Optimization of cleaning strategies based on ANN algorithms assessing the benefit of soiling rate forecasts

Cite as: AIP Conference Proceedings **2126**, 220005 (2019); https://doi.org/10.1063/1.5117764
Published Online: 26 July 2019

Felix Terhag, Fabian Wolfertstetter, Stefan Wilbert, Tobias Hirsch, and Oliver Schaudt

Optimization of Cleaning Strategies Based on ANN Algorithms Assessing the Benefit of Soiling Rate Forecasts

Felix Terhag[1], Fabian Wolfertstetter[2,a),b)], Stefan Wilbert[2], Tobias Hirsch[3] and Oliver Schaudt[4]

[1]*University of Cologne, Mathematical Institute, Weyertal 86-90, 50931 Cologne, Germany*
[2]*German Aerospace Center (DLR), Institute of Solar Research at Plataforma Solar de Almería, Crtra Senés km 4, 04200 Tabernas, SPAIN*
[3]*DLR, Institute of Solar Research, Pfaffenwaldring 38-40, 70569 Stuttgart, Germany*
[4]*RWTH Aachen University, Department of Mathematics, Pontdriesch 10, 52062 Aachen, Germany*

[a)]Corresponding author: fabian.wolfertstetter@dlr.de
[b)]URL: http://www.dlr.de

Abstract. Soiling puts operators of solar power plants before the challenge of finding the right strategy for the cleaning of their solar fields. The trade-off between a low cleanliness and thus low revenues on the one hand and elevated cleaning costs and high field efficiency on the other hand has to be met.

In this study we address this problem using a reinforced learning algorithm. Reinforced learning is a trial and error based learning process based on a scalar reward. The algorithms improve with an increasing number of training runs, each performed on a different one-year data set. The reward being the profit of the CSP project. In order to prevent overfitting to a special case, the training data has to be sufficiently large. To increase our 5 year soiling-rate and 25 year meteorological measurement data set from CIEMAT's Plataforma Solar de Almeria (PSA). We first present a method to create artificial long term data sets based on these measurements that are representative of the sites' weather conditions. With the extended datasets we are able to train the algorithm sufficiently before testing it on the validation dataset.

The algorithm is given the daily choice to deploy up to two cleaning units in day and/or night shifts. In a second step, it is given soiling rate forecasts with different forecast horizons. At PSA our trained algorithm can increase a project's profit by 1.28 % compared to a reference constant cleaning frequency (RPI) if only the current cleanliness of the solar field is known. If it is given a one day soiling-rate forecast the profit can be increased by 1.33 %. A three day soiling-rate-forecast can increase the profit by 1.37 %. An extended forecast horizon does not seem to increase the RPI further. For sites with higher dust loads than PSA the RPI is expected to be significantly higher than at PSA.

Reinforced learning in combination with the data extension algorithm can be a useful method to increase a CSP project's profit over its lifetime.

INTRODUCTION

Because concentrating solar power (CSP) projects are developed in regions with elevated dust loads, project planners and plant operators are seeking methods to better deal with the higher soiling rates encountered at these sites. Knowledge of the soiling rate and weather conditions at a site is one step to better predict the yield of a power plant. In order to optimize the profit of the project, sophisticated mitigation strategies should be elaborated before the construction of the plant. This enables a decision on how many cleaning units to purchase and which strategy to apply in order to find the best compromise between the efficiency of the plant and spendings on cleaning procedures.

Early studies on the issue assumed constant cleaning intervals [1, 2] and optimized only the frequency of cleaning throughout the year. As soiling and weather conditions change throughout the year, this cannot be the perfect approach. For example, on summer days with high direct normal irradiance (DNI) where dumping is likely to occur, cleaning will not increase the power plant's output. In such cases, money, water and fuel can be saved instead.

In our preceding work, we investigated the performance of threshold based cleaning decisions on the profit of a parabolic trough power plant [3]. In this study, we compared the reference cleaning strategy consisting in a fix

SolarPACES 2018
AIP Conf. Proc. 2126, 220005-1–220005-10; https://doi.org/10.1063/1.5117764
Published by AIP Publishing. 978-0-7354-1866-0/$30.00

cleaning frequency with one vehicle in two daily shifts to candidate threshold based strategies. The investigation was tested on a 3 year measurement data set taken in southern Spain. It found, that condition based cleaning can increase a parabolic trough plant's profit by up to 0.8 % for the unaltered dataset on a plant layout with 50 MW_{el} nominal turbine power and 7.5 h thermal storage. The best strategy consisted in deploying all cleaning units, if the current mean field cleanliness fell below a threshold value. In certain conditions it showed to be best to introduce a second, lower cleanliness threshold value, for which external personnel is hired to perform additional cleaning action.

In reference [4] condition based cleaning is performed not only on the cleanliness of the solar field, but also on the electricity price. This study showed that cleaning costs could be lowered by 5-30 % if a threshold for the mean solar field cleanliness is employed. The study was based on artificially created soiling and weather data based on statistical methods.

In the present work we apply reinforced learning algorithms to find the best cleaning strategies and use this adaptive method to compare strategies fitted for different known parameters. We use the net profit of the CSP project over its lifetime as the parameter of comparison.

MEASUREMENT DATA AND IT'S EXTENSION

The data used in this study has been measured at CIEMAT's Plataforma Solar de Almería (PSA) in Spain. The measurements include direct normal irradiance (DNI), global horizontal irradiance (GHI), diffuse horizontal irradiance (DHI), atmospheric pressure (P), relative humidity (rH), ambient temperature (T_{amb}) and wind speed and direction. The data is available in time resolutions of down to 1 min since the year 2001 [5]. Additionally to these standard measurements, the soiling rate has been measured continuously since 2012 with the Tracking Cleanliness Sensor (TraCS) in daily time resolution [6, 7, 8].

In order to increase the learning data base for artificial neural network (ANN) based algorithms, it was necessary to find a method to increase this already broad dataset artificially. The desired method should reproduce weather characteristics present in the measurement data set. A special focus is set on maintaining temporal dependencies that are vital for cleaning strategy optimization: If cleaning is enforced one day, it has effects on the following days, because it alters the state of the solar field. So consecutive days should have a realistic dependency between each others.

Big data approaches, like reinforcement learning algorithms, tend to overfit to small datasets. That is why we need to extend our measurement data, to get a result which generalizes on unseen data. This is one reason we need different datasets for training and testing. The test dataset should be big enough to represent many eventualities and not only a few, possibly atypical years.

Extension of Meteorological Measurement Data

As a basis we applied the categorization of days into irradiance variability classes following reference [9]. There are eight available classes which are decided upon based on the temporal variation of DNI. Class 1 classifies as clear-sky hour, i.e. there is minimal to no deviation to the theoretical DNI level. Going towards the higher numbered classes, irradiance conditions tend to deviate more from the clear sky case, usually caused by clouds or important aerosol plume passages. Class 8 shows very low DNI compared to the clear-sky case and is almost completely overcast. The class separation in our study is based on 1 min datasets and the class decision is calculated in a resolution of 1 h. The intermediate classes don't only show lower irradiances than the clear sky case but also higher temporal variability.

The class distribution has been reduced to 3 classes here, using the following criteria: if more than 76 % of a day are classified as variability class 1 or 2 from the above reference, it is assigned class 1 here, if more than 50 % of a day refers to variability classes 7 or 8 it is assigned class 3 in this work. We define class 2 as all the remaining variability classes 3 to 6. The classification is performed for every measurement day in the database.

Figure 1 shows the frequency of occurrence for the 3 classes as defined here over the complete measurement interval. It can be seen, that clear sky days are more frequent in summer and have a share of roughly 1/3 during the rest of the year.

In a next step, the transition probabilities for the irradiance conditions to change from one day to the next are determined. Table 1 shows these probabilities. The first column is the class of the current day, the top line gives the class of the following day. It can be seen that for all classes the highest probability is for it to continue, i.e. a clear sky day ("class 1" in the left column) tends to be followed by a clear sky day; the same is true for the other classes. A transition from class 1 to class 3 and vice versa is the most improbable case. To keep track of the seasonal changes in

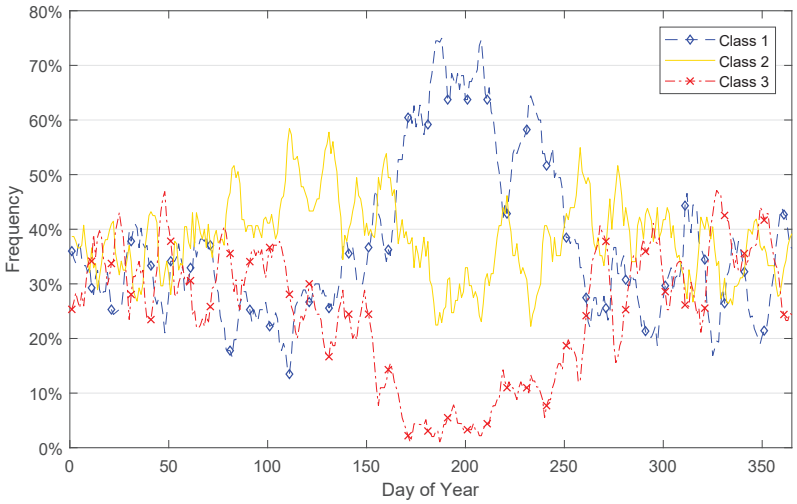

FIGURE 1. Distribution of classes over the year from all measurement days. The shown frequencies are a running average with a time window of 15 days.

frequencies, these transition probabilities are not used, but transition probabilities are calculated for each day of the year considering a time window of ±7 days. This means that transition probabilities for each day of the year are the mean value of the transitions of 14 days around the current day times the number of available measurement years. The result is a more statistically secure probability.

To create an artificial year, the class of the first day is selected following its probability of occurrence determined from measurements for the first day of the year ±7 days. The classes of the following days are drawn according to the daily transition probabilities. After the class for every day of the year is determined, a measurement day from the same class and a two week time window around the same day of the year (day of year ±7 days) is selected randomly for every day. Only complete measurement days are selected in order to produce a realistic series of weather data, including all but the soiling rate. The latter does not cover the whole measurement period such that a statistical approach is applied to the soiling rate data. This will be described below. The procedure is repeated to create 2000 training years from the measurement data and 20 years for the validation set.

TABLE 1. Probabilities of irradiance class transition from one day to the next as determined from the measurement dataset. The first column is the class of the current day, the top line gives the class of the following day.

Current day \ Following day	Class 1	Class 2	Class 3
Class 1	58 %	32 %	10 %
Class 2	31 %	45 %	24 %
Class 3	17 %	38 %	45 %

Temporal Extension of the Soiling Rate

To complete these synthetic meteorological years, soiling rate values have to be assigned to each day of the year. Therefore we use rain measurements to find a soiling rate. It is important to note, that rain is not needed elsewhere in the calculation of power plant outcome. The soiling rate and rain is divided into three bins by intensity, which are considered separately. The soiling rate is divided into high soiling ($SR < -0.02$), common soiling ($SR \in [-0, 02, 0)$) and natural cleaning events, meaning a positive soiling rate. Rain is divided by the total daily precipitation. Precipitation over 5 mm/day is considered heavy rain, 0.05-5 mm/day light rain and percipitation below 0.05 mm/day is considered no significant rain.

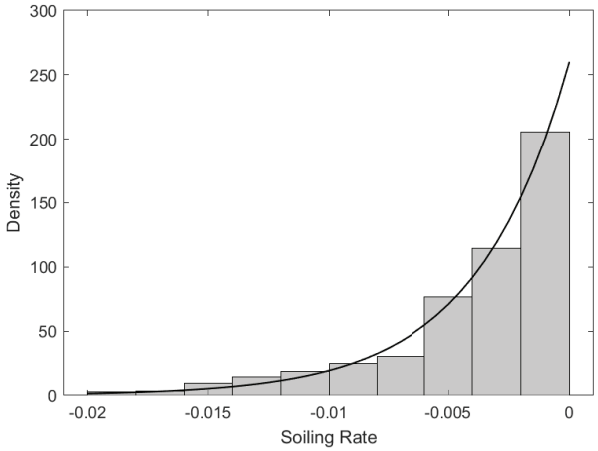

FIGURE 2. Density histogram of the daily soiling as measured at PSA since 2012. Only Soiling rate classified as *normal soiling* is shown ($SR \geq -0.02$). The black line shows the exponential fit used as the density function for the creation of artificial time series.

For creating artificial soiling rate we make several assumptions: Firstly the mean of high soiling is independent of the day of year. This is because there are not enough datapoints with high soiling to obtain a mean depending on seasonal change. Furthermore, we assume that high soiling only occurs during light rain. This is due to the fact, that most high soiling events originate from deposition processes such as red rain or deposition of Saharan dust through long range transport. Lastly, we suppose, that natural cleaning events occur if and only if there is heavy rain. Therefore we set a threshold for cleaning rain at 5 mm. This threshold is based on several studies on cleaning effects of rain on PV modules. Some studies find a small threshold of 2.5 mm [see 10] while others find 5 mm [see 11], 5-10 mm [see 12] or even 13 mm [see 13] of daily precipitation. From this range we chose a moderate threshold of 5 mm.

With these assumption we model the soiling rate as follows: For every day of year and class we calculate the probabilities for heavy and light rain. According to these probabilities we draw the type of rain (no, light, heavy) for every given class and day of year.

If heavy rain is drawn, the soiling rate is positive. For this purpose we draw from a uniform distribution between 0.7 and 1. The result is the fraction which gets cleaned during this natural cleaning event. E.g. if the cleanliness of a collector is $\xi_0 = 0.9$ and we draw a natural cleaning fraction $f = 0.8$ the cleanliness after this day is $\xi_1 = \xi_0 + (1 - \xi_0) \cdot f = 0.98$.

On a dry day the soiling is drawn from a exponential distribution with a mean corresponding to the day of year. The exponential distribution fits well to the measured data, see Figure 2. In the synthetic years, the soiling rate histogram will thus be reproduced.

On a day with light rain the soiling rate is either normal or high. Whether or not there is a high soiling event, is determined randomly. The probabilities are set, so that the expected number of days with high soiling matches the measured frequency. The soiling rate for high soiling events is determined by a exponential distribution which is offseted by -0.02. In case of a normal soiling event the procedure is the same as for dry days.

POWER PLANT LAYOUT AND CLEANING COSTS

The power plant used as an example in this study is chosen similar to the Andasol III power plant with 50 MW_{el} nominal electrical output and 7.5 h of thermal storage. The solar field has roughly 500 000 m^2 of mirror surface subject to soiling. The costs of cleaning, power plant layout and other financial and technical parameters have been listed in our previous publication [3] and will not be repeated here. For this study we only investigate the case of the Andasol type power plant in Spain. The cleaning vehicles or cleaning units are the same as in the previous study. The main parameters are listed in Table 2.

The consumption of water, fuel, manpower and depreciation of cleaning vehicles is calculated separately and added to the operation costs. The specific operation costs in table 2 does not include cleaning costs.

TABLE 2. Technical and financial data used in the study

Parameter	Value
Nominal turbine power	$49,9 MW$
Number of loops in Solar Field	156
Aperture area of solar field	510.000 m^2
Thermal storage	7.5 h
Cooling	water
Planned lifetime	25 years
DNI-yearly sum at PSA	2388 kWh/m^2/a
Equity ratio	30 %
Specific operating costs	1.8 EUR/m^2/a
Feed-in tariff	0.27 EUR/kWh
Cleaning velocity for one unit	9 loops / shift
Number of personnel per vehicle	1
Cleaning vehicle fuel consumption	6 – 8 l/loop
Cleanliness after cleaning	0.986
Demin. water consumption of cleaning unit	1 m^3/loop
Estimated lifetime of cleaning unit	15 years

REINFORCED LEARNING ALGORITHM DEVELOPMENT

To estimate the benefit of soiling rate forecasts one needs strategies, that adapt to different forecast horizons. In this work we propose to use an policy gradient algorithm, which optimizes the cleaning strategy to the inputs. This way we try to find a quasi optimal strategy for an arbitrary input, so that we can estimate the profit increase for different inputs. Therefore we have to specify which actions are possible to create a strategy and which inputs we used.

Cleaning Options and Effects on the Power Plant

It is assumed here, that up to two cleaning units are available to clean the solar field. We call the cleaning scheduling decision algorithm *agent*. The agent can choose every day from cleaning with each unit during night, day, in both or no shifts. It can thus choose from 9 possible actions. The cleaning decision has effects on the cleanliness of the solar field and the cleaning costs.

The movement of each cleaning vehicles is saved for each simulated day of the investigated year. The cleanliness of each collector of the solar field is traced individually. Every day, the soiling rate of that day lowers the cleanliness of every trough. If a cleaning vehicle cleans a trough, its reflectivity is set to 0.986 as given in Table 2. A cleaning vehicle does not clean a trough to 1 because of imperfect cleaning action and unaccessible mirror sections in a trough. If cleaning is performed during the daytime, the collector that is currently cleaned is defocused, reducing the availability of the solar field. If two vehicles clean the same parabolic trough at the same time, the availability is the same as if only one vehicle cleaned that trough.

The cleanliness of the solar field of the next day is thus influenced by the soiling rate and the cleaning decisions of the preceding days. If no cleaning is performed, the cleanliness is lowered by the soiling rate of the coming day.

The resulting solar field cleanliness is tracked and saved in a time series. The cleaning costs of that year are summed up from the daily values.

In order to calculate the power plant's revenues and profit, the software greenius [14] is used. It requires financial, technical and meteorological input parameters to be able to calculate the power plants output parameters. The cleanliness of each day is multiplied to the hourly DNI values of the same day to give the adapted DNI that is used in

greenius. The technical configuration is chosen as in the preceding paper [3]. The DNI is handed over as the cleanliness corrected DNI. The cleaning costs are added to greenius' yearly maintenance cost parameter. After one run of a cleaning agent, a greenius simulation is run and the output parameters are saved into a database. The output parameter used in this work is the profit of the project over its lifetime. The relative profit increase, RPI, is the main comparison parameter used in this study. It is defined as the ratio between the profit of the CSP project over its lifetime as achieved with the candidate cleaning strategy and the profit of the same project (using the same meteo and financial data) achieved with the reference cleaning strategy. The latter consists in two shifts of one cleaning unit on all days of the year [see 3].

Input Parameters for Decision Taking

The agent has to choose which action to take prior to each day. It can select these actions based on the known parameters of the current state. For every agent the current mean field cleanliness is known. As shown in [3] knowledge of the current mean cleanliness can significantly improve the profit compared to a constant cleaning strategy. To asses the benefit, agents are provided different forecast horizons of the soiling rate. For every forecasted day the agent receives the soiling rate in three bins and the class of the forecasted day. The soiling rate is separated the same way as in soiling rate characterization. The binned soiling rate forecast is used because it is more realistic to have knowledge of the quality of the soiling rate – whether it is normal soiling, high soiling or natural cleaning, rather than the exact value. Both the soiling rate and the class is provided in one-hot encoding, so that the input dimension is increasing by six for every forecast day.

Reinforced Learning Procedure

After we defined the inputs for the agent and the possible actions, we need to find a method that finds a way to adapt its strategy to the given inputs. In an Environment where an agent sees states – in our case the input parameters described above – and can choose between a finite set of actions, *reinforcement learning* is used to train an agent for taking actions, which maximize a reward signal. In this work we use a policy gradient method: The *REINFORCE* method described in [15, chap. 13] based on [16]. We use a two layer ANN as a parameterized policy. As an optimizer we use the *Adam* optimizer, which has shown to work well on a wide range of problems [17, 18]. The agent – in this case the parameterized strategy encoded in the ANN – gets a state as an input and outputs probabilities for each action. In this work we will limit our discussion of policy gradient methods and reinforcement learning, so that the reader will gain a conceptual understanding how policy gradient methods work and how they were applied. For a deeper discussion of the topic of reinforcement learning please refer to [15] or [19].

The synthetic meteorological years are divided into two sets. One set of 2000 years is used as training data and one set of 20 years is used as validation data. The agent starts with a randomly initialized strategy. During training the strategy is applied ten times on one year of the training set. Each day of the year the agent gets the current state. It calculates the probabilities for each action and the action taken is drawn according to these probabilities. This way the results differ slightly each time. The project's profit of each year are obtained via greenius as stated above. The total profits are saved and normalized by subtracting the mean and dividing by the standard deviation. The normalized total dividends are the reward for the actions in the corresponding year. The strategy is updated in accordance with the update rules of the REINFORCE Algorithm, so that the actions taken with negative rewards are slightly discouraged and the actions which led to positive rewards are encouraged in the future. After this update the strategy is applied to the next year in the training set. Every 15 updates the strategy is tested on the independent validation set. For this purpose the strategy is applied to every year of the validation set. This time the action taken is not drawn randomly, but the action with the highest probability is chosen. So that the strategy is transformed into a deterministic strategy in the validation case. This is done to keep track of the training progress on a comparable set. For this purpose the mean RPI of the 20 validation years is used. The training ends when the agent has not improved on the latest best result for 20 consecutive validation steps. The best agent on the validation set is picked as the final strategy. For better robustness, every setting is run twice and the better agent is used as the final result. This way we reduce the chance that the final result is trapped in a suboptimal local maximum.

The layout of the feedforward ANN is not changed between the different trials. The first hidden layer consists of 35 and the second hidden layer of 30 tanh-neurons. The layout is kept constant between the different runs to maintain comparability. The ideal way would be to perform a hyperparameter optimization for every setting, but this is not feasible due to limited computation time. This approach might be favorable for no or small forecast horizons,

because it is possible that the layout is not complex enough to deal with tasks of higher complexity, posed by higher-dimensional inputs. We tested random variations of layouts that showed that the results are fairly robust towards different architectures. If this algorithm would be used to optimize one specific setting of CSP-plant it would be recommended to also try and optimize the hyperparameter of the reinforcement learning method.

Cleaning Agent Behavior

To investigate the behavior of the cleaning agents, we look into the cleaning actions on one year of the validation set. In figure 3 we compare agents trained with a forecast horizon of three days and one day to the constant reference strategy. Both agents learned to clean in no, two or four shifts a day. If they use two shifts, they deploy both cleaning units in a night shift. That way it avoids the need to defocus the collectors at daytime during the cleaning action, that would lower the availability of the solar field and thus power output as described in [3]. Because two cleaning units can clean one collector at once, no further defocussing is necessary if two cleaning units are deployed, but twice as much is cleaned during daytime. This is why it makes sense to deploy pairs of cleaning units at daytime, if any.

Comparing the mean field cleanliness of the agents, to the constant reference strategy, it can be seen that if the cleanliness drops low – e.g. after a heavy soiling event – the trained agents are faster to recover to a moderate mean cleanliness. This is done by cleaning four shifts in one day for several days. On the other hand, trained agents save cleaning costs, if the mean cleanliness is already high, although the constant strategy decides to preserve at a higher mean cleanliness for these episodes (see for example day of year 252-268 in figure 3). Those observations hold true for the agent trained on one respectively three forecast days alike.

Comparing the differences between the two trained agents displayed, it can be observed that both tend to stop cleaning and thus save cleaning costs, in expectation of a natural cleaning event. The agent with the largest forecast horizon can do so earlier. This is best visible at the natural cleaning events of day of year 98, 169 and 268 in figure 3. While the agent with a forecast horizon of one day can only stop cleaning at the day of the natural cleaning event, the agent with a forecast of three days, stops cleaning two days earlier. That means it uses its full forecast horizon of three days to save cleaning costs if possible.

RESULTS AND DISCUSSION

For assessing the benefit of soiling rate forecasts, an agent is trained for five different settings. One setting where only the mean cleanliness is known and other settings with soiling rate forecasts of 1,2,3 and 6 days. The performance indicator is the relative profit increase (RPI).

The training progress for the five different settings is displayed in figure 4. After the random initialization the agents start off with a fairly bad strategy. During the first 20 validation steps the RPI of the strategy increases rapidly, from a mean RPI of $-3.7\,\%$ at validation step zero to $1.1\,\%$ at step 20. After that the training progress slows down in the following 10 validation steps, before it levels out for all settings between validation step 30 and 40. The mean RPI is $1.3\,\%$ at validation step 30. At this step each setting has been evaluated on 4500 years and thus taken more than 1.6 million decisions.[1]

In the detailed view of the training courses in figure 4 right, it is visible that the agent without a soiling rate forecast remains on a lower level. Also the agent with only one forecast day settles on a lower level, whereas the best results of the other three agents seem to be on a very similar level. This is also supported by the numerical results in table 3. While the agent with no forecast reaches an RPI of $1.28\,\%$ and one forecast day reaches $1.33\,\%$, the remaining agents all reach an RPI of of $1.36\,\%$ or $1.37\,\%$.

An agent with no soiling rate forecast, can still react to the mean field cleanliness. So that it can recover when the cleanliness drops low and save cleaning cost, when the cleanliness is already high. This is why, although it has no soiling rate forecast, this setting finds a strategy that achieves a good improvement compared to the constant cleaning strategy.

The soiling rate forecast allows the agents to react to future events. This is why agents with access to soiling rate forecast further improve the RPI. But after a forecast horizon of two days, the agents do not improve significantly with further forecasted days. This is likely due to the fact that if there is a special event coming in three days or more the current field cleanliness is more crucial than the event coming in a few days. For example, we have seen, that the

[1] Every validation step consists of 15 training updates and for each update one year is evaluated 10 times, with slightly different actions determined by the stochastic strategy. In every year the agent has to choose one action for all of the 365 days.

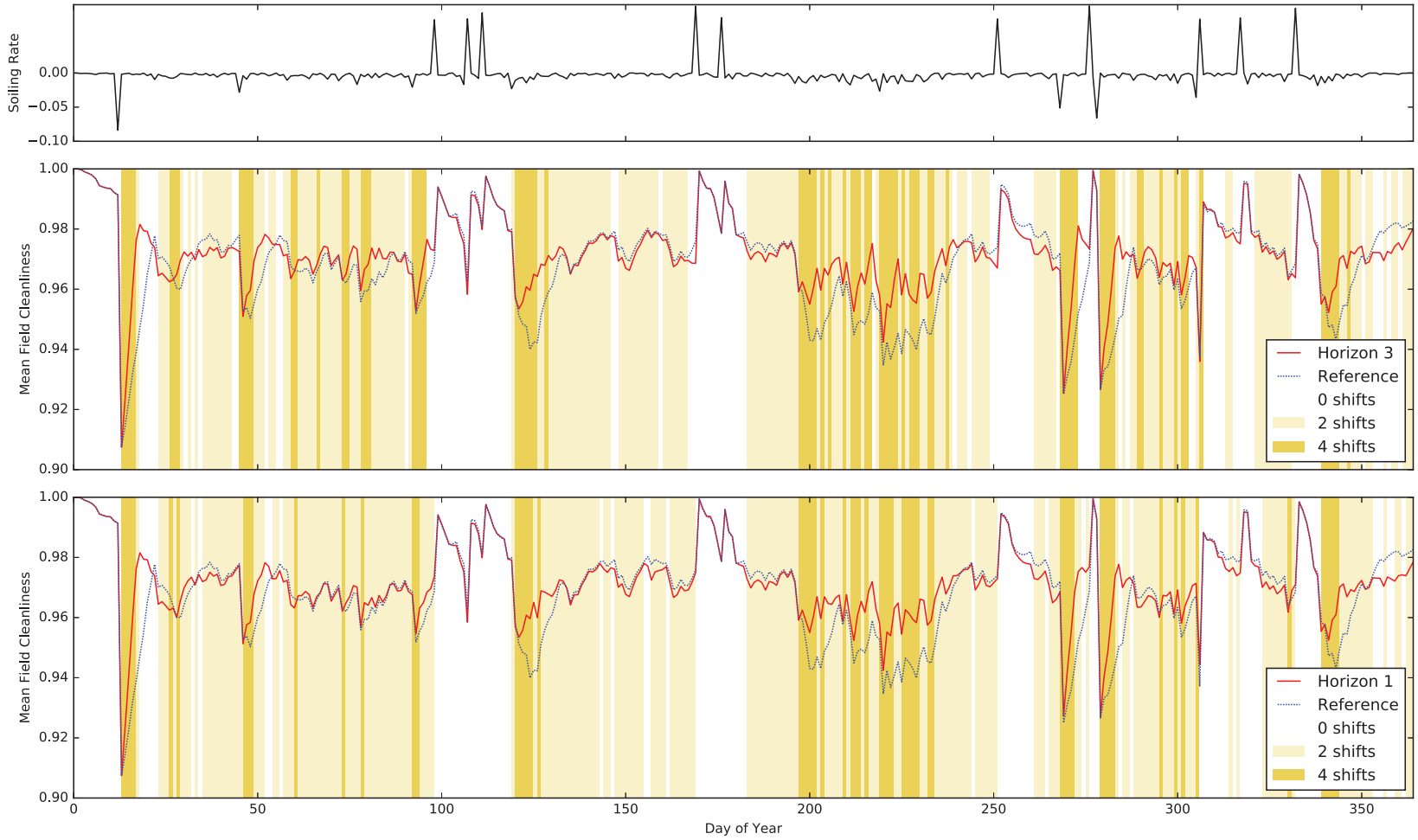

FIGURE 3. Comparison of the cleaning agent behavior on one example year in the validation set. On the top the soiling rate of this specific validation year is displayed. The soiling rate of natural cleaning events is dependent on the current cleanliness as described in a former chapter. The positive soiling rate for a cleanliness of 0.9 is shown. Below is the cleaning agent's behavior of the fully trained agents with a forecast horizon of three days (middle) and one day (bottom). For this the mean field cleanliness is compared in each case to the constant cleaning reference. The constant reference cleans with one cleaning unit in two shifts per day. The colors in the background indicate the specific action the agent took at that day of year.

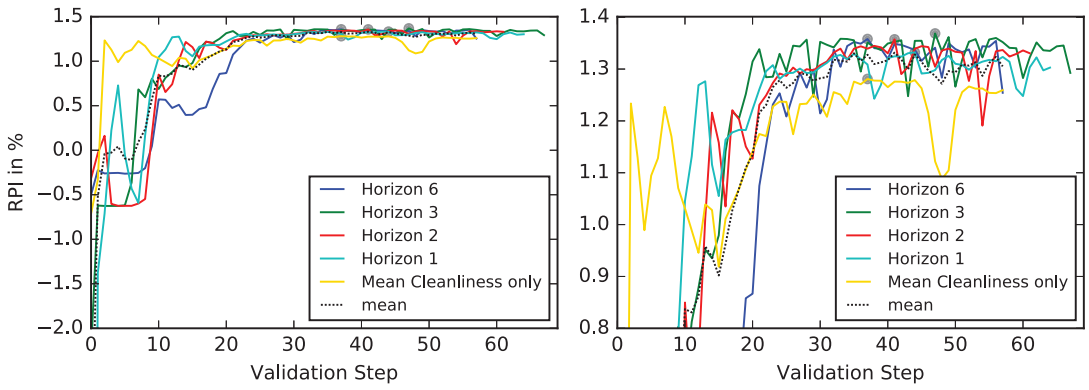

FIGURE 4. Training progress of the different settings. The colored curves show the relative profit increase for the five different forecast horizons studied. While the left side should provide an overview of the overall training course, the right side provides a zoomed in, more detailed, view of the results. The black, dotted line indicates the mean of the five different settings.

TABLE 3. Relative profit increase for different forecast horizons. All values are the mean value over the Validation set of 20 years. Forecast horizon ø means that the agent gets no soiling rate forecast input.

Forecast Horizon in days	**RPI in [%]**	**water used for cleaning in [m^3]**	**average cleaning shifts per day**	**annual net el output in [MWh_{el}]**
ø	1.28	6398	1.95	157851
1	1.33	6108	1.86	157828
2	1.36	5987	1.82	157821
3	1.37	6003	1.83	157831
6	1.36	5846	1.78	157794

agent with a forecast horizon of three utilizes all three days to save cleaning costs, but it performs only marginally better than the agent with a forecast horizon of two days. Even though the agent saves cleaning costs by not cleaning it looses solar revenue because of a slightly less clean field during those days. This is one reason why a forecast horizon of six days does not improve the RPI. In fact the RPI is slightly lower than the RPI for a horizon of three, even though it has the same and additional information. In tests performed beyond the scope of this paper, we estimated the error rate as 0.02 % RPI. So a difference of 0.01 % is well within this uncertainty.

It is worth noting that the problem gets more challenging with every further forecast day which results in longer calculation times. Whereas the agent with a horizon of three days has an input dimension of 19, an agent with six forecast days has an input dimension of 37.[2] The training process of the agent with six forecast days takes the longest to reach an acceptable level. Between validation step 10 and 20 it is well below all other agents, see figure 4.

CONCLUSION

The application of reinforced learning algorithms in cleaning strategy optimization can increase the profit of a CSP project. The issue of a database typically too small for neural network based algorithms can be overcome with the creation of synthetic datasets for training that reproduce the site characteristics from measurements. We have seen, that the agents learn reasonable behavior on those synthetic datasets.

A profit increase of 1.28 % can be achieved, if only the current state of the solar field cleanliness is known to the cleaning agent during training and application. If a soiling rate forecast with low resolution (three bins) of two or more days is available, the profit can be increased by up to 1.37 %. This result obtained with reinforced learning algorithms is more than 50 % better than the profit increase with simple field clanliness based strategies.

[2] As described in a former chapter, with every forcast day the input dimension increases by 6, because of the forecasted class and the soiling rate, both in three bins and one-hot encoding.

This shows that investigations in the direction of soiling rate modeling and forecasting have their justification: Cleaning decisions can be made on the basis of upcoming DNI and soiling rate data. Similar to the results in [3] it is expected that the improvement is higher at sites with elevated dust loads, increased soiling rate forecasting resolution or different plant layouts. The improvement is also expected to be higher for CSP plants with less thermal storage.

The presented approach can be directly used to test the influence of other parameters on the cleaning strategy. For example the influence of higher water costs, temporally variable costs of some kind or adaption to higher soiling or different plant layouts. The presented algorithm can find a cleaning policy that adapts to any of those given settings.

ACKNOWLEDGEMENT

This project has received funding from the European Unions Horizon 2020 research and innovation programme under grant agreement nr. 792103, project SOLWATT.

REFERENCES

[1] A. Heimsath, M. Heck, G. Morin, W. Kiewitt, and W. Platzer, "Soiling of aluminum and glass mirrors under different climatic conditions and techno-economic optimization of cleaning intervals," 2010.

[2] K. Kattke and L. Vant-Hull, "Optimum target reflectivity for heliostat washing," 11. - 14. SAeptember 2012 2012.

[3] F. Wolfertstetter, S. Wilbert, J. Dersch, S. Dieckmann, R. Pitz-Paal, and A. Ghennioui, "Integration of soiling-rate measurements and cleaning strategies in yield analysis of parabolic trough plants," *Journal of Solar Energy engineering*, vol. 140, no. 4, p. 041008, 2018.

[4] H. Truong Ba, M. E. Cholette, R. Wang, P. Borghesani, L. Ma, and T. A. Steinberg, "Optimal condition-based cleaning of solar power collectors," *Solar Energy*, vol. 157, pp. 762–777, 2017.

[5] S. Wilbert, B. Reinhardt, J. DeVore, M. Roeger, R. Pitz-Paal, C. Gueymard, and R. Buras, "Measurement of solar radiance profiles with the sun and aureole measurement system," *Journal of Solar Energy Engineering*, vol. 135, pp. 041002–041002, 2013.

[6] F. Wolfertstetter, K. Pottler, A. Alami Merrouni, A. Mezrhab, and R. Pitz-Paal, "Novel method for automatic real-time monitoring of mirror soiling rates," 2012.

[7] F. Wolfertstetter, K. Pottler, N. Geuder, R. Affolter, A. A. Merrouni, A. Mezrhab, and R. Pitz-Paal, "Monitoring of mirror and sensor soiling with tracs for improved quality of ground based irradiance measurements," *Energy Procedia*, vol. 49, pp. 2422–2432, 2014.

[8] F. Wolfertstetter, *Auswirkungen von Verschmutzung auf konzentrierende solarthermische Kraftwerke*. Phd, 2016.

[9] M. Schroedter-Homscheidt, M. Kosmale, S. Jung, and J. Kleissl, "Classifying ground-measured 1 minute temporal variability within hourly intervals for direct normal irradiances," *Meteorologische Zeitschrift*, 2018.

[10] F. A. Mejia and J. Kleissl, "Soiling losses for solar photovoltaic systems in california," *Solar Energy*, vol. 95, pp. 357–363, 2013.

[11] R. Hammond, D. Srinivasan, A. Harris, K. Whitfield, and J. Wohlgemuth, "Effects of soiling on pv module and radiometer performance," in *Photovoltaic Specialists Conference, 1997., Conference Record of the Twenty-Sixth IEEE*, pp. 1121–1124, IEEE, 1997.

[12] 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 *Photovoltaic Energy Conversion, Conference Record of the 2006 IEEE 4th World Conference on*, vol. 2, pp. 2391–2395, IEEE, 2006.

[13] M. Naeem and G. Tamizhmani, "Climatological relevance to the soiling loss of photovoltaic modules," in *Smart Grid (SASG), 2015 Saudi Arabia*, pp. 1–5, IEEE, 2015.

[14] J. Dersch, K. Hennecke, and V. Quaschning, "Free greenius new options and availability," 2012.

[15] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. The MIT Press, second ed., 2018.

[16] R. J. Williams, *Simple Statistical Gradient-Following Algorithms for Connectionist Reinforcement Learning*, pp. 5–32. Boston, MA: Springer US, 1992.

[17] S. Ruder, "An overview of gradient descent optimization algorithms," *CoRR*, vol. abs/1609.04747, 2016.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *CoRR*, vol. abs/1412.6980, 2014.

[19] C. Szepesvri, *Algorithms for Reinforcement Learning*, vol. 4. Morgan & Claypool Publishers, 01 2010.