Predicting solar radiation using a parametric cloud model

Kristian Skeie[1,*] and *Arild* Gustavsen[1]

[1] University of Science and Technology NTNU, Architecture and technology, Trondheim, Norway

Abstract. In this paper, we evaluate a method to calculate hourly global solar radiation and improve the calculation of diffuse and vertical surface radiation on building facades by accounting for ground conditions based on publicly available data of cloud coverage, temperature and precipitation from a forecast service covering the Nordic countries. The detailed weather forecasts produced by MET Norway provide hourly forecasts for the next 60 hours, and 6-hour predictions for the next week. To calculate solar radiation on cloudy days the clear and cloudy sky MAC model developed by Davies and Mckay (1982) is used. Instead of basing the prediction on ground observations as in the original method, cloud coverage in three levels and total cloud cover is used as input in a cloud product parameterisation. The resulting global horizontal irradiance is validated against the output of the numerical weather prediction (NWP) model and compared to a year of hourly ground measurements in Trondheim, Norway. To evaluate applicability to the building sciences, vertical irradiance measurements are compared to tilted surface irradiance calculated with the ISO 52010:2017 method. For the location, six-hour forecasting performance is on par with the GHI output of the NWP model (using the cloud layer model and the available weather parameters of the location forecast API). To account for the unpredictability of clouds and improve the short-term forecasting performance beyond 38 % RMSD, 38 % SD and 0.80 R^2 a different approach is needed, like combining model and sky observations.

1 Introduction

Solar radiation is rarely measured in buildings and local data is often a missing component in building energy modelling. Even when it is part of an on-site weather station, maintaining high-quality measurements is difficult and the data is rarely used for control or followed up by analysis. This is changing following the availability of satellite derived solar radiation (such as Copernicus Atmosphere Monitoring Service (CAMS) covering Europe), and data from numerical weather prediction (NWP) models representing detailed atmospheric conditions run in forecast, nowcast or reanalysis mode for respectively long and short-term predictions or historic time-series [1]. Different techniques have been developed to generate solar forecasting and nowcasting data [2, 3]. Nowcasting is a relatively new concept, referring to weather forecasting for the next few minutes to six hours using all immediately available weather data. Sophisticated techniques used for short-term solar predictions may involve combining NWP models with observations from sensor-networks like radiometers or PV-installations, shadow cameras, satellite-derived images, statistics and machine learning methods. In the following work, we simply rely on empirical modelling making use of weather forecasts from an operational NWP model covering the Nordic countries. As it turns out, anyone with basic computer skills can download these location forecasts up to multiple days ahead through web application program interfaces (web API's) from various weather service providers. Some national weather centres, like the Norwegian Meteorological Institute (MET Norway), offer additional model and research data. In the building sciences, ability to access past forecasts can, for example, be valuable to train a model predictive controller for energy management in buildings on actual forecast data.

However, a recent review by Du et. al. (2019) shows that solar radiation forecasts tend not to be offered freely to the public. Their study reveals that only 1 out of 8th public weather API's has started to offer solar radiation parameters [4]. The particular service, Weatherbit, was found to offer meaningful short-term predictions of global horizontal solar radiation, but the diffuse horizontal component was only valid under clear sky conditions. A decomposition into direct and diffuse radiation is essential to calculate the amount of incoming solar radiation on building facades and tilted surfaces, like PV-systems. Empirical models exist to account for differences in the diffuse part between overcast and cloudless weather and variations in ground reflection from snow cover. Interestingly, the other parameters that are available from public weather API's like cloud coverage and precipitation (that are rarely monitored) represent an opportunity to easily implement these phenomena in a processing step to improve solar predictions for building energy management. In this paper, we test if cloud coverage, air temperature, humidity and other parameters are sufficient to estimate diffuse and global solar radiation, even when solar radiation is not available.

We evaluate a method to calculate hourly global solar radiation and improve the calculation of diffuse and

* Corresponding author: kristian.skeie@ntnu.no

vertical surface radiation by accounting for ground reflection based on publicly available data of cloud coverage, temperature and precipitation from a location forecast product covering the Nordic countries (api.met.no). The weather forecasts on yr.no, delivered by MET Norway and NRK, present the same parameters.

The detailed weather forecasts produced by MET Norway provide hourly forecasts for the next 60 hours and 6-hour predictions for the next week. To calculate solar radiation on cloudy days the cloudy sky MAC model developed by Davies and McKay (1982) is used [5]. Instead of basing the prediction on ground observations as in the original method, cloud coverage in three levels and total cloud cover is utilized as input in a model product parameterisation. The resulting global horizontal irradiance is validated against the native output of the NWP model and compared to a year of hourly ground measurements in Trondheim, Norway. To evaluate the applicability to the building sciences, vertical irradiance measurements are compared by calculating tilted surface irradiance with the Perez sky model (ISO 52010:2017) [6]. In an intermediate step, the Engerer diffuse separation model is utilized [7]. Furthermore, to investigate the influence of reflections, an albedo calculation method is also evaluated using additional parameters provided by the forecast; precipitation (to estimate snow cover) and temperature (to account for melting) as input [8].

1.1 Irradiance in forecasts covering Scandinavia

The location forecast of MET Norway is a public weather API with parameters listed in Table 1. Cloud cover in multiple levels and a number of other parameters are available, but solar irradiance is not included. A possible explanation found in literature is that up until recently "this data was not provided as model output because they were deemed unimportant relative to temperature, wind, humidity, and cloud cover" [2]. Acknowledging interest in solar forecasts from electricity operators faced by the intermittency of renewable energy sources, the Swedish Meteorological and Hydrological Institute (SMHI) compared the performance of two NWP systems that cover Scandinavia [12]. The two models are:

• The global Integrated Forecast System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). The two models High Resolution forecast – "HRES" (Deterministic) and Ensemble forecast – "ENS" are evaluated, with a grid resolution of about 9 km and 18 km, respectively.

• The MetCoOp Ensemble Prediction System (MEPS) run by the meteorological services of Sweden, Norway and Finland. The grid spacing of MEPS is 2.5 km and 65 atmospheric levels are covering the Nordic region (higher horizontal resolution but less vertical levels than the IFS model of 137 and 91 levels for HRES and ENS).

The main differences between the involved models are horizontal and vertical resolution, number of ensemble members and perturbation technique of the ensemble members. Shortwave radiation is parametrized the same way in both NWP models based on the Morcrette radiation scheme "McRad" [13, 14]. In July 2017, a new model, ecRad was made operational in ECMWF's integrated forecasting system, it improves on the McRad model in the treatment of clouds and optical gas properties [15]. However, during the analysis period of the SMHI study, the old McRad scheme was still operational. The results display the applicability of NWP models by showing a good fit to the distribution of global and direct horizontal irradiance with better performance from MEPS ensemble mean, and secondly the MEPS control member.

Table 1. Current parameters available in location forecast 1.9 and parameters announced in version 2.0 (in parenthesis).

Available parameters	api.met.no
air pressure at sea level	✓
weather symbol	✓
precipitation amount	✓
air temperature	✓
wind from direction	✓
wind speed	✓
wind speed of gust	✓
relative humidity	✓
cloud area fraction	✓
fog area fraction	✓
cloud area fraction low	✓
cloud area fraction medium	✓
cloud area fraction high	✓
dew point temperature	✓
solar irradiance	-
uv-index clear sky max	(✓)

The results were significant, as previous studies cited in the Swedish paper found that regional NWP models do not necessarily perform better than coarser models. The premise is that coarser models will represent smoother cloud cover that better agree with the observations. In detailed models (grid size<10 km), wrongly placed clouds may be counted twice, first where it is missing and later when it appears. It is suggested that the unpredictability of clouds can be improved by using the ensemble prediction system which is available for both IFS and MEPS. Another common approach to evaluate location forecasts from high resolution models, is to average the global horizontal radiation from all nearby model grid points within a distance to the location. In the post processed forecasts of MET Norway, the cloud area fractions are the median value of a neighbouring 37.5 x 37.5 km grid. This is one of the techniques outlined by Larson (2013) who describe the challenges of modelling solar radiation with NWP in [37].

Forecasting for the next minutes to 6 hours is pursued with different approaches. One is based on observations and the other on the use of NWP models, in combination with postprocessing techniques such as downscaling and the assimilation of ground-based or remote-sensing in situ measurements. How much value post processing can add depends on the weather variable. In the case of wind and temperature, bias-correction and persistence may be used extensively on the assumption that short-term projections will be similar to present conditions. The situation is more

difficult with solar radiation, where rapid changes in local cloud cover may lead to substantial changes in the flux. That is why some commercial forecasting services primarily rely on rapid update cloud tracking provided by meteo-satellite imaging to produce high-quality predictions on less than hourly time scales [37].

MET Norway has released two operational post-processed products that can be characterized as nowcasts, integrating output from MEPS as well as temperature and precipitation measurements from professional stations, citizen stations, and radar. The forecasts are updated every hour as new observations and model fields become available. Wind, temperature, and precipitation is downscaled from 2.5 km to 1 km, whereas Surface Solar Radiation Downwards (W/m^2), a synonym for global horizontal radiation (GHI), relative humidity, air pressure and cloud area fraction also are available parameters from the operational MEPS model. Product documentation is publicly available on (github.com/metno/NWPdocs). This recent development is making GHI from the operational MEPS forecast covering Scandinavia more accessible than before, even though it is not currently included in the location forecast web-API (Table 1).

2 Method

2.1 Approach to modelling

Many packages exist for solar calculations & PV-forecasting, such as the recent open source tools “Solar Forecast Arbiter”, and pvlib for Python [27-28], but no recent scripts have been found to include cloud layering models such as the cloudy sky MAC model. Several scripts are developed to process the weather forecasts in steps using Node-RED, a visual programming tool originally developed by IBM for wiring together hardware devices, APIs and online services using JavaScript functions. A second implementation of the model was written in R for analysis and parameter optimization with rStan, an interface for the probabilistic programming language Stan.

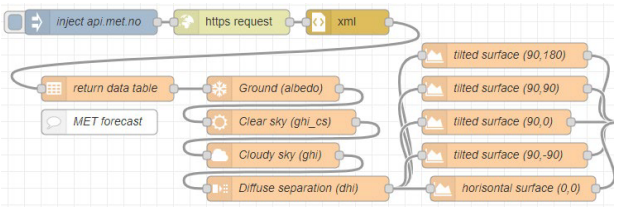

Fig. 2. Flow based script implemented in Node-RED to download forecasts for the next hours, calculate tilted surface radiation, and post the solar forecast to a time-series database.

The runtime of Node-RED is built on Node.js and the interface is accessed from a web browser (Figure 2). A package written in JavaScript called "soljs" was found in the Node package manager that already included solar positioning algorithms and the Perez sky model from ISO 52010:2017. The solar positioning algorithm in this code (MIT lisence) was used as a basis for the ground albedo, clear sky and cloudy sky parametric models and the diffuse separation model (Figure 2).

2.2 Observations

The observations are hourly average measurements of global horizontal irradiance from the station SN68173 operated by MET Norway. Their stations use first-class rated pyranometers (according to ISO 9060) from Kipp Zonen. SN68173 is located on NTNUs campus Gløshaugen, Norway. In Node-RED, the observations from SN68173 were queried by calling the Frost API of MET Norway and sent to a time-series database together with the post-processed forecast. This approach opens the possibility to include the ground observations in the processing steps to improve short-term predication. For vertical irradiance, measurements from the south facade of a lab building on campus was used, the ZEB Living Lab, located ca. 100 meters from the other station. The building facade is oriented 7 degrees westwards from south position. The vertical irradiance is measured by a "LP PYRA 03 AC" from Delta Ohm, which is a second class pyranometer according to ISO 9060.

2.3 Model performance metrics

For model performance assessment, the suggestions from [9] are followed. Therefore, thirteen performance metrics are considered that can be categorised into three classes. A fourth class is included for visual indicators:

• Class A: indicators of the dispersion (e.g. error) of individual points. These are the mean bias difference (MBD), root mean square difference (RMSD), mean absolute difference (MAD), standard deviation of the residual (SD), coefficient of determination (R^2), slope of best-fit line (SBF), uncertainty at 95% (U95) and the t-statistic (TS). A value of 0 is perfect performance for class A metrics, with exception of the SBF and R2 where 1 rather indicates perfect performance.

• Class B: indicators of overall performance. These are Nash-Sutcliffe’s efficiency (NSE), Willmott’s index of agreement (WIA), and Legates’s coefficient of efficiency (LCE). A maximum value of 1 indicates a perfect model.

• Class C: indicators of distribution similitude in percentage. These are the Kolmogorov-Smirnov test integral (KSI), the critical limit overestimation (OVER) and the combined performance index (CPI). A value of 0 for all Class C metrics indicates excellent performance.

• Class D: visual (qualitative) indicators. The Taylor diagram combines information about RMSD, SD and R^2 into one single figure. Information about cumulative distribution and similarity can also be shown graphically.

Most of the statistical indicators in class A except R^2 and SBF) are presented in percentage of the mean observed GHI, O_m = 228.5 W/m2. Multiplying each metric with $O_m/100$ convert them from relative to absolute terms (W/m^2). Only observations with solar height five degrees over the horizon or higher are considered ($\varphi_{zen} < 85$). A condition that is met for N = 3446 hours in the dataset. In Trondheim, the sun angle at noon is below the threshold before 12.1.2019 and after 30.11.2019.

The indicators of distribution likeness, KSI and OVER, are calculated from the definitions in [10]. The

combined performance integral CPI may be the most versatile metric of all, as it combines information about dispersion and bias (through RMSE) and distribution likeness (through KSI and OVER). Therefore, it is argued by [9] that if only one statistical indicator is to be considered, the best choice would be CPI. It is defined as CPI = (KSI + OVER + 2RSMD)/4 as proposed by [9, 11].

2.4 Clear sky model

What makes solar radiation data stand out from other types of time series, is the diurnal cycle due to the apparent position of the sun [3]. Image-based forecasting methods and NWP methods are driven by meteorology, but atmospheric parameters may also play an important role as inputs to other methods. For example, most statistical or machine learning solar forecasting consider the diurnal sun cycle before building a model, and the most common approach is to apply a clear sky model or calculate the solar radiation above the atmosphere (extra-terrestrial flux) using a solar position algorithm [2,3]. Clear sky models can be divided into complexity, or by the number of inputs required. Among the most advanced and well-validated is the REST2 model [16]. Broadband models that fall into the middle category regarding input requirements, such as the MAC model also consistently performed well in validation exercises, especially at low sun angles [7]. A recent state-of-the-art worldwide performance assessment ranked the MAC and RESTv5 as the best scoring among 75 global clear-sky models [17]. In the following section, the MAC model is implemented. MAC is also described and validated in several other studies and has proved well at high latitudes [18–21]. This model was also used by Olseth and Skartveit (1993) in their formulation of an hourly global irradiance model from ground cloud observations in Bergen [22]. Code for a version called MAC2 is conveniently made available on GitHub by the authors of the recent validation study [17]. The inputs of the MAC2 model is broadband aerosol optical depth, total amount of precipitable water vapour (cm), local barometric pressure (mbar) and ground albedo. Equations to calculate aerosol transmittance in MAC2 from Ångström β, α are found in [17].

The performance of clear sky models inevitably depends on the input data. In the past, monthly means or flat assumptions have often been made due to the difficulty in obtaining certain dynamic variables such as aerosols. Today, high temporal atmospheric data is available from satellite (i.e. CAMS Aerosol forecasts) and as parameters in weather forecasts and reanalysis products (i.e. MERRA-2, ERA5) [17]. Nevertheless, monthly gridded data was used below. The Linke turbidity factor is a simple approximation used to describe the atmospheric absorption and scattering of solar radiation under a clear sky. Several datasets exist formatted as monthly tables on a latitude/longitude grid covering the whole world [23, 24]. For this work, a function in the pvlib-python package [26] is used to extract data for Trondheim and interpolate from monthly to daily values. Many studies have found that the Linke turbidity factor (T_{L2}, for an air mass equal to 2) is correlated with precipitable water vapour (u_{H2O}) and the Ångström turbidity beta (β). Six different formulations from literature were found to produce a considerable variation and impact on clear sky model performance in [17]. In the following, one of the formulations (T_{L2Gr}) was reversed to calculate the Ångstrøm beta from the Linke dataset and hourly calculations of water vapour column (derived from forecasted dewpoint temperature). This approach was taken to introduce some dynamic to the model, but may not be without implications in terms of model performance. Comparisons carried out in an intermediate step show that the clear sky output is generally well attenuated to ground measurements from Trondheim April 2019 ("SN68173" in Figure) and the year overall, but some dynamics are lost when compared to models that rely on satellite data, such as Copernicus Atmosphere Monitoring Service (CAMS) clear sky model McClear.

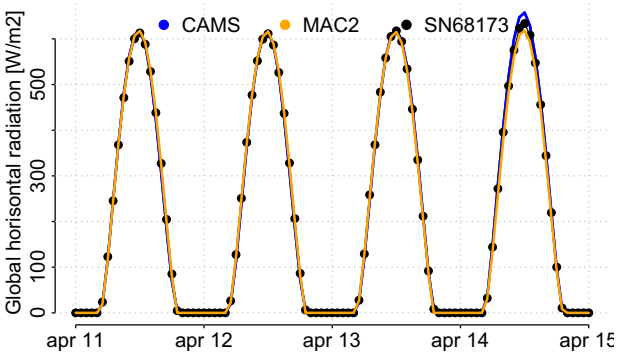

Fig. 3. The clear sky MAC2 model in yellow (with monthly atmospheric data and hourly variables from the NWP model) compare well to ground measurements and to hourly estimates of the CAMS radiation service in blue (W/m^2).

2.6 Ground reflectivity

The ground albedo calculation method uses additional parameters provided by the forecast; precipitation (to estimate snow cover) and temperature (to account for melting) [8]. Dumitrascu (2019) provide an overview of different methods to calculate ground reflectivity for building energy simulation and propose an empirical model with separate calculations for:

- Ground free of snow period
- Snow accumulation or non-melting period
- Snow melting period, where significant variation in ground albedo from one day to another was observed because of ambient temperature above zero and sporadic snowfalls. The calculation uses two melting rates controlled by a threshold temperature increasing towards spring [8].

The inputs of the proposed model are snow age (in hours), outdoor temperature, day of year, solar zenith angle and sky transmissivity (K_t). Hence, the hours since last snowfall was calculated from the forecast. A conservative approach was taken in this study due to Trondheim's location in a coastal climate, hence the non-melting period was neglected and the highest albedo of the snow melting and the ground free of snow calculations was assumed.

2.7 Cloudy sky model

In the 1982 paper introducing the clear sky MAC model, Davies and McKay evaluated four cloudy sky models [5]. Among them, the cloud layer MAC model has the best performance, and with its three-layer product, it matches the information available in the location forecast web API. The percentage of total cloud coverage and cloud amounts reported in three layers (high, medium and low clouds) and the clear sky global horizontal radiation are inputs. In another study (1995) using ground observations of clouds in the Arctic, the MAC cloud layer model produced the lowest range of error compared to 11 other simple cloudy sky models [29]. More advanced methods for cloud detection and classification at mid- and high latitudes exploits remote sensors on geostationary and polar-orbiting satellites, such as the software packages of the EUMETSAT Satellite Application Facilities on Support to Nowcasting and Very Short Range Forecasting (NWC-SAF) providing operational cloud imaging products over the Arctic and Europe [30-32].

Simple cloud layer models like MAC take into consideration the variations in cloud transmittance with cloud type. Total cloud transmission is the product of the individual layer transmission $[(1 - C_i) + t_i C_i]$. The equation can be expanded into a geometric series to account for multiple reflections between the ground and atmosphere. In the following equation only one reflection from the ground and one from the cloud is considered (or one can consider the effect of multiple reflections to be incorporated in the denominator).

$$\mathrm{GHI} = G_{cs} \frac{\prod_{i=1}^{j}[(1 - C_i) + t_i C_i]}{1 - \alpha_b \alpha_s} \quad (1)$$

In Eq.1 G_{cs} is the hourly global horizontal radiation from a clear sky and C_i and t_i are cloud amount and transmittance for the ith layer, α_b is the atmospheric albedo for radiation reflected from a ground surface with albedo α_s.

In the paper by Davies McKay, cloud transmissivity for 16 cloud types are presented following the cited work of Haurwitz (1946) on insolation in relation to cloud type:

$$t_i = A_i \, exp(-B_i \, m_r) \quad (2)$$

Were A_i and B_i are table values for different cloud types dependent on optical airmass m_r. Other studies have used constant t_i values like 0.32, 0.42 and 0.78 for low medium and high cloud layers [33]. As a starting point, the table values of [26] are used (table 2).

Table 2. Typical cloud albedo and absorptivity of clouds [26].

Level	Cloud type	α_{ci}	t_i
High	Cirrus	0.25	0.75
Middle	Alto	0.60	0.38
Low	Low cloud	0.70	0.36

Atmospheric and surface albedo is calculated using the formulas below, described in the studies of [18, 19, 33].

$$\alpha_b = \alpha_b(R)(1 - C_0) + \alpha_b(a) + \overline{\alpha_c}\, C_0 \quad (3)$$

where constant $\alpha_b(R) = 0.0685$ is molecular scattering affecting only the clear sky part, $\alpha_b(a)$ is the aerosol scattering below clouds and $\overline{\alpha_c}$ is the cloud albedo. In this paper, no attempts are made to cover the complex reflections between various cloud layers. For simplicity, we use total cloud coverage C_0 and the cloud albedo is approximated according to cloud level by equation 4.

$$\overline{\alpha_c} = \frac{\sum \alpha_{ci} C_i}{\sum C_i} \quad (4)$$

Another change from the original model is that the aerosol scattering below clouds $\alpha_b(a)$ is set to zero so that the model attenuation under a clear sky is close to zero (GHI $\approx G_{cs}$). Finally, an extension of the model is tested that include precipitation, derived from the assumption that cloud transmissivity is lower during forecasted showers:

$$t_i = (t_i + (t_{rain}\mathrm{Rainfall}[0,1]) \quad (5)$$

Rainfall events are not parametrised directly in the MAC cloud model but can be considered indirectly when choosing between the 16 cloud type transmissivities to represent current weather conditions. In a report by the same authors [18] another cloud layer model variant is evaluated. The Josefson model differs from MAC in some aspects notably by applying fixed cloud layer transmissivities and 30 % reduction in total cloud transmissivity if precipitation occur during the hour, and 20 % if it ended within the hour. In our model (Eq. 5), rain transmissivity is added or subtracted from t_i in each layer.

2.8 Diffuse separation model

The diffuse separation model is the Engerer separation model [34] that was recently reparametrized as the Engerer2 model and released with source code available by the authors on GitHub [35]. The model requires global horizontal irradiance and ground albedo as input as well as solar position and clear sky irradiance which in this case was provided with MAC2.

2.8 Inclined surface irradiance (ISO 52010)

See section 2.1 for description. An alteration to the equations in ISO 52010 [6] is enabling setting the time shift parameter of the solar positioning algorithm to other values than -0.5 hours, in order to allow other time steps than one hour. With hourly solar data, the convention is that the timestamp 14:00 refers to the time interval 13:00 to 14:00, so the time shift should be set to the mid of the interval (in this case hour) leading up the current one.

3 Results

3.1 Cloudy sky model parametrisation

First the cloudy sky model was tested with values from literature (Table 4). Then a simple model fitting of the parameter was performed using Stan, one time on the GHI output of the forecast model (NWP) and one time on the observations from station (SN68173). The model form is:

$$GHI_{obs} - GHI_{pred} = normal(0, \sigma) \qquad (6)$$

The result of the model fitting is presented in Table 4. The values from literature were used as initial model parameters, assigned weakly informed normally distributed priors with $\sigma = 0.5$, a lower bound of 0 and upper bound of 1, to help provide identifiability (except t_{rain} that had bounds -1 to 1).

Table 4. Parameters from literature, and results of training the model on either global horizontal irradiance from forecast (NWP) or observations (SN68173).

Parameters	MAC_L *Literature*	MAC_N *Fit to NWP*	MAC_O *Fit to obs.*
$t_{i;low}$	0.36	0.45	0.49
$t_{i;med}$	0.38	0.41	0.51
$t_{i;hig}$	0.75	0.87	0.77
$\alpha_{i;low}$	0.70	0.32	0.03
$\alpha_{i;med}$	0.60	0.94	0.12
$\alpha_{i;hig}$	0.25	0.03	0.11
t_{rain}	0.00	t_i-0.08	t_i+0.03
$\alpha_b(R)$	0.07	0.04	0.006

Notice the increased transmittance during hours with rainfall when fitted to observations versus the expected decrease when fitted to the GHI output of the forecast model. Figure 4 show density plots of the posterior parameter distributions for the model, when fitted to observations (MAC_O in table 4).

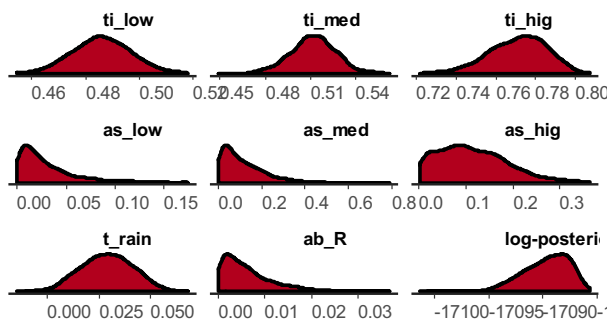

Fig. 4. Parameter density distribution from the cloudy sky model training on the observed GHI (see table 4).

Performance assessment following the framework introduced in the method section (Table 2) show that the NWP GHI model output has the overall best score, but even the cloudy MAC_L model with parameters from literature perform reasonably well with slightly lower class C scores, indicating less distribution likeness.

Table 2. Performance statistics of the NWP and cloudy sky MAC model validated against 60-min measured data. The percent statistics refer to the mean observed GHI, 228.5 W/m^2 for N=3446 hours ($\varphi_{zenith} < 85^{\circ}$).

	NWP	MAC_L	MAC_O
Mean GHI (W/m^2)	**237**	**254**	**238**
Class A (%)			
MBD	3.90	11.34	4.60
RMSD	41.98	41.89	37.77
MAD	26.47	26.99	25.10
SD	41.80	40.32	37.46
R^2	0.78	0.79	0.80
SBF	0.92	0.90	0.83
U95	116.11	113.96	104.22
TS	5.47	16.51	7.21
Class B			
NSE	0.75	0.75	0.80
WIA	0.94	0.94	0.94
LCE	0.62	0.61	0.64
Class C (%)			
KSI	62.7	117.5	73.6
OVER	4.3	44.2	25.5
CPI	37.7	61.4	43.6

The Taylor diagram show that the difference between the model scores on key Class A metrics are minor, indicating the same level of error of individual points. An additional linear correction factor was applied to the NWP output and the optimized cloudy sky model "MAC_O" to see how much it impacts the overall scores. Results of the scaling are given in light green and pink.

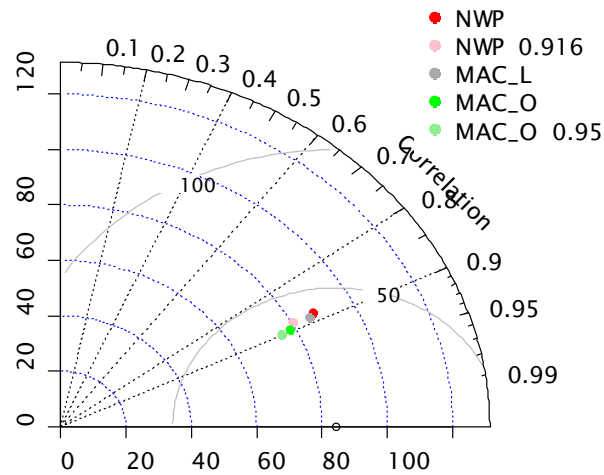

Fig. 5. The Taylor diagram show relatively little difference in RMSD (%), SD (%) and Pearson correlation between models.

Finally, the distribution likeness is investigated further using plots of CDF (Figure 6, left side) and absolute difference in CDF to the observation data (right side). In [10] we read that *"The KSI parameter (Kolmogorov–Smirnov test Integral) is defined as the integrated differences between the CDFs of the two data sets"* while in the OVER-score *"the integration is calculated only for those differences between the CDFs that exceed a critical limit, D_c"*. The critical limit, D_c is shown as a dotted line in the right plot.

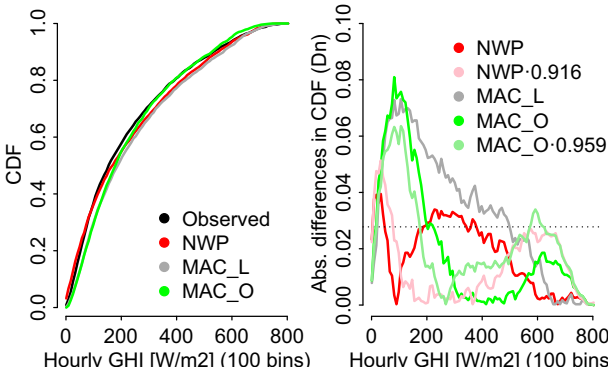

Fig. 6. Plot (left) of the cumulative distribution function for the models compared to the observations and (right) the absolute difference between them and the observations for Trondheim. The dotted line represent the critical value used in the KSI, OVER and CPI statistical tests presented in Table 2.

3.2 Calculated vertical irradiances

The diffuse part is calculated from the MAC_O GHI model output and used in the calculation of vertical irradiance to be compared to the measurements from a south facing façade pyranometer (see section 2.2). The focus is on the monthof March when the vertical irradiance is high due to a low solar position in Trondheim, little shading from trees, and ground reflectance that varies due to snowfall and melting.

Figure 7 show that the day to day variations in solar energy on the façade is generally well represented, but on some days with mixed clouds, the forecast error is relatively large. 10-minute observations of direct irradiance on the façade are also shown with black points, (lines are hourly data) to indicate the cloud effects. The overestimation of solar irradiance on clear days is not entirely due to modelling error but could be partly prescribed to the measurement accuracy and maintenance (cleaning, snow covering the sensor). It was noticed that another pyranometer of the same type mounted horizontally also output less than the station maintained by Met Norway (SN68173) located <100 m away (Figure 8). This mismatch makes further analysis difficult. Method section 2.2 refer to the verical and horisontal pyranometers on the facility having lower rated accuracy. This may explain parts of the overestimation on clear days, while snow may cover the sensor and influence the measurements on days with heavy snowfall.

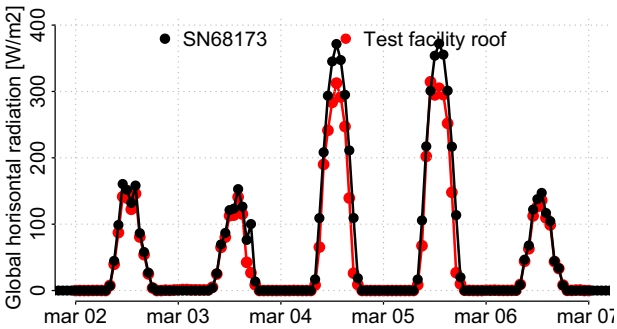

Fig. 8. Global horizontal irradiance measured on the test facility (in red) and the station SN68173 used in the study.

Conclusions

The study shows that it is possible to get six-hour forecasting performance close to the GHI output of the NWP model using the parametric cloud layer model and the available weather parameters of the location forecast web API. To account for the unpredictability of clouds and improve the short-term forecasting performance beyond 38 % RMSD, 38 % SD and 0.80 R^2 a different approach is needed. Weighting ground measurements with a Kalman filter or regression would be one way to improve short term performance and bias. Replacing short term predictions of the cloudy sky model and NWP with a model relying on real-time satellite image projections is another approach that is taken in many commercial nowcasting services [36].

In the validation, the MEPS control member, or more specifically the post-processed forecasts, was used. To our knowledge, it is the basis for the location forecasts and available sooner than the ensemble. The study of [12], showed that the ensemble mean had better forecasting performance than the control member. Future studies should investigate if the ensemble describes the probability well for day-ahead forecasts. Taking into consideration sky conditions in the

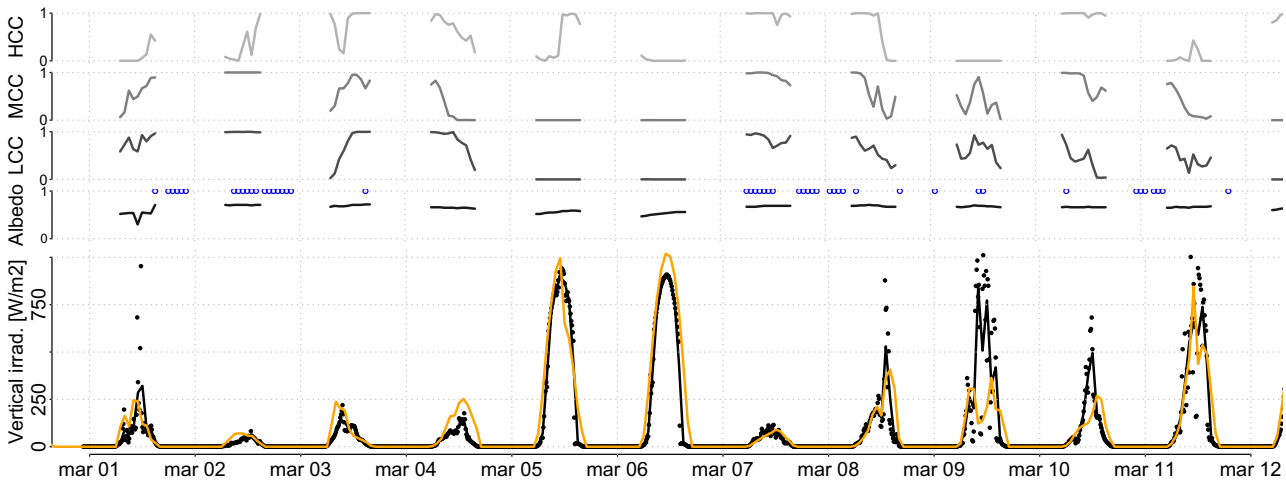

Fig. 7. Observations of hourly vertical irradiance in March (black line), 10-minute values (black dots) and calculated total hourly direct (yellow line). Layered cloud cover (HCC, MCC, LCC 0 to 1) and ground albedo (snow events in blue) are illustrated above.

nearby 2.5 km model grid points is another suggestion. In the post processed forecasts used in the validation, the cloud area fractions in each layer of a grid point is the median value of a neighbouring area (37.5 x 37.5 km), while the solar radiation was extracted directly from the model.

The clear sky radiation and ground albedo used in the cloud layer model are also utilized to calculate diffuse radiation and ground reflections and are input to the vertical surface radiation calculations. The results show that these variables can easily be calculated in a scripting environment when a new forecast is available to be used in solar predictions for energy management in buildings or local energy generation. A closer study of the impact of the value of these effects and using other weather variables from the location forecast is planned.

Acknowledgements

Based on data from MET Norway, under Norwegian license for public data (NLOD). This paper has been written within the Research Centre on Zero Emission Neighbourhoods in Smart Cities (FME ZEN). The authors gratefully acknowledge the support from the ZEN partners and the Research Council of Norway.

References

1. Y.H. Huang, Using Satellite-Derived Solar Radiation to Create Weather Files of Unprecedented Accuracy and Reliability, in [25]
2. D. Yang, J. Kleissl, C.A. Gueymard, H.T.C. Pedro, C.F.M. Coimbra, Solar Energy 168, 60 (2018)
3. R. Inman, H. T.C. Pedro, C. F.M. Coimbra, Solar forecasting methods for renewable energy integration, Vol. 39 (2013)
4. H. Du, et. al. Nowcasting methods for optimising building performance, in [25]
5. J.A. Davies, D.C. McKay, Solar Energy 29, 55 (1982)
6. ISO 52010:2017, Standard, International Organization for Standardization, Geneva, CH (2017)
7. N.A. Engerer, F.P. Mills, Solar Energy 120, 9 (2015)
8. L. Dumitrascu, I. Beausoleil-Morrison, On the Prediction of Ground-Reflected Solar Radiation and its Relevance in the Context of Building Performance Simulation (BPS), in [25]
9. C.A. Gueymard, Renewable and Sustainable Energy Reviews 39, 1024 (2014)
10. B. Espinar, L. Ramírez, A. Drews, H.G. Beyer, L.F. Zarzalejo, J. Polo, L. Martín, Solar Energy 83, 118 (2009)
11. C.A. Gueymard, Solar Energy 86, 2145 (2012), progress in Solar Energy 3
12. T. Landelius, M. Lindskog, H. Körnich, S. Andersson, Adv. Sci. Res. 15, 39 (2018)
13. J.J. Morcrette, H.W. Barker, J.N.S. Cole, M.J. Iacono, R. Pincus, Monthly Weather Review 136, 4773 (2008)
14. L. Bengtsson, U. Andrae, M. others et al., Monthly Weather Review 145, 1919 (2017)
15. R. Hogan, A. Bozzo, Submitted to J. Adv. Modeling Earth Sys., May 2018 (2018)
16. C.A. Gueymard, Solar Energy 82, 272 (2008)
17. X. Sun, J.M. Bright, C.A. Gueymard, B. Acord, P. Wang, N.A. Engerer, Renewable and Sustainable Energy Reviews 111, 550 (2019)
18. J.A. Davies, D.C. McKay, G. Luciani, M. A. Wahab, Report, Atmospheric Environment Service (1988)
19. J.A. Davies, D.C. McKay, Solar Energy 43, 153 (1989)
20. Godøy, Report, Norwegian Meteorological Institute (DNMI) (2000)
21. J.S. Stein, C.W. Hansen, M.J. Reno, Report 2012-2389 SAND National Laboratories (2012)
22. J.A. Olseth, A. Skartveit, Solar Energy 51, 197 (1993)
23. J. Remund et al., Worldwide linke turbidity information (2003)
24. J. Remund et al., Aerosol optical depth and linke turbidity climatology (2010)
25. V. Corrado, A. Gasparella, eds., Proceedings of Building Simulation 2019: 16th Conference of IBPSA (IBPSA, 2019), ISBN 978-1-7750520-1-2
26. A. Henderson-Sellers, K. McGuffie, A climate modelling primer, book (Wiley,1987)
27. Holmgren F. W. et. al., Journal of Open Source Software, 3(29), 884 (2018).
28. Holmgren F. W. et. al., Solar forecast arbiter, presentation US Department of Energy DOE (2019)
29. K.L. Young, M.K. Woo, D.S. Munro, Solar Energy 54, 33 (1995)
30. A. Dybbroe, K.G. Karlsson, A. Thoss, Journal of Applied Meteorology 44, 39 (2005)
31. F. Kaspar, R. Hollmann, M. Lockhoff, K. et. al., Advances in Science and Research 3 (2009)
32. J. Garcia P. et. al., Use of NWCSAF NWC/GEO software package with MSG, Himawari-8/9 and GOES-13/16 satellites (2019), in Proceedings of the 2019 Joint EUMETSAT/AMS/NOAA Conference.
33. M.A. Wahab, Solar Energy 51, 241 (1993)
34. N. Engerer, Solar Energy 116, 215 (2015)
35. J.M. Bright, N.A. Engerer, Journal of Renewable and Sustainable Energy 11, 033701 (2019)
36. P. Blanc, J. Remund, L. Vallance, Short-term solarpower forecasting based on satellite images, Wood-head Publishing Series in Energy (2017)
37. Larson, Solar Energy Forecasting and Resource Assessment pp. 299–318 (2013)