Endogenous Soiling Rate Determination and Detection of Cleaning Events in Utility-Scale PV Plants

Åsmund Skomedal, Halvard Haug, and Erik Stensrud Marstein

***Abstract*—As the deployment rate of PV power plants continues to soar, the need for robust, scalable methods for performance analytics increases. In this paper, we demonstrate the usefulness of one approach for quantifying soiling rates in utility-scale PV power plants endogenously, i.e., directly from the production data. The temperature corrected performance ratio, normalized to a clean state, is used to derive the soiling ratio (SR). Cleaning events, caused by either rain or manual cleaning, are automatically detected by positive shifts in the running median of the SR time series. Soiling rates are then estimated by the rate of change of the SR between the cleaning events, which is determined by linear regression. The method is validated on data from three utility-scale PV power plants in the Middle East, yielding soiling rates that are in the range 0%–0.18%/day at least 50% of the time, with a median of 0.1%/day.**

***Index Terms*—Data analysis, monitoring, PV systems, regression analysis, solar power generation, time series analysis.**

I. Introduction

THE global production capacity of PV power plants installed in 2017 alone was 98 GWp [1]. This amounted to a growth of about 32% in the global capacity. Growth rates of this magnitude have been observed consistently the past decade, and this trend is expected to continue in years to come [2]. Most of the PV power built in 2017 was implemented in utility-scale PV power plants, where the need for operation and maintenance (O&M) efforts is significant [3]. The parallel development of data science and artificial intelligence has made it natural for stakeholders to work toward data enhanced monitoring solutions. Such "smart" monitoring solutions may increase the performance of the power plants as well as decrease the need for personnel and available spare parts [4].

Manuscript received December 4, 2018; revised January 15, 2019; accepted February 12, 2019. Date of publication March 14, 2019; date of current version April 19, 2019. This work was supported by the ENERGIX Project 282404 co-funded by the Research Council of Norway and Industry Partners. *(Corresponding author: Åsmund Skomedal.)*

Å. Skomedal is with the Institute for Energy Technology, Kjeller 2007, Norway (e-mail: asmunds@ife.no).

H. Haug and E. S. Marstein are with the Institute for Energy Technology, Kjeller 2007, Norway and also with the Department of Technology Systems, University of Oslo, Oslo 0315, Norway (e-mail: halvard.haug@ife.no; erik.stensrud.marstein@ife.no).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JPHOTOV.2019.2899741

One topic that has received increasing attention in the PV community in recent years is soiling. The severity of soiling varies from area to area, with some locations having a reported soiling rate of either 0.4%/day or more [5]. If the array is cleaned every 30 days, a constant soiling rate of 0.4%/day amounts to a 12% power loss at the end of the soiling period, and on average a 6% production loss over time. In areas with high soiling rates, cleaning of the PV arrays is one of the main tasks of the O&M team, and in order to optimize the cleaning schedule, quantification of the soiling loss is of high importance. Accurate measurements at a given location, however, require a dedicated soiling measurement station, which might not always be available [5]. Furthermore, the severity of the soiling may vary throughout a PV power plant because of several factors, including topography, location within the PV plant, and either exposure to roads or open areas. A soiling station only measures the soiling at the particular location where it is located. Varying soiling throughout the plant will lead to varying losses in the performance. This raises the question of whether it is possible to estimate the soiling ratio (SR) directly from the production data, i.e., *endogenously*.

In this paper, we utilize and further develop a previously suggested method for endogenous quantification of the SR and automatic detection of cleaning events. All the quantities needed to do the calculation are already measured within the standard monitoring system of the PV plant, and no extra instrumentation, such as a dedicated soiling station, is required. Another advantage of this method is that the SR is derived from the power output of the PV array. In the end, the main relevance of measuring the soiling levels is to assess the power loss. A nonuniform soiling pattern may lead to a nonlinear relationship between the short-circuit current and the power output [6]. Furthermore, the reduction in transmission due to soiling tends to be stronger at short wavelengths, which causes a nonlinear relationship between the transmission loss and the power loss [7]. It, therefore, makes more sense to derive the SR from the power output than, for instance, either a measure of the transmission or the short-circuit current.

Various endogenous methods for estimating the SR (or analogous quantities) have been suggested previously [6], [8]–[10]. In particular, this paper builds on the work of Deceglie *et al.* [8], [9], who developed the stochastic rate and recovery (SRR) method. The SRR method, in addition to estimating the soiling loss directly from PV production data, automatically detects soiling intervals, finds the soiling rate in these periods, and stochastically generates possible soiling profiles for the PV system in question. With the exception of generating new soiling profiles, this paper aims at improving the method proposed in [8] and [9], demonstrating the viability of using it in the monitoring of utility-scale PV power plants.

II. Theory and Method

The corrected performance ratio (CPR) is a measure of the performance of a PV array over a period, corrected for irradiation and temperature [11]. It is defined as

$$\mathrm{CPR} = \frac{E_{\mathrm{DC}}/P_{\mathrm{STC}}}{G_{\mathrm{POA}}/1000\ \mathrm{W/m^2}} \cdot \frac{1}{1+\gamma\left(T_{\mathrm{mod}} - 25\,^{\circ}\mathrm{C}\right)} \tag{1}$$

where E_{DC} is the dc energy output of the array over a given period, P_{STC} is the nominal power rating of the array at standard testing conditions, G_{POA} is the incident irradiation energy in the plane of the array in the given period, γ is the temperature coefficient of the array (a negative number), and T_{mod} is the module temperature in degrees Celcius. In this paper, daily values of the CPR have been used.

The CPR changes over time. In the long term, we expect it to slowly reduce because of degradation in the system components. In the short term, however, there are various elements that may cause changes in the CPR. If there is no hard shading, no low incident angle effects, no low light effects, and no faults in the system, we may assume that soiling is the most important factor for the daily changes in the CPR. In this sense, we may treat the CPR as a measure of soiling.

Of course, the assumption of no hard shading, no low angle effects, no low light conditions, and no faults does not always hold in reality. One strategy to deal with this is to treat all phenomena not related to irradiance, temperature, and soiling as noise, and to filter the data robustly for noise before they are used. Invalid data values due to loss of communication in the monitoring system must also be handled by filtering. It may be argued that, in this context, filtering and noise handling is more important than the actual metric used for estimating the SR.

One very useful concept when discussing noise levels is the signal-to-noise ratio (SNR), which for stochastic variables is defined as

$$\mathrm{SNR} = \frac{E\left(S^2\right)}{\sigma_N^2} \tag{2}$$

where $E(S^2)$ is the expected value (mean) of the square of the signal *S,* and σ_N^2 is the variance of the noise, which itself has a mean value of zero. Here we choose to define the noise as the deviation of each point from the running median. The following strategies for noise removal have been optimized with the criteria of maximizing the SNR of the CPR.

First, knowing that soiling can only decrease the CPR and never reduce the production below a certain minimum level, we keep only CPR values that fulfill the following condition:

$$\mathrm{CPR}_{\max} > \mathrm{CPR} > \mathrm{CPR}_{\min}$$

where $\mathrm{CPR}_{\max}$ and $\mathrm{CPR}_{\min}$ must be determined from the data. Furthermore, sudden shifts in the CPR time series larger than a certain threshold, followed by a recovery of the CPR, are assumed to be outliers, and subsequently removed. In this case, the threshold is the standard deviation in the noise σ_N so that the algorithm adapts to the noise level. The algorithm for doing this is as follows:

$$\begin{aligned}\text{IF } &(\mathrm{CPR}(t) - \mathrm{CPR}(t-1) > \sigma_N \text{ AND } \mathrm{CPR}(t+1) \\ &\quad - \mathrm{CPR}(t) < \sigma_N) \\ \text{OR } &(\mathrm{CPR}(t) - \mathrm{CPR}(t-1) < \sigma_N \text{ AND } \mathrm{CPR}(t+1) \\ &\quad - \mathrm{CPR}(t) > \sigma_N) \\ \text{THEN } &\text{Remove } \mathrm{CPR}(t).\end{aligned}$$

When calculating daily values of the CPR, instead of using one aggregate value for each day, the CPR has been calculated for 10-min periods in the 4 h around solar noon. The median value of CPR during these hours is chosen as the daily CPR. Strictly speaking, by doing this, we are not using the conventional definition of CPR, as defined by (1). However, there are good reasons for this choice: Calculating the CPR for each 10-min period is to correctly account for the effect of the temperature; using only the 4 h around solar noon limits the high uncertainty in the power measurements at low light conditions, low incident angle effects in morning and afternoon, as well as the fact that the SNR necessarily is lower when the signal is lower; finally, using the median, as opposed to the mean, gives robustness to outliers.

An array consists of a number of strings, and the power of each array is fed into an inverter. The energy may be calculated as the sum of all string currents I_i multiplied by the voltage over the array V_{array} and the time period Δt, i.e.,

$$E_{\mathrm{DC}} = \sum_{i=1}^{N} I_i \cdot V_{\mathrm{array}} \cdot \Delta t \tag{3}$$

where N is the number of strings. If one of the strings has a fault, it will affect E_{DC}. To avoid this, we use the median energy

$$E_{\mathrm{DC}}^{*} = N \cdot I_{\mathrm{median}} \cdot V_{\mathrm{array}} \cdot \Delta t \tag{4}$$

where I_{median} is the median current of the strings. In this paper, E_{DC}^{*} is used in (1) instead of E_{DC}. In this case, V_{array} is measured by the inverters, which are usually designed to measure current and voltage with a high internal stability, albeit with a low absolute accuracy. The uncertainty in the absolute value of V_{array}, however, does not affect the uncertainty of the SR due to the normalization that is described in the next step.

In order to estimate the SR of an array, a "clean state" of the array has to be identified. We define the "clean state" as the 95th percentile of the CPR time series, denoted by CPR_0. We may then estimate the SR by

$$\mathrm{SR}\,(t) \cong \frac{\mathrm{CPR}\,(t)}{\mathrm{CPR}_0} \tag{5}$$

where *t* is time. It must be noted, though, that the exact value of CPR_0 cannot be induced from the time series. Furthermore, because of degradation, either changes in the environment or changes in the PV system, the CPR_0 may change with time. In fact, the choice of CPR_0 is the largest source of systematic error in the estimation of the SR, contributing to an absolute uncertainty of at least 5%. A true estimate of the uncertainty cannot be determined without an accurate measure of the SR.

A cleaning detection algorithm was written following the method suggested by Deceglie in [8]. First, the running median of the daily $\mathrm{SR}(t)$ is found. Deceglie recommends a 14-day window for the running median. In the dataset used in this paper, however, it was found that eight days was a more suitable window length. The reason for this may be a higher frequency of cleaning. The optimum window length may change from dataset to dataset.

The change in the running median from day to day is denoted by $\Delta\mathrm{SR}(t)$, and any positive value of $\Delta\mathrm{SR}(t)$ above a certain threshold is identified as a cleaning event. In [8], a recommended threshold of $Q_3 + 1.5 \cdot \mathrm{IQR}$ is given, where Q_i is the ith quartile of $|\mathrm{SR}\Delta(t)|$, and $\mathrm{IQR} = Q_3 - Q_1$ is the interquartile range. This threshold adapts to the noise level. However, because of high noise levels in the data used in this paper, it was necessary to adjust the threshold to a lower value. The optimal threshold was found to be $a \cdot (Q_3 + \mathrm{IQR})$, where $a = 0.8$ in sites 1 and 3, and $a = 0.7$ in site 2. Since, for different reasons, the SNR sometimes varies with time, an attempt at adjusting the threshold to the SNR as it varies with time was also made. To do this, a running calculation of the SNR over a window of 30 days was made, and the threshold was adapted to this level. However, this was not found to increase the accuracy of the inference of cleaning events.

We define the *soiling rate* (not to be confused with the SR) as the rate of change in the SR, denoted by lower case sr. Between each cleaning event, a least squares fit was made on the daily values of the SR. The slope of the fitted line was taken to be the soiling rate, the uncertainty given by the square root of the corresponding diagonal element of the covariance matrix of the least squares fit.

The production data used in this paper comes from three utility-scale PV plants in a desert area in the Middle East, with an accumulated capacity of nearly 50 MWp. The three sites have an O&M team that does manual cleaning of the arrays at regular intervals. The pyranometers will normally be cleaned every third day although some irregularities in this routine have been observed. The module cleaning is done subarray by subarray, that is; all modules connected together in one subarray are cleaned at a time. The identification of cleaning events has been compared with a cleaning log, which has been made available from November 2017 onward.

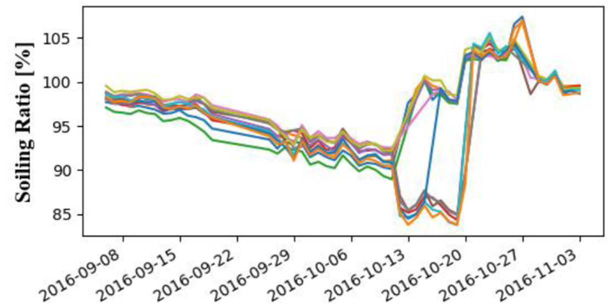


Fig. 1. SR of different arrays at site 1 during a period in the fall of 2016. As can be seen, half the plant was cleaned on October 12, one array was cleaned on October 15, and cleaning of the remainder of the plant was completed on October 19, resulting in steep positive shifts in the SR at each of these dates.

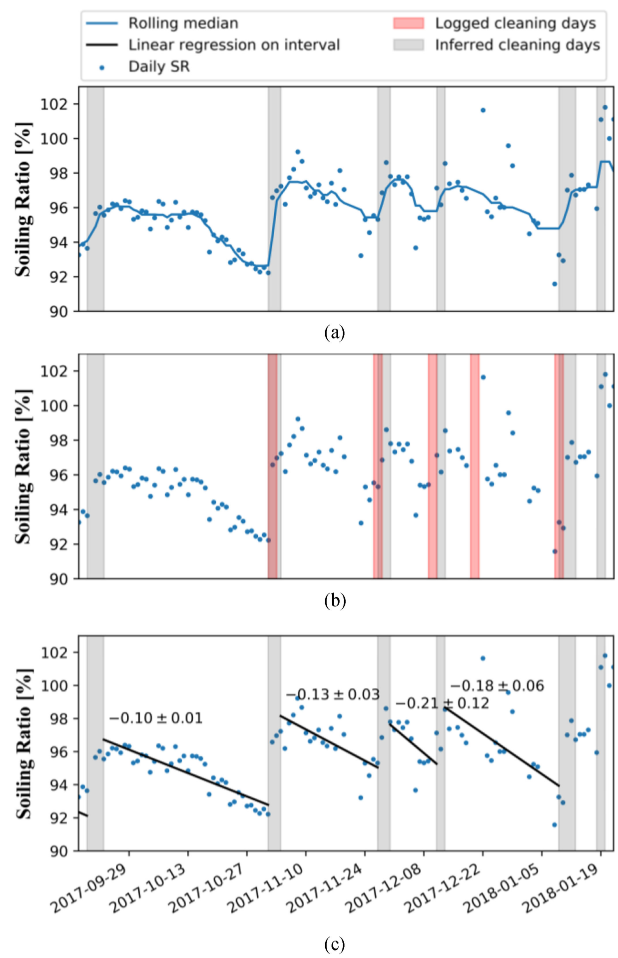


Fig. 2. Daily values of the SR with inferred cleaning events marked with vertical gray lines at times when there is a large positive shift in the rolling median. (a) Rolling median of the SR. (b) Logged cleaning events. (c) Piecewise linear regression lines on the sections between the inferred cleaning events. The numbers quantify the SR in %/day.

III. Results and Discussion

In Fig. 1, a time series of the SR at one of the sites is shown. The method developed in this paper only works if the magnitude of the change in SR due to cleaning is significantly larger than the noise in the dataset. If this is the case, it should be possible to identify periods with a gradual decrease in the SR with time, which is caused by soiling, followed by sudden positive shifts, which is caused by cleaning. This is clearly the case in Fig. 1. Furthermore, it can be seen that the different arrays have been cleaned at different times. Of course, this is an extraordinary example, where the arrays were left uncleaned for an unusually long period. For other periods in the time series, the SNR is much lower, and sometimes it is too low for the cleaning detection algorithm to work.

Fig. 2 shows the daily SR of an inverter at site 1, along with inferred cleaning events, the rolling median of the SR,

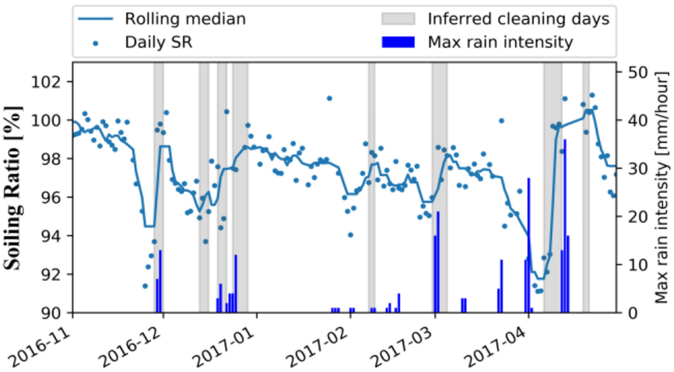
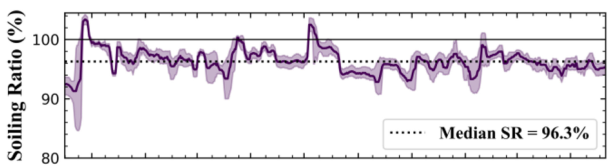
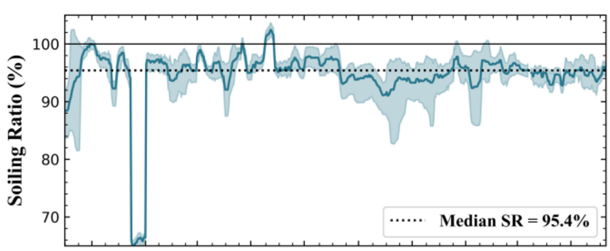


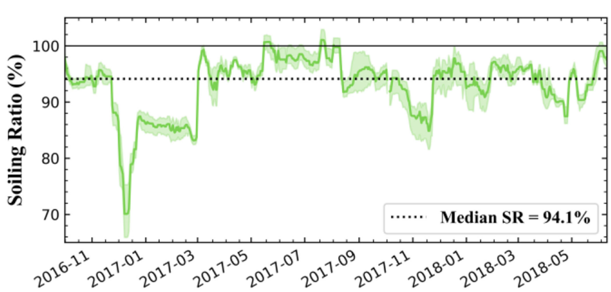
Fig. 3. Daily SR plotted together with inferred cleaning days and rain intensity.



(a)



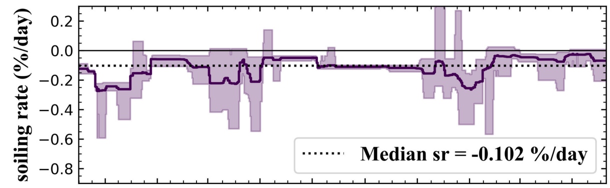
(b)



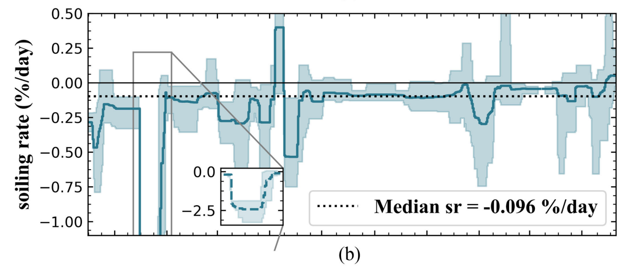
(c)

Fig. 4. SR as a function of time at (a) site 1, (b) site 2, and (c) site 3. The shaded region around the curves marks the variation between the arrays at each site, ranging from the minimum to the maximum value.

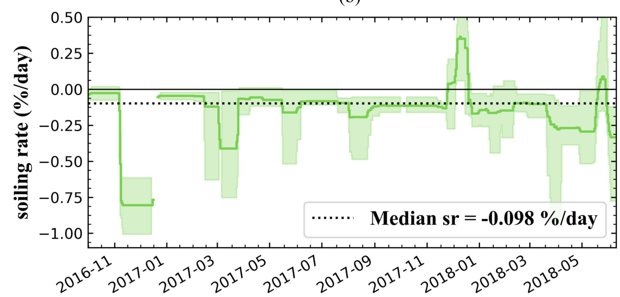
logged cleaning events, and soiling rate regression lines. This figure effectively illustrates the method used in this paper. In Fig. 2(a), it is shown how the cleaning detection algorithm marks significant positive shifts in the running median as cleaning events. As can be seen in Fig. 2(b), there are both false positives (the algorithm detects a cleaning event even though no manual cleaning has been logged) and false negatives (the algorithm does not detect a cleaning event, even though manual cleaning has been performed). A false positive may be due to either rainfall or wind, in which case the term "false" is misleading. A false negative may come about if the positive change in SR is too small relative to the noise in the data, which underlines a



(a)



(b)



(c)

Fig. 5. Soiling rates (in %/day) at (a) site 1, (b) site 2, and (c) site 3. The shaded region around the curves shows the variation in the soiling rates between the arrays at each site, ranging from the minimum to the maximum value.

weakness in the algorithm previously mentioned; namely, it is very sensitive to the noise level. Fig. 2(c) shows a linear least-squares fit to the daily SR between each inferred cleaning event. The slope of this regression line is taken as the soiling rate for that period.

Fig. 3 shows the daily SR of an inverter at site 1 in a period where there was no logging of manual cleaning, but when there was rain. It is clear that the cleaning detection algorithm is also able to detect cleaning by rain. In this case, too, there are "false positives" and "false negatives." Of course, these may be attributed to (not logged) manual cleaning and inefficient cleaning by rain respectively. In this case, it seems that a maximum rain intensity of about 5 mm/h is necessary, but not sufficient, for effectively cleaning an array.

In Fig. 4, the SR at each of the three sites is shown as a function of time. The median of the SR is 96.3%, 95.4%, and 94.1% for the three sites respectively. The shaded regions in the graphs span the minimum and the maximum SR of the arrays at each site, and as such shows the variation in SR between the arrays. Some periods have $\mathrm{SR} > 100\%$. Of course, this is a necessary result of choosing CPR_0 as the 95th percentile of $\mathrm{CPR}(t)$. Still, the times where $\mathrm{SR} > 100\%$ can be ascribed to pyranometers that have not been consistently cleaned, since an unclean pyranometer causes the calculated CPR value to be higher than it would otherwise have been. In addition, the SR going down in discrete steps, rather than gradually, can most likely be related to a pyranometer that has been left to soil together with the system

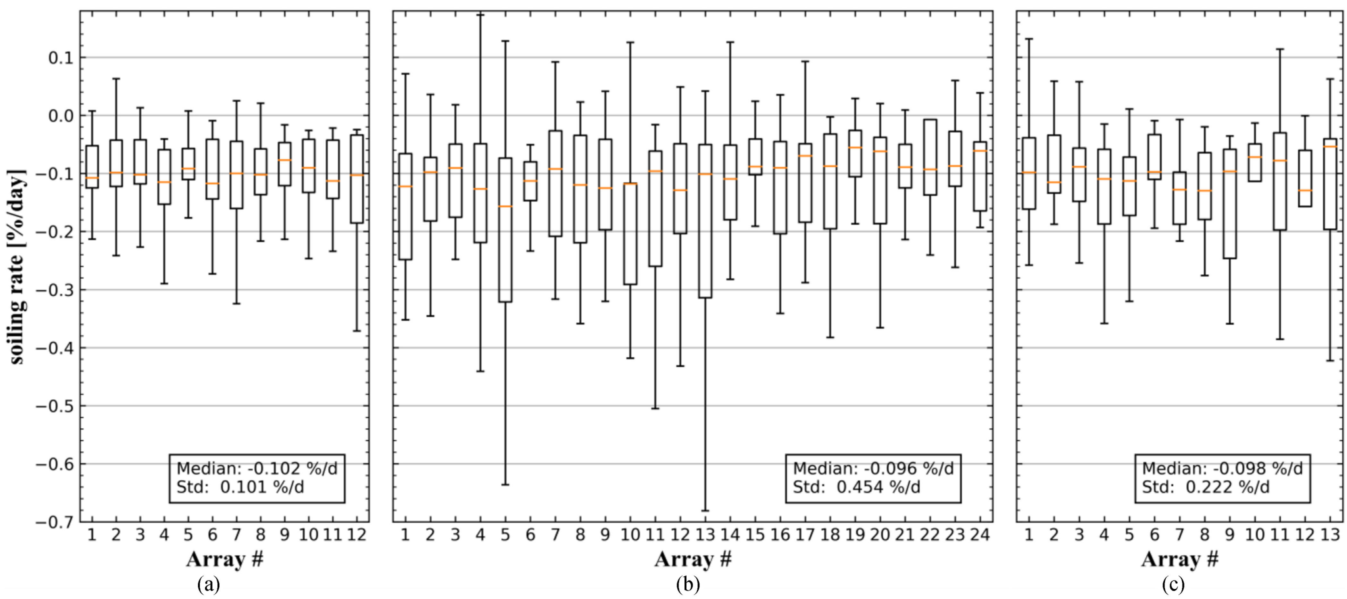


Fig. 6. Soiling rates of each array at (a) site 1, (b) site 2, and (c) site 3. The edges of the boxes mark the 25th and 75th percentile of the soiling rate of each array. The whiskers mark the minimum and maximum soiling rate of the array, and the orange horizontal line marks the median. The textboxes show the median and standard deviation of the soiling rate of each site as a whole (overall arrays for the whole period).

(leaving the CPR constant). When the pyranometer is cleaned, there is a sudden drop in the CPR, causing the stair-like shape in the SR time series. There are also periods when the SR does not reach 100%, even after cleaning. This may either be due to imperfect cleaning, or it may be due to effects that are not adequately accounted for by the algorithm. Two candidates for this effect are module degradation and wrong temperature coefficients. As mentioned above, module degradation is expected to slowly reduce the CPR over time. This has not been accounted for since we expect the magnitude of this effect to be much smaller than the uncertainty of the SR. On the other hand, there is a high uncertainty in the estimate of the temperature coefficients. These have been derived from the slope of $\mathrm{PR}(T_{\mathrm{mod}})$, where PR is the performance ratio.

The abnormally low SR in December 2016 and January 2017 at sites 2 and 3 may be ascribed to a dust storm and a subsequent lack of immediate cleaning.

The soiling rate, given in %/day, is presented in Fig. 5 for each of the three sites. The median soiling rate fluctuates between 0 and 0.3%/day through the period for all the sites, with a few notable exceptions. The high soiling rates at sites 2 and 3 in December and January 2017 may again be ascribed to a dust storm. The spread in soiling rates between the arrays within the different sites is visualized by the shaded regions, which lie between the maximum and the minimum values. The fact that the soiling rates are piecewise constant is because of the method used for extracting the soiling rates; for each soiling interval and each array, a constant soiling rate is derived based on the linear regression of the SR. Since these soiling rates are piecewise constant, the median of these is also piecewise constant. It is interesting to note that no seasonal trend can be seen in Fig. 5. The frequency of cleaning was increased in sites 1 and 2 from January 2018. In effect, the average value of the SR after January 2018 is higher, causing the SNR to be lower. This makes the cleaning detection algorithm less effective, and fewer of the actual cleaning events are detected. This, in turn, leads to a systematic error, in that the derived soiling rates are systematically lower than the true soiling rates in this period. This probably explains the relatively low soiling rates in site 1 and 2 in the period after January 2018.

The variation in the soiling rates from array to array within each site is illustrated in Fig. 6. As can be seen, the median soiling rate, represented by the horizontal line in each box, varies within and between the sites but is consistently between –0.05%/day and –0.16%/day. The edges of the boxes mark the first and third quartiles. Note that the standard deviation in the sr is lowest in site 1, twice as high in site 3, and highest in site 2. There is also a lower consistency in the position of the first and third quartiles and the minimum and maximum soiling rates in site 2. In other words, there seems to be a larger variation in the soiling rates derived for site 2 than for sites 1 and 3. This can be ascribed to various problems with the input data. As has been discussed above, the CPR is not a direct measure of soiling. Instead, it is derived from the dc power at each inverter, as well as the irradiance and module temperature. This means that it is affected by a whole range of phenomena other than soiling, e.g., module faults, disconnected temperature sensors, soiling of the pyranometers, tracker faults, and all other faults on the dc side of the inverters. This is in part controlled by filtering the data before calculating the SR. However, the challenge is to make an algorithm that filters away all unwanted conditions. Unfortunately, this is not always possible. As a consequence, if there are too many system faults (including communication loss in the monitoring system) on the dc side of the inverters, the method described in this paper might yield inconsistent results, as it does at site 2. In other words, we do not think the variation between the arrays in site 2 reflects reality. Instead, we believe that it is a consequence of the inability of the algorithm to adequately filter the data. Thus, we have a higher confidence in the results from sites 1 and 3 than from site 2.

IV. Conclusion

The endogenous method described in this paper allows the estimation of the SR on an array to array basis in a PV plant without the need for dedicated soiling stations. The algorithm is shown to be able to automatically detect cleaning events although there are both false positives and false negatives. This is because the algorithm is only able to detect cleaning events that cause changes in SR larger than the noise level of the dataset, and not all cleaning events fulfill this requirement. Furthermore, this paper demonstrates a way of estimating the soiling rates from the SR data. The accuracy of the estimate, however, is largely dependent on the noise level in the input data, as well as the accuracy of the power/energy measurements. In the case considered in this paper, sites 1 and 3 had much less noise than site 2, which leads to a larger confidence in the soiling rates derived for these sites. Judging from these data, the median soiling rate at sites 1 and 3 is 0.1%/day.

The method and data filtering algorithms used in this paper show a way of quantifying soiling rates without the need for costly instrumentation. It makes it possible to quantify the power loss because of soiling at any given time, at any given part of a PV plant. This is a general, cheap, and simple alternative to installing a dedicated soiling station and may be a valid alternative for many PV systems.

References

[1] "2018 snapshot of global photovoltaic markets," Int. Energy Agency Photovol.Power Syst., St. Ursen, Switzerland, Rep. IEA PVPS T1-33:2018, pp. 1–16, 2018.

[2] "International Technology Roadmap for Photovoltaic—Results 2017," Int. Technol. Roadmap Photovolt., Frankfurt, Germany, 2018.

[3] N. Enbar, D. Weng, and G. T. Klise, "Budgeting for solar PV plant operations and maintenance: Practices and pricing," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SAND-2016-0649R, 2015.

[4] A. Youssef, M. El-Telbany, and A. Zekry, "The role of artificial intelligence in photo-voltaic systems design and control: A review," *Renewable Sustain. Energy Rev.*, vol. 78, pp. 72–79, 2017.

[5] J. R. Caron and B. Littmann, "Direct monitoring of energy lost due to soiling on first solar modules in California," *IEEE J. Photovolt.*, vol. 3, no. 1, pp. 336–340, Jan. 2013.

[6] M. Gostein, B. Littmann, J. R. Caron, and L. Dunn, "Comparing PV power plant soiling measurements extracted from PV module irradiance and power measurements," in *Proc. Conf. Rec. IEEE Photovolt. Spec. Conf.*, 2013, pp. 3004–3009.

[7] H. Qasem, T. R. Betts, H. Müllejans, H. Albusairi, and R. Gottschalg, "Dust-induced shading on photovoltaic modules," *Prog. Photovolt.*, vol. 22, no. 2, pp. 218–226, 2014.

[8] M. G. Deceglie, M. Muller, Z. Defreitas, and S. Kurtz, "A scalable method for extracting soiling rates from PV production data," in *Proc. Rec. IEEE Photovolt. Spec. Conf.*, 2016, pp. 2061–2065.

[9] M. G. Deceglie, L. Micheli, and M. Muller, "Quantifying soiling loss directly from PV yield," *IEEE J. Photovolt.*, vol. 8, no. 2, pp. 547–551, Mar. 2018.

[10] A. Kimber, L. Mitchell, S. Nogradi, and H. Wenger, "The effect of soiling on large grid-connected photovoltaic systems in California and the southwest region of the United States," in *Proc. Rec. IEEE 4th World Conf. Photovolt. Energy Convers.*, 2007, vol. 2, pp. 2391–2395.

[11] T. Dierauf, A. Growitz, S. Kurtz, and C. Hansen, "Weather-corrected performance ratio," Nat. Renewable Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-5200-57991, 2013.

Åsmund Skomedal was born in Søgne, Norway, in 1991. He received the M.S. degree in technical physics from the Norwegian University of Science and Technology, Trondheim, Norway, in 2017. He is currently working toward the Ph.D. degree in PV systems at the Institute for Energy Technology, Kjeller, Norway, collaborating with the University of Oslo, Oslo, Norway.

Halvard Haug received the Ph.D. degree in physics from the University of Oslo, Oslo, Norway, in 2014.

He received the Ph.D. degree from the Institute for Energy Technology (IFE), Kjeller, Norway. He was a Postdoctoral Fellow and is currently a Research Scientist with the IFE, working on surface passivation, lifetime characterization, and device and system modeling of silicon solar cells. He is also an Adjunct Associate Professor with the Department of Technology Systems, University of Oslo.

Erik Stensrud Marstein received the Dr. Scient. degree in physics from the University of Oslo, Oslo, Norway, in 2003.

He has performed research in the field of solar cell technology with the Institute for Energy Technology (IFE), Kjeller, Norway. He is currently the Chief Scientist with the Department for Solar Energy, IFE, in addition to the Director of the Research Center for Sustainable Solar Cell Technology (FME SUSOLTECH) and Adjunct Professor with the Department of Technology Systems, University of Oslo. His research interests include the entire value chain of silicon-based solar cells, from feedstock, through ingot, wafer cell, and module production through to PV systems.