:: maximum decorrelation distances (in m) of the against wind direction for 1 minute with different methods for days with similar weather with averaged data.

against wind average	02-02-2014	02-23-2014	03-30-2014	05-25-2014	11-09-2014	23-11-2014	07-26-2015	08-09-2015
Elsinga	469	1333	5294	234	905	488	312	337
Perez	max	max	max	max	max	max	max	max
DeltaElsinga	238	362	788	275	530	291	391	299
DeltaPerez	458	761	1206	449	1286	369	611	908

Attachment 6. Direction of maximum decorrelation data specified further.

Table 18: Direction with maximum wind decorrelation distances with different methods for days with similar weather with non-averaged data

	Median wind direction	5 second Elsinga	5 second Delta Elsinga	5 second Delta Perez	20 minutes Elsinga	20 minutes Delta Elsinga	20 minutes Delta Perez
02-02-2014	20	120-150	150-180	120-150	150-180	60-90	150-180
02-23-2014	180	120-150	**150-180**	90-120	90-120	30-60	0-30
03-30-2014	150	30-60	**120-150**	90-120	90-120	60-90	0-30
05-25-2014	40	120-150	150-180	150-180	150-180	150-180	**30-60**
11-09-2014	160	120-150	120-150	60-90	60-90	130-150	**150-180**
23-11-2014	160	**150-180**	120-150	**150-180**	30-60	30-60	30-60
07-26-2015	10	90-120	150-180	30-60	30-60	**0-30**	90-120
08-09-2015	3	60-90	60-90	150-180	150-180	90-120	150-180

Table 19: Direction with maximum wind decorrelation distances with different methods for days with similar weather with averaged data

	Median wind direction	5 second Elsinga	5 second Delta Elsinga	5 second Delta Perez	20 minutes Elsinga	20 minutes Delta Elsinga	20 minutes Delta Perez
02-02-2014	20	120-150	**0-30**	90-120	90-120	60-90	0-30
02-23-2014	180	90-120	**150-180**	0-30	90-120	30-60	0-30
03-30-2014	150	90-120	90-120	0-30	60-90	60-90	0-30
05-25-2014	40	150-180	90-120	90-120	150-180	150-180	**30-60**
11-09-2014	160	60-90	30-60	120-150	30-60	120-150	**150-180**
23-11-2014	160	**150-180**	**150-180**	**150-180**	30-60	30-60	30-60
07-26-2015	10	90-120	90-120	90-120	30-60	**0-30**	90-120
08-09-2015	3	150-180	60-90	90-120	30-60	30-60	150-180

Attachment 7. The decorrelation distance versus time resolutions of all days.

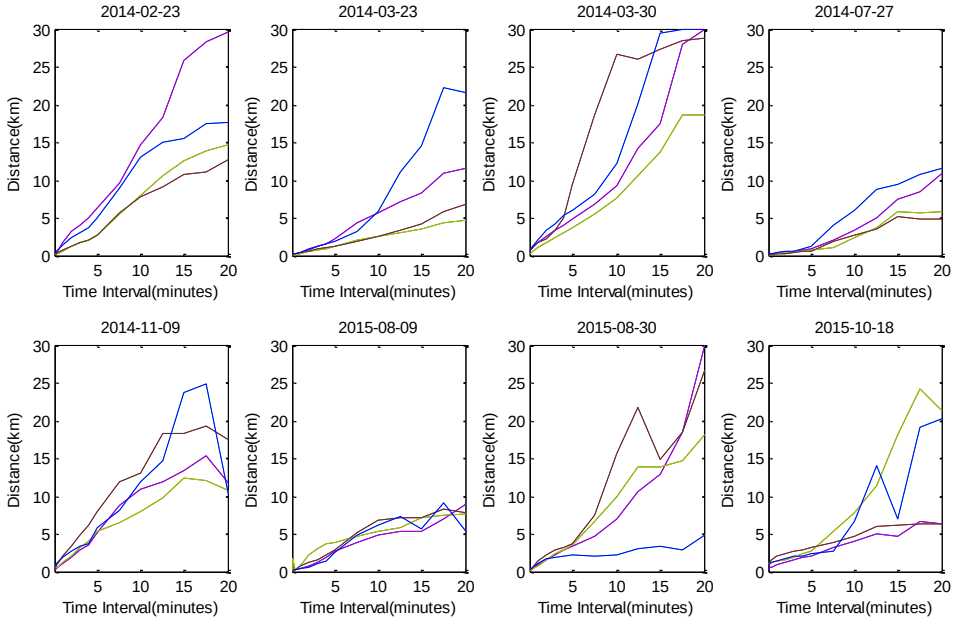


Figure 26 The decorrelation distance versus the time interval of all days for different wind directions

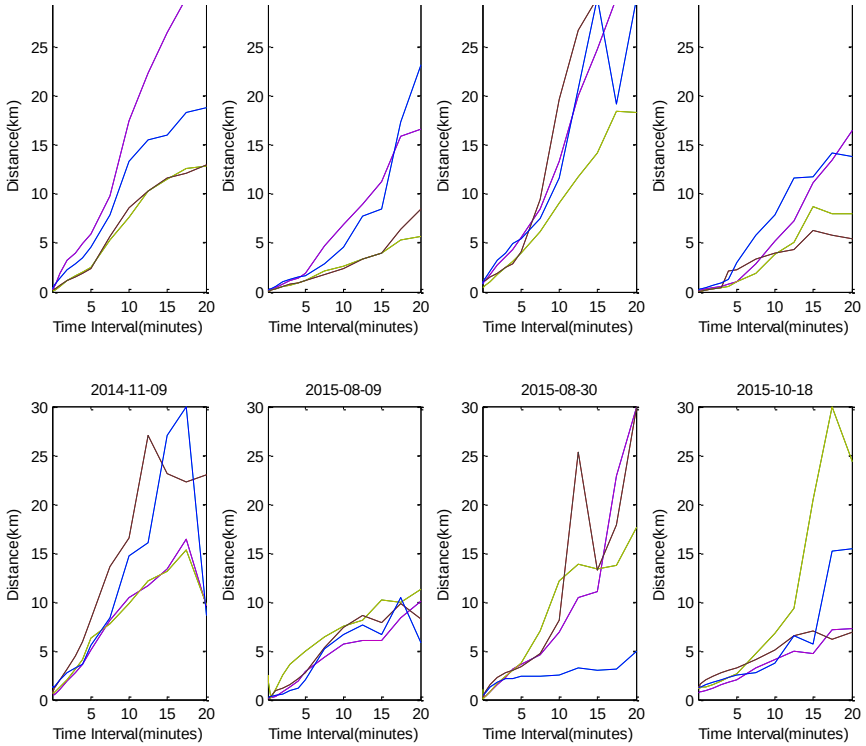


Figure 27. The decorrelation distance versus the time interval with averaged data of all days for different wind directions

Attachment 8. Extract from matlab code

```
for k=1:SizeClearnessIndex(2);
 for j=1:SizeClearnessIndex(2);
  if j==k
   EDcor(j,k)=0 ;
  else
   EDcor(j,k)=sqrt(var(DeltaClearnessIndex(:,k)-…
DeltaClearnessIndex(:,j)));
  end
 end
end
PDcor= corr(DeltaClearnessIndex(:,:),'type','Pearson');
Pcor= corr(ClearnessIndex(:,:),'type','Pearson');

for k=1:(360/AziBin)
 for j=1:size(indexrow{k},1)
  strindexrow=indexrow{k};
  strindexcolumn=indexcolumn{k};
  indexa=strindexrow(j,1);
  indexd=strindexcolumn(j,1);
  Ecorbig{k}(j)=Ecor(indexa, indexd);
  Pcorbig{k}(j)=Pcor(indexa, indexd);
  EDcorbig{k}(j)=EDcor(indexa, indexd);
  PDcorbig{k}(j)=PDcor(indexa, indexd);
  Dkmbig{k}(j)=dkm(indexa, indexd);
 end

 Dkmmult{k}=ceil(Dkmbig{k}*(1000/Dkmbin));

 for e=1:112
  EcorAveStore{e,k}(1)=999999999;
  PcorAveStore{e,k}(1)=999999999;
  EDcorAveStore{e,k}(1)=999999999;
  PDcorAveStore{e,k}(1)=999999999;
 end

 for e=1:112
  StorePlace=0;
  for i=1:size(Dkmmult{k},2)
   if Dkmmult{k}(i)==e
    StorePlace=StorePlace+1;
    EcorAveStore{e,k}(StorePlace)=Ecorbig{k}(i);
    PcorAveStore{e,k}(StorePlace)=Pcorbig{k}(i);
    EDcorAveStore{e,k}(StorePlace)=EDcorbig{k}(i);
    PDcorAveStore{e,k}(StorePlace)=PDcorbig{k}(i);
   end
  end

 end
```

```
IndexEcorIndAve=zeros(MaxDistance*(1000/Dkmbin),(360/AziBin),…
size(ClearnessIndex,2));
  for e=1:MaxDistance*(1000/Dkmbin)
   for k=1:(360/AziBin)

     if EcorAve(e,k) <1000
      IndexEcorIndAve(e,k)=1;
     end
   end
  end

Elsfittype = fittype('a * (1-exp(-xdata/b))+c',...
 'dependent',{'ydata'},'independent',{'xdata'},...
 'coefficients',{'a','b',ç});
Elsoptions = fitoptions(Elsfittype);
Elsoptions.Lower=[0 0];
Elsoptions.Upper=[1 30];
Elsoptions.StartPoint=[0.01 1];
Elsfittype = fittype('a * (1-exp(-xdata/b))+c',...
 'dependent',{'ydata'},'independent',{'xdata'},...
 'coefficients',{'a','b','c'},'options', Elsoptions);

 [curveave1, goodness, output]=fit(xdata,ydata,Elsfittype);
  ABvalueave(k,1:2)=(coeffvalues(curveave1));
  rsqvalueave(k,1)=goodness.rsquare;
 strrsq= ['Rsquare= ',num2str(rsqvalueave(k,1))] ;
```