



Article Minimization of Economic Losses in Photovoltaic System Cleaning Schedules Based on a Novel Methodological Framework for Performance Ratio Forecast and Cost Analysis

Fabian Zuñiga-Cortes ^{1,*}, Juan D. Garcia-Racines ¹, Eduardo Caicedo-Bravo ¹, and Hernan Moncada-Vega ²

- ¹ Electrical and Electronic Engineering Department, Universidad del Valle, Calle 13 # 100-00, Cali 760032, Colombia; juan.garcia.racines@correounivalle.edu.co (J.D.G.-R.); eduardo.caicedo@correounivalle.edu.co (E.C.-B.)
- ² Maintenance Planning Department of T&D, Celsia S.A., Calle 15 # 29B-30, Autopista Cali-Yumbo, Yumbo 760502, Colombia; hmoncada@celsia.com
- * Correspondence: fabian.zuniga@correounivalle.edu.co

Abstract: The growing interest in deploying photovoltaic systems and achieving their benefits as sustainable energy supplier raises the need to seek reliable medium-term and long-term operations with optimal performance and efficient use of economic resources. Cleaning scheduling is one of the activities that can positively impact performance. This work proposes a methodological framework to define the optimal scheduling of the cleaning activities of photovoltaic systems. The framework integrates a forecast model of the performance ratio, including the environmental variables' effect. In addition, an economic analysis involving the economic losses and maintenance costs of cleaning is used. This framework is applied to a case study of a photovoltaic system located in Yumbo, Colombia. Based on the historical data on irradiance, active energy, temperature, rainfall, and wind speed, the obtained forecast model of the photovoltaic system's performance ratio in a 60-day horizon has a mean absolute percentage error lesser of than 11%. The next cleaning date is forecasted to be beyond the horizon in a 19-day range, which will decrease as time goes by. This framework was applied to historical data and compared to actual cleaning dates performed by the utility company. The results show a loss of USD 33.616 due to unnecessary, early, or late cleaning activities.

Keywords: cleaning schedule; economic losses; forecast model; performance ratio (PR); photovoltaic systems (PV)

1. Introduction

In recent years, interest in distributed generation technologies has increased, mainly in those based on renewable energy resources, with photovoltaic solar power one of the most prominent alternatives in the supply of sustainable energy [1,2]. This growing interest is due in part to current government policies, the level of maturity of technology, the reduction of costs, and the benefits that this type of technology promises to the energy sector by facilitating remote access to energy with the use of local energy potential, avoiding/postponing centralized network updates, and stimulating new business models, among others [3]. However, adequate operation, reliability, optimal performance, and efficient use of economic resources must be ensured to achieve the benefits offered by photovoltaic (PV) systems [4].

In the search for optimal performance, it is important to highlight that the accumulation of dust on the surface of the PV modules is one of the main factors limiting the photoelectricity conversion of PV energy. Dust sedimentation is influenced by various environmental factors, such as relative humidity, wind speed, rainfall, and the concentration of atmospheric particles [5], depending on the conditions of the area where PV modules are installed. For this reason, if the surface of the PV modules is not cleaned from time



Citation: Zuñiga-Cortes, F.; Garcia-Racines, J.D.; Caicedo-Bravo, E.; Moncada-Vega, H. Minimization of Economic Losses in Photovoltaic System Cleaning Schedules Based on a Novel Methodological Framework for Performance Ratio Forecast and Cost Analysis. *Energies* **2023**, *16*, 6091. https://doi.org/10.3390/en16166091

Academic Editors: Hao-Ying Lu and Yu-Pei Huang

Received: 22 July 2023 Revised: 11 August 2023 Accepted: 17 August 2023 Published: 21 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). to time, the efficiency in the generation of the PV system can decrease drastically and the corrosion of the modules can accelerate, which would also be reflected in a reduction of financial income.

Some research papers have analyzed the various effects of the accumulation of dirt on its surface on PV systems' performance. The work presented in [6] exposes three factors of the effects of dust in the generation of PV energy: the loss of efficiency, an increase in temperature, and the acceleration of corrosion. These factors were analyzed with I-V and P-V characteristic curves, indicating that as the dust density increased, the short-circuit current, the open-circuit voltage, and the system's output power were reduced. The results showed that dust with a density of 10 g/m^2 can reduce the maximum power by approximately 34%. In [7], an investigation quantified the impact of dust accumulation on power generation and the bifacial gain of a PV system. The daily rate of dust deposition and its correlation with relative humidity, wind speed, particle concentration, and ambient temperature were determined based on experiments. The results showed how a clean bifacial PV system registered a higher performance ratio (0.83) compared to the performance ratio (0.78) of a dirty bifacial PV system.

In [8], the authors developed an experiment to estimate the energy loss due to dirt in a newly installed 6.3 kWp PV system in Burydah, Saudi Arabia. The experiment's conclusions indicate that about 3% of the power was reduced after the first month of system installation. This power reduction went up to 18% within four months of installation. The work presented in [9] analyzed the performance of PV modules installed near rice farms in Thailand. Dust deposition was analyzed for periods of two weeks by microscopy and spectrometry. According to the results, it was found that the dust particles' sizes were from 10 to 20 μ m during the dry season period, which implies energy losses of around 3–4% per month. Since open-pit coal mining activities occur in South Sumatra, the effects of these activities on the formation of coal dust and fouling on PV panels installed near a mine were investigated in [10]. The last month of the six months of data collection showed that a clean panel had a 1.57% higher efficiency than a dirty panel.

Other works that analyze the effect of dust sedimentation on PV systems' performance are presented in [11,12]. The conclusions presented agree on how dust accumulation can significantly reduce the efficiency of PV systems. However, they do not offer a way to set a maintenance schedule for cleaning. In [13], a cleaning cycle optimization model was developed for a 12 MW PV plant with 46,200 panels in the western province of Jilin, China. With constant efficiency reduction, they established the maximum output power and the minimum economic loss as objectives. The authors proposed an optimization method based on adaptive dynamic programming (ADP), whose results indicate that the minimum economic loss was achieved by cleaning the system every 17 days. In [14], the performance ratio (PR) of an industrial solar roof plant located in Bangladesh was analyzed based on cleaning frequencies, and the economic benefits of increasing the frequency of panel cleaning were estimated. The study revealed that the system performance increased by up to 12% after cleaning, with a cleaning frequency of three times per month. The study conducted by [15] investigated the impact of soiling on the power generation of an 85 kW solar plant in Tulkarm, Palestine. The effect of different cleaning periods on the system's efficiency was analyzed. Over a year, the plant's production data were compared to evaluate the cleaning effectiveness in six groups of solar panels, each with a different cleaning period, including one group that was not cleaned throughout the year. The results revealed that not cleaning the panels throughout the year resulted in a power loss of 13.1%. In contrast, cleaning every six months caused an average loss of 9.1%, and cleaning every two months caused an average loss of 4.4%.

Analyses based on constant efficiency loss due to dirt accumulation in PV systems do not properly correspond to the conditions of a real installation, which is why [16] analyzed the introduction of change points in historical performance trends of a 1 MW commercial PV system in southern Spain. Fouling data were analyzed using a piecewise regression method and three change point detection algorithms. The results show that taking change points into account is important for soil modeling, particularly in studies aiming to optimize cleaning schedules. The urban areas in Dhaka, Bangladesh, have high construction activity, generating an artificial increase in dust. Therefore, to efficiently operate the PV system installed there, in [17], the soiling rates were experimentally analyzed for horizontal single-sided PV, single-sided PV with a 24° inclination, and vertical two-sided PV. These rates were used to numerically estimate the performance of each of the three configurations for different cleaning intervals. The results show that performance was maximized at cleaning cycles of 5, 6, and 28 days, respectively. Single-sided PV with a 24° inclination can produce 5.3% more revenue than vertical two-sided PV under these cleaning conditions. A different study conducted in [18] showed the application of a soiling model in five grid-connected PV plants in Spain. Based on the environmental conditions, this model considers dust accumulation and the natural cleaning produced by rain. Model outputs were compared with data collected over two years by dirt sensors installed on each grid. The results revealed a difference of 0.71% between the values obtained by the sensors and the values predicted by the model.

On the other hand, some investigations based their analyses on image processing techniques to determine the soiling level and its effect on the efficiency of PV systems. In [19], a model was proposed to optimize the cleaning cycle of a PV system in Northeast China using a dust deposition monitoring method with image recognition. In addition, the maintenance cost of two cleaning technologies was evaluated in dry and wet conditions. The results showed that the power conversion efficiency was reduced linearly while the dust deposition density increased. According to the proposed model, the optimal cleaning cycles for the PV system are approximately 10.1 and 22.8 days when its efficiency is reduced by 4.5% and 10.2%, respectively. In [20], an artificial light source was used in a laboratory environment, and the output power values of a 60 W PV module were compared among three artificially deposited dust accumulation densities. At each level of soiling, images of the module were captured, and the features were obtained using the gray-level co-occurrence matrix (GLCM). Then, the data were classified by an artificial neural network to determine the level of dust as low, medium, or high. Based on this classification and the effect on the PV module PR, criteria were established to define the cleaning cycle.

According to the various studies in which the negative effects of soiling in PV modules have been analyzed, it is clear that the effective scheduling of PV system cleaning activities can positively impact the PV system's operational and economic performance. However, most studies have established a cleaning schedule that resorts to human experience, which can cause efficiency and financial losses. On the other hand, since some research [7,13,16,17,19] assumes that the soiling rate is constant within each cleaning period, the performance profiles exhibit a sawtooth shape. This assumption may not correspond to reality by not considering the change points in soiling rates, which are related to sudden variations in environmental conditions, for example, dust storms or prolonged periods of rain.

This paper proposes a methodological framework based on a PR forecast model of PV systems to maximize energy production efficiency and reduce possible economic losses associated with soiling. Together with an economic analysis involving the economic losses and maintenance costs of cleaning, this model helps the decision maker to define the optimal schedule for the cleaning activities of PV systems in a planning period. In this model, exogenous variables are integrated, including environmental variables, to identify seasonal effects or the impact of dry and rainy periods on the PR. The proposed analysis is convenient for defining the maintenance operations and the economic performance of PV systems in the medium and long terms.

The different methodological steps were applied to a case study of a PV system located in Yumbo, Colombia. Based on the historical data on the irradiance, active energy, temperature, rainfall, and wind speed, the forecast of the performance profile of the plant in a 60-day horizon, including the next cleaning date, was defined. The results show that the forecast model could predict the behavior of the PR with a mean absolute percentage error (MAPE) of less than 11% for a 60-day horizon. Forecasts beyond this horizon increased the uncertainty of the predicted PR value; however, as time passed and new PR values entered the model, this uncertainty decreased. An analysis performed on previous cleaning activities showed a total loss of close to USD 31.616 generated by performing unnecessary, early, or late cleaning activities by the utility company.

The methodological framework can be applied step by step in other case analyses considering the particular conditions of the PV system. In addition, different environmental variables could be added to the proposed framework depending on their availability, quality, and relationship to the PR.

2. Methodological Framework for PR Forecast and Cost Analysis Cleaning Schedule

The impact of dirt on solar panels can be seen mainly in the reduction in electricity generation. Therefore, it is necessary to define an indicator that represents these changes, and which helps to identify when cleaning is necessary to recover the original generation level of the solar panels. For this reason, based on the studies by several authors [21–23], the use of the performance ratio (PR) is proposed as the main indicator of panel fouling. The PR indicator is calculated by Equation (1).

$$PR = \frac{P_{PV} * G_{STC}}{G * P_p * (1 + |\beta| * (T_{mod} - T_{pv})) * (1 - Deg)} * 100\%,$$
(1)

where P_{PV} is the power generated by the solar panels (kW); G_{STC} is the reference irradiance in standard test conditions with a value of 1 kW/m²; *G* is the irradiance measured at the site (kW/m²); P_p is the peak power of the solar panels (kW); β is the coefficient of the power variation by temperature (%/°C); T_{mod} is the mean temperature reached by the solar panels (°C); T_{pv} is the temperature measured on the solar panels (°C); Deg is the annual degradation rate of the solar panels.

For this work, a computational model that forecasts the PR value from its historical values was created and applied to a solar plant in Colombia. As inputs of the model, the historical values of the PR, irradiance, delivered active energy, temperature, wind speed, and rainfall were used. This model was expected to determine the most convenient date to carry out the next solar panel cleaning, balancing the cleaning costs and the costs of energy lost due to the decrease in the PR. As a solution, the methodological framework presented in Figure 1 is proposed and explained in the following sections.

2.1. Historical Data

The first step in the framework is to identify the variables to be used as inputs in the forecast model. The PR variable is the primary source of information for forecasting. However, depending on the data availability, quality, and PR behavior, environmental variables can be used as exogenous variables.

This paper used the data of six variables: PR, irradiance, active energy, wind speed, temperature, and rainfall. These variables were measured on-site and were selected due to their availability in the particular case analysis described in Section 3. However, any other environmental variable can be chosen and evaluated in the next step of the methodology framework.

The PR variable is the only input to the prediction model, while the other five variables are used to ensure the quality of the PR variable. Having as many historical values as possible is advisable to represent the seasonal influence throughout the year.



Figure 1. Proposed methodological framework for PR forecast and cost analysis of the cleaning schedule.

2.2. Preprocessing

In this step, the outliers of all previously defined variables are removed based on the variable limits or historical maximum and minimum values registered on-site to ensure success in creating the model.

Then, it is defined which variables are relevant for the PR forecast using the Granger [24] causality test. This is important, as these variables can be used as exogenous input in the forecast model or to impute missing PR values.

We performed a multivariate analysis of five environmental variables to impute the missing values in the PR: irradiance, active energy, wind speed, temperature, and rainfall. This analysis estimates the missing values of the PR variable based on the other variables through linear regressions. This way, it is possible to use techniques to obtain complete data to train the forecast model. However, the values have an associated error sensitive to data adequacy, so it is best to keep the number of imputed values to a minimum.

2.3. Forecast Model

The PR behavior can be analyzed as a time series using three main components: longterm trend, seasonal trend, and stationary behavior. The first two components depend on recurrent factors like wind, rainfall, or climate. The third component includes random movements of the variable and fluctuations that are difficult to explain. To train the forecast model, the PR time series obtained after the imputation was used as the only input, to which the trend and seasonality, both annual and weekly, were calculated to model the impact of the changes in the climatic seasons. Depending on the case study, seasonality may have a greater or lesser impact on the prediction of the PR.

On the other hand, cleaning activities performed on a PV system affect the PR. The PR is expected to increase once the modules have been cleaned, breaking the PR variable's normal trend due to soiling. To prevent this phenomenon from impacting the model, the cleaning activities performed were treated as "shocks" to the PR variable during training. Once the model was obtained, the daily PR forecast was made within 60 days after the last known PR value.

In this study, the forecast model accuracy was measured using four metrics:

- Mean squared error (*MSE*), described in Equation (2);
- Root mean squared error (*RMSE*), described in Equation (3);
- Mean absolute error (*MAE*), described in Equation (4);
- Mean absolute percentage error (*MAPE*), described in Equation (5).

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \left| PR_i - \widehat{PR_i} \right|$$
(2)

 $RMSE = \sqrt{MSE} \tag{3}$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} \left(PR_i - \widehat{PR_i} \right)^2 \tag{4}$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{PR_i - \widehat{PR_i}}{PR_i} \right| * 100\%$$
(5)

where *N* is the total number of observations. PR_i is the actual value of *PR* at observation *i*. $\widehat{PR_i}$ is the forecasted value of PR at observation *i*.

2.4. Next Cleaning Date Forecast

Finally, a cost analysis is performed to determine when the right time is to perform the next cleaning of the PV system. This analysis compares the economic losses due to the decrease in the PR with the cost of cleaning. The economic losses are calculated using Equation (6).

$$C_{loss} = (PR_{clean} - PR) * \frac{E_{pv}}{PR} * C_{energy},$$
(6)

where C_{loss} is the cost of energy losses due to the decrease in the *PR* due to soiling. *PR*_{clean} is the value of the *PR* measured once the previous cleaning was performed. *PR* is the historical or forecast value of the *PR*. E_{pv} is the daily active energy delivered by solar panels (kWh), C_{energy} is the cost of energy (USD/kWh).

Figure 2 shows an example of how the reduction of PR generates economic losses, and Figure 3 shows an example of the comparison between these losses and the cost of cleaning.

To determine the appropriate date to carry out the cleaning, the cost of energy lost due to soiling since the last cleaning date is calculated based on PR values previously forecasted and compared to the cost of the cleaning activity. Then the date when these two costs are equal is selected as the next cleaning date, as shown in Figure 3. However, this date varies by including the influence of the PR forecast confidence intervals. As a result, a range of dates with different probabilities of occurrence is found. This effect is reflected as a probability density function.



Figure 2. Example of PR value and economic losses due to soiling.



Figure 3. Example of comparison between economic losses due to soiling and the cost of cleaning.

In addition, as a rule, the PR annual average of PV systems must be calculated, and this must not be less than a pre-established lower threshold defined by the system operator. Therefore, in this analysis, the annual average was calculated based on historical and forecast values, and it was determined at what moment cleaning should be carried out in case this average was to cross the threshold.

3. Case Analysis

3.1. Historical Data

Following the proposed framework, the first step was to obtain the historical values of the PR and the meteorological variables. In this work, the forecast was based on data recorded daily for about three and a half years in a solar power plant in Yumbo, Colombia. The characteristics of this plant are presented in Table 1.

Location	Yumbo, Colombia (3.5518, –76.4986)		
Number of solar panels	34,560		
Rated power	11,232 kW		
Cleaning cost	USD 9382		
Ū.	0.027 USD/kWh weekday		
Energy cost	0.024 USD/kWh Saturday		
	0.023 USD/kWh Sunday and holidays		
PR lower threshold	Annual average of 63%		

Table 1. Solar power station characteristics.

The Yumbo power plant registers the daily irradiance value, active energy delivered, and PR. In addition, a nearby meteorological station records the wind speed, temperature, and rainfall.

3.2. Preprocessing

The Granger test was performed between the PR and each meteorological variable to determine their relevance in prediction. The *p*-value obtained is shown in Table 2 for different delay values in days.

Variable	1 Day	7 Days	15 Days	30 Days	60 Days
Delivered active energy	0.00%	2.90%	1.19%	0.17%	0.69%
Irradiance	0.00%	0.11%	0.15%	2.96%	2.21%
Wind speed	0.00%	0.00%	0.00%	0.03%	0.00%
Temperature	0.00%	0.00%	0.00%	0.00%	0.01%
Rainfall	2.25%	3.84%	3.92%	6.28%	6.12%

Using a significance level of 5% for the results in Table 2, the variables of delivered active energy, irradiance, wind, and temperature could be used to forecast the PR value at least up to 60 days in the future. However, the rainfall variable lost its relationship with the PR after 15 days.

The results are specific to the conditions of this case analysis and could differ depending on the PV system's geographic location and the local conditions to which this system is subjected. Therefore, it is suggested to perform the Granger test on every variable for every different case analysis.

Subsequently, outliers from each of the input variables are removed based on the percentiles and the historical lowest or highest values registered on-site, as follows:

- PR: Negative values and values above 100;
- Irradiance: Negative values, values less than the 1% quantile and greater than the 99% quantile [25];
- Active energy: Negative values, values less than the 1% quantile and greater than the 99% quantile [25];
- Wind speed: Values over 72 km/h [26];
- Temperature: Values below 14 °C and above 50 °C [26];
- Rainfall: Negative values, values less than the 1% quantile and greater than the 99% quantile [25].

Once all outliers have been removed, imputing the missing values in all variables is necessary. We used multivariate linear regression analysis to estimate the missing values across all six variables.

3.3. Forecast Model

This step adjusts a forecast model based on the preprocessed historical data. However, despite the previously obtained results, using one or all of the five meteorological variables

as exogenous inputs to the model did not yield better results than using the PR as the only input. Hence, these five variables were used only to estimate missing values.

Analyzing the three main components of the PR variable behavior is necessary to create an adequate forecast model. In this paper, we used the forecasting library Prophet v1.1 in Python to analyze the PR behavior and to forecast its value for 60 days.

First, the long-term trend is modeled using a piecewise linear function [27]. This function comprises eight linear pieces whose durations are determined based on the most significant changes in the PR time series. A different number of linear pieces were tested, and eight were selected based on their performance metrics. As time passes and new PR values are measured on-site, the number of linear pieces may increase, yielding a more accurate model. The long-term trend is shown in Figure 4.



Figure 4. Historic PR and long-term trend.

Subsequently, the seasonal trend is modeled through the Fourier series [27] adjusted to two different periods: 7 days and 365 days. These periods were selected based on the geographical conditions of the PV system of this case analysis. Colombia, a tropical country, generally has two seasons repeated yearly: the dry and rainy seasons. Moreover, Yumbo is located very close to an industrial zone that follows a well-defined weekly schedule that affects the air quality and probably the solar panels' soiling level. The seasonal trend is shown in Figure 5.



Figure 5. Historic PR and seasonal trend.

Seasonal behavior is closely related to the geographic location of the PV system. Therefore, in other case analyses, there may be, for example, a monthly seasonality or a four-month seasonality that must be represented in this step.

Finally, the stationary behavior needs to be modeled. Normally, this behavior is composed of random motion and the magnitude of the past values of the time series. Therefore, we used autocorrelation and partial autocorrelation as tools to determine the degree of the relationship that each value of the time series had with its past values.

The results are shown in Figures 6 and 7. The autocorrelation analysis shows that each stationary behavior value was related to its past seven values. Furthermore, the partial autocorrelation shows a relation with its three past values.



Figure 6. Stationary behavior autocorrelation.



Figure 7. Stationary behavior partial autocorrelation.

An extra component in the modeling had to be considered in this analysis case. The cleaning activities performed previously impacted the behavior of the PR series and modified its normal trend. Therefore, this influence must not be reflected in the forecast model.

According to the cleaning reports, this activity can take between 15 and 30 days to complete. Therefore, in this paper, a 30-day window is defined starting on each of the previous cleanup dates. The values contained in each window had a significantly reduced impact on the model training.

Based on the previously obtained results, the forecast model was trained. The 60-day PR value forecast was obtained starting on 16 August 2022, as shown in Figure 8. It can be seen that the predicted PR value follows the trend of the actual PR, whose value rose or fell between 7% and 15% daily, which impacted the size of the confidence interval. In this area, the dry and rainy seasons throughout the year affect the soiling of the modules, especially between August and November, where the effect is more noticeable. In addition, the PV system is located close to an industrial zone whose weekly activity appears to be represented too.



Figure 8. Historic PR, in-sample and out-sample forecasted PR values, and previous cleaning dates.

The model's performance shown in Table 3 shows that as the forecast horizon increased, so did the error. According to Lewis [28], a forecast model is highly accurate if its mean absolute percentage error (MAPE) is less than 10% and accurate if it is less than 20%. According to this, the forecast model had a high accuracy by obtaining a MAPE of less than 11%.

Horizon (Days)	MSE	RMSE	MAE	MAPE
7	13.30	3.65	2.73	3.98%
14	19.93	4.46	3.42	4.90%
21	31.22	5.59	4.10	5.77%
30	49.98	7.07	5.70	8.02%
37	59.14	7.69	6.57	9.34%
44	62.47	7.90	6.78	9.72%
51	58.73	7.66	6.82	9.86%
60	66.72	8.17	7.47	10.90%

Table 3. Forecast model performance for Yumbo solar power station.

3.4. Next Cleaning Date Forecast

Once the PR forecast was obtained, the energy loss cost was obtained using the values of Table 3 and compared with the cost of cleaning. Figure 9 shows the results obtained, where the probability density function represents the next cleaning date. This function is centered at the date where the economic losses and the cleaning cost are equal, and its width is defined by the confidence interval obtained from the PR forecast.



Figure 9. Comparison between economic losses due to soiling and the cost of cleaning.

According to the results, the next cleaning in the Yumbo photovoltaic system must be carried out between 13 November 2022 and 2 December 2022. The forecast cleaning date exceeds the 60-day forecast horizon proposed in the methodological framework. Therefore, its uncertainty defines a 19-day range. However, in a constantly updated model, the uncertainty of the predicted PR value decreases as time passes and new PR values enter the model. Consequently, the range of dates of the next cleaning also decreases until it becomes a single value. In this case, the cost analysis defined the next cleaning date since the PR annual average did not fall below the PR lower threshold defined by the operator, as shown in Figure 10.



Figure 10. Comparison between PR annual average and PR lower threshold.

In this case analysis, the PR variable was used as the only input to the forecast model; however, the methodological framework allowed the data on this variable to be complemented by including exogenous variables. In addition, the works mentioned in this paper's introduction directly relate solar panels' energy losses with the degree of soiling. This can increase the precision of a forecast model since it eliminates external factors other than soiling, which can affect the PR variable.

3.5. Analysis of Previous Cleaning Dates

This section evaluates the adequacy of some of the cleaning activities previously performed by the power company. The model previously trained in the methodological framework was used to forecast the PR values and then to perform the cost analysis to obtain the next cleaning date.

Figure 11 shows the suggested cleaning date when the previous cleaning date was 1 April 2019. In this case, a cleaning activity was performed on 1 July 2019. Therefore, the cost analysis used the historical PR values from 1 April 2019 to 30 June 2019, and the PR forecasted values from 1 July 2019 to find the next cleaning date. This distinction between the PR values before and after a cleaning activity was determined to avoid any influence these events could have had on the following PR historical value behavior.



Figure 11. Suggested cleaning date with a previous cleaning on 1 April 2019.

The results show a suggested cleaning date between 4 August 2019 and 8 August 2019, about one month after the actual cleaning performed by the utility company. This difference represents a loss of close to USD 4.417 due to an early cleaning.

In the case shown in Figure 12, the PR historical values from 5 June 2020 to 7 August 2020 and the PR forecasted values from 8 August 2020 were used to find the next cleaning date. The results show a suggested cleaning date between 31 January 2021 and 8 March 2021. This cleaning date exceeds the 60-day forecast horizon proposed in the methodological framework, so the analysis generated a wide range of dates.

According to the cost analysis, the utility company performed an unnecessary cleaning activity on 8 August 2020 and an early cleaning activity on 22 January 2021, generating a loss of close to USD 13.681.

Figure 13 shows a suggested cleaning date when the previous cleaning date was 22 January 2021. The PR historical values from 22 January 2021 to 12 July 2021 and the PR forecasted values from 13 July 2021 were used in the cost analysis. The results show that the utility company performed a cleaning activity about one week before the suggested cleaning date between 21 July 2021 and 23 July 2021. This situation represents a loss of close to USD 653.



Figure 12. Suggested cleaning date with a previous cleaning on 5 June 2020.



Figure 13. Suggested cleaning date with a previous cleaning on 22 January 2021.

Figure 14 shows the suggested cleaning date when the previous cleaning date was 13 July 2021. In this case, the cost analysis used the PR historical values from 13 July 2021 to 30 November 2021 and the PR forecasted values from 1 December 2021 to obtain a suggested cleaning date between 10 March 2022 and 3 April 2022. Once again, the suggested cleaning date exceeded the 60-day forecast horizon proposed, thus generating a wide range of dates.

The results show that the utility company performed an unnecessary cleaning activity on 1 December 2021 and a late cleaning activity on 11 April 2022, generating a loss of close to USD 11.375.

In the case shown in Figure 15, the cost analysis used only the PR historical values, yielding a suggested cleaning date between 31 July 2022 and 1 August 2022. The results show a late cleaning activity performed by the utility company on 18 August 2022, generating a loss of close to USD 1490.



Figure 14. Suggested cleaning date with a previous cleaning on 13 July 2021.



Figure 15. Suggested cleaning date with a previous cleaning on 11 April 2022.

To summarize, all five cases analyzed show a difference between the suggested cleaning date and the actual cleaning performed by the utility company. This difference generated a total loss of close to USD 31.616, thus highlighting the importance of the applied methodological framework and the benefits that can be achieved in the future.

4. Conclusions

Effective scheduling of PV system cleaning activities is one of the measures that can positively impact their operational and economic performance. However, this scheduling is generally carried out based on the human experience. In addition, most of the time, the rate at which dirt accumulates is assumed to be constant within each cleaning period, which neglects the impact of variations in environmental conditions.

In this paper, a methodological framework is proposed to help define the optimal scheduling of the cleaning activities of PV systems in a planning period. The proposed framework integrates a forecast model of the performance ratio, including the environmen-

tal variables' effects. In addition, an economic analysis involving the economic losses and maintenance costs of cleaning is used. The different methodological steps were applied to a case study of a PV system located in Yumbo, Colombia. Based on the historical data on the irradiance, active energy, temperature, rainfall, and wind speed, the forecast of the performance profile of the plant in a 60-day horizon, including the next cleaning date, was defined. The cost analysis associated with the predicted values of the PR was economically justified and allowed for determining the next cleaning date.

According to the results, the forecast model determined the behavior of the PR and established the date to carry out the cleaning. Defining the behavior of the PR is an effective way to set the date to clean photovoltaic systems. However, to ensure the high accuracy of a forecast model, it is necessary to ensure that the input data are equally accurate and complete.

The results show that the forecast model can predict the behavior of the PR with a mean absolute percentage error (MAPE) of less than 11% for a 60-day horizon. The cost analysis associated with the predicted values of the PR was economically justified and allowed for determining the next cleaning date.

An analysis performed on five previous cleaning dates showed a total loss of close to USD 31.616 generated by performing unnecessary, early, or late cleaning activities by the utility company.

Forecast models can be retrained to improve their fitness and increase their accuracy as time passes and new historical data are introduced. This is particularly true if the seasonality of environmental variables, such as the temperature and rainfall, change. The framework can be applied to other PV systems and specific models that work with the local environmental conditions can be obtained.

On the other hand, the forecast model can be generalized to fit the behavior of several PV systems or a cluster of systems with similar characteristics. In addition, the methodological framework could be modified to automatically adjust to the conditions of any type of PV system based on information from the environmental variables.

The quality of the input variables determines the scope or limitations of the methodological framework. The model's accuracy and forecast horizon could be extended by relating a PV system's energy losses with the degree of soiling since it eliminates external factors other than soiling that can affect the PR variable.

New variables can be added to the proposed framework. However, its suitability to forecast the PR must be evaluated using the Granger test or another method that measures the degree of causality.

Author Contributions: Conceptualization, F.Z.-C., J.D.G.-R., E.C.-B. and H.M.-V.; Methodology, F.Z.-C. and J.D.G.-R.; Software, F.Z.-C. and J.D.G.-R.; Validation, F.Z.-C., J.D.G.-R., E.C.-B. and H.M.-V.; Formal Analysis, F.Z.-C. and J.D.G.-R.; Investigation, F.Z.-C., J.D.G.-R., E.C.-B. and H.M.-V.; Resources, E.C.-B. and H.M.-V.; Data Curation, F.Z.-C. and J.D.G.-R.; Writing—Original Draft Preparation, F.Z.-C. and J.D.G.-R.; Writing—Review and Editing, F.Z.-C., J.D.G.-R., E.C.-B. and H.M.-V.; Visualization, F.Z.-C. and J.D.G.-R.; Supervision, F.Z.-C. and E.C.-B.; Project Administration, F.Z.-C., E.C.-B. and H.M.-V.; Funding Acquisition, E.C.-B. and H.M.-V. All authors have read and agreed to the published version of the manuscript.

Funding: The authors gratefully acknowledge the financial support provided by the Colombia Scientific Program within the framework called *Ecosistema Científico* (Contract No. FP44842-218-2018).

Data Availability Statement: Not applicable.

Acknowledgments: The authors gratefully acknowledge the financial support provided by the Colombia Scientific Program within the framework called *Ecosistema Científico* (Contract No. FP44842-218-2018). Acknowledgements are extended to the energy utility company CELSIA for the technical support they provided.

Conflicts of Interest: The authors declare no conflict of interest. The sponsors had no role in the design, execution, interpretation, or writing of the study.

References

- Ullah, Z.; Elkadeem, M.R.; Kotb, K.M.; Taha, I.B.M.; Wang, S. Multi-Criteria Decision-Making Model for Optimal Planning of on/off Grid Hybrid Solar, Wind, Hydro, Biomass Clean Electricity Supply. *Renew. Energy* 2021, 179, 885–910. [CrossRef]
- Ngabesong, R.; McLauchlan, L. System Planning and Modeling of a Renewable Energy Self-Sufficient Community. In Proceedings of the IEEE Green Technologies Conference, Austin, TX, USA, 4–6 April 2018; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 21–26.
- Piacentino, A.; Duic, N.; Markovska, N.; Mathiesen, B.V.; Guzović, Z.; Eveloy, V.; Lund, H. Sustainable and Cost-Efficient Energy Supply and Utilisation through Innovative Concepts and Technologies at Regional, Urban and Single-User Scales. *Energy* 2019, 182, 254–268. [CrossRef]
- 4. National Renewable Energy Laboratory (NREL). *Model of Operation-and-Maintenance Costs for Photovoltaic Systems;* National Renewable Energy Laboratory (NREL): Golden, CO, USA, 2020.
- Sivagami, P.; Jamunarani, D.; Abirami, P.; Harikrishnan, R.; Pushpavalli, M.; Geetha, V. Review on Soiling Implications and Cleaning Methodology for Photovoltaic Panels. In Proceedings of the 2021 IEEE International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems, ICSES, Chennai, India, 24–25 September 2021; pp. 1–10.
- 6. Chen, Y.; Liu, Y.; Tian, Z.; Dong, Y.; Zhou, Y.; Wang, X.; Wang, D. Experimental Study on the Effect of Dust Deposition on Photovoltaic Panels. *Energy Procedia* **2019**, *158*, 483–489. [CrossRef]
- Raina, G.; Sharma, S.; Sinha, S. Analyzing the Impact of Dust Accumulation on Power Generation and Bifacial Gain. *IEEE Trans. Ind. Appl.* 2022, 58, 6529–6535. [CrossRef]
- Kumar, A.; Alaraj, M.; Rizwan, M.; Alsaidan, I.; Jamil, M. Development of Novel Model for the Assessment of Dust Accumulation on Solar PV Modules. *IEEE J. Photovolt.* 2022, 13, 150–157. [CrossRef]
- Sakarapunthip, N.; Chenvidhya, D.; Chuangchote, S.; Chenvidhya, T. Dust Accumulation and Its Effect on PV Performance in Tropical Climate and Rice Farm Environment. In Proceedings of the IEEE 48th Photovoltaic Specialists Conference (PVSC), Online, 20–25 June 2021; pp. 1848–1854.
- Nurjanah, S.; Dewi, T. Rusdianasari Dusting and Soiling Effect on PV Panel Performance: Case Study Open-Pit Mining in South Sumatra, Indonesia. In Proceedings of the IEIT 2021: 1st International Conference on Electrical and Information Technology, Malang, Indonesia, 14–15 September 2021; pp. 251–256. [CrossRef]
- Kampira, T.; Isaacs, S.; Beem, H. Experimental Design for Measuring the Effect of Soiling on Power Production of Solar Panels in Ghana. In Proceedings of the 2022 IEEE Global Humanitarian Technology Conference (GHTC), Santa Clara, CA, USA, 8–11 September 2022; pp. 407–410. [CrossRef]
- Rai, S.; Khanna, R.; Bora, B.; Singh, S.; Jain, S.; Banerjee, C. Performance Assessment of Atmospheric Soiling Dust on SPV Technology for Composite Climate. In Proceedings of the 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering, WIECON-ECE, Bhubaneswar, India, 26–27 December 2020; pp. 143–146. [CrossRef]
- Fan, S.; Yao, X.; Cao, S.; Zhao, B. An Optimization Method Based on Adaptive Dynamic Programming for Cleaning Photovoltaic Panels. In Proceedings of the Chinese Control Conference, CCC, Shenyang, China, 27–29 July 2020; pp. 1565–1568.
- Alam, M.N.; Aziz, S.; Karim, R.; Chowdhury, S.A. Impact of Solar PV Panel Cleaning Frequency on the Performance of a Rooftop Solar PV Plant. In Proceedings of the ICDRET 2021—6th International Conference on Development in Renewable Energy Technology, Dhaka, Bangladesh, 28–30 December 2021; pp. 2021–2024. [CrossRef]
- 15. Abdallah, R.; Juaidi, A.; Abdel-Fattah, S.; Qadi, M.; Shadid, M.; Albatayneh, A.; Çamur, H.; García-Cruz, A.; Manzano-Agugliaro, F. The Effects of Soiling and Frequency of Optimal Cleaning of PV Panels in Palestine. *Energies* **2022**, *15*, 4232. [CrossRef]
- 16. Micheli, L.; Theristis, M.; Livera, A.; Stein, J.S.; Georghiou, G.E.; Muller, M.; Almonacid, F.; Fernandez, E.F. Improved PV Soiling Extraction through the Detection of Cleanings and Change Points. *IEEE J. Photovolt.* **2021**, *11*, 519–526. [CrossRef]
- Jahangir, J.B.; Al-Mahmud, M.; Sarker Shakir, M.S.; Hasan Mithhu, M.M.; Rima, T.A.; Sajjad, R.N.; Khan, M.R. Prediction of Yield, Soiling Loss, and Cleaning Cycle: A Case Study in South Asian Highly Construction-Active Urban Zone. In Proceedings of the 2020 47th IEEE Photovoltaic Specialists Conference (PVSC), Online, 15 June–21 August 2020; pp. 1371–1374.
- Redondo, M.; Platero, C.A.; Moset, A.; Rodríguez, F.; Donate, V. Soiling Modelling in Large Grid-Connected PV Plants for Cleaning Optimization. *Energies* 2023, 16, 904. [CrossRef]
- 19. Zhao, B.; Zhang, S.; Cao, S.; Zhao, Q. Cleaning Cycle Optimization and Cost Evaluation of Module Dust for Photovoltaic Power Plants in China. *Clean Technol. Environ. Policy* **2019**, *21*, 1645–1654. [CrossRef]
- Unluturk, M.; Kulaksiz, A.A.; Unluturk, A. Image Processing-Based Assessment of Dust Accumulation on Photovoltaic Modules. In Proceedings of the 2019 IEEE 1st Global Power, Energy and Communication Conference, GPECOM, Nevsehir, Turkey, 12–15 June 2019; pp. 308–311.
- 21. Ghiani, E.; Pilo, F.; Cossu, S. Evaluation of Photovoltaic Installations Performances in Sardinia. *Energy Convers. Manag.* **2013**, 76, 1134–1142. [CrossRef]
- Urrejola, E.; Antonanzas, J.; Ayala, P.; Salgado, M.; Ramírez-Sagner, G.; Cortés, C.; Pino, A.; Escobar, R. Effect of Soiling and Sunlight Exposure on the Performance Ratio of Photovoltaic Technologies in Santiago, Chile. *Energy Convers. Manag.* 2016, 114, 338–347. [CrossRef]
- 23. Tossa, A.K.; Soro, Y.M.; Thiaw, L.; Azoumah, Y.; Sicot, L.; Yamegueu, D.; Lishou, C.; Coulibaly, Y.; Razongles, G. Energy Performance of Different Silicon Photovoltaic Technologies under Hot and Harsh Climate. *Energy* **2016**, *103*, 261–270. [CrossRef]
- 24. Brockwell, P.J. Time Series Analysis. In *International Encyclopedia of Education*; Princeton University Press: Princeton, NJ, USA, 2010; pp. 474–481. ISBN 9780080448947.

- Yang, J.; Rahardja, S.; Fränti, P. Outlier Detection: How to Threshold Outlier Scores? In Proceedings of the AIIPCC '19: 2019 International Conference on Artificial Intelligence, Information Processing and Cloud Computing, Sanya, China, 19–21 December 2019. [CrossRef]
- 26. ATLAS—IDEAM. Available online: http://www.ideam.gov.co/web/tiempo-y-clima/atlas (accessed on 1 March 2023).
- 27. Taylor, S.J.; Letham, B. Forecasting at Scale. Am. Stat. 2018, 72, 37–45. [CrossRef]
- 28. Lewis, C.D. Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting; Butterworth Scientific: Oxford, UK, 1982; ISBN 1099-131X.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.