Meteorologische Zeitschrift, PrePub DOI 10.1127/metz/2016/0676

Verification of ECMWF and ECMWF/MACC's global and direct irradiance forecasts with respect to solar electricity production forecasts

M. Schroedter-Homscheidt[1*], A. Benedetti[2] and N. Killius[1]

[1]German Aerospace Center (DLR), Earth Observation Center, Oberpfaffenhofen, Germany
[2]European Centre for Medium-Range Weather Forecasts (ECMWF), Reading, United Kingdom

(Manuscript received February 27, 2015; in revised form May 10, 2016; accepted May 13, 2016)

Abstract

The successful electricity grid integration of solar energy into day-ahead markets requires at least hourly resolved 48 h forecasts. Technologies as photovoltaics and non-concentrating solar thermal technologies make use of global horizontal irradiance (GHI) forecasts, while all concentrating technologies both from the photovoltaic and the thermal sector require direct normal irradiances (DNI). The European Centre for Medium-Range Weather Forecasts (ECMWF) has recently changed towards providing direct as well as global irradiances. Additionally, the MACC (Monitoring Atmospheric Composition & Climate) near-real time services provide daily analysis and forecasts of aerosol properties in preparation of the upcoming European Copernicus programme. The operational ECMWF/IFS (Integrated Forecast System) forecast system will in the medium term profit from the Copernicus service aerosol forecasts. Therefore, within the MACC-II project specific experiment runs were performed allowing for the assessment of the performance gain of these potential future capabilities. Also the potential impact of providing forecasts with hourly output resolution compared to three-hourly resolved forecasts is investigated. The inclusion of the new aerosol climatology in October 2003 improved both the GHI and DNI forecasts remarkably, while the change towards a new radiation scheme in 2007 only had minor and partly even unfavourable impacts on the performance indicators. For GHI, larger RMSE (root mean square error) values are found for broken/overcast conditions than for scattered cloud fields. For DNI, the findings are opposite with larger RMSE values for scattered clouds compared to overcast/broken cloud situations. The introduction of direct irradiances as an output parameter in the operational IFS version has not resulted in a general performance improvement with respect to biases and RMSE compared to the widely used Skartveit et al. (1998) global to direct irradiance conversion scheme. Cloudy situations and especially thin ice cloud cases are forecasted much better with respect to biases and RMSE, but large biases are introduced in clear sky cases. When applying the MACC aerosol scheme to include aerosol direct effects, an improvement especially in DNI biases is found for cloud free cases as expected. However, a performance decrease is found for water cloud cases. It is assumed that this is caused by the lack of an explicit modelling of cloud-aerosol interactions, while other meteorological forcings for cloud processes like the temperature field are modified by the aerosols.

Keywords: irradiance forecast, verification, global irradiance, direct normal irradiance, scattered clouds, aerosols, MACC

1 Introduction

Solar surface irradiance forecasting is of major importance for the management of solar energy providing larger shares of our electricity supply due to the energy system transition process. Especially, applications from the grid management sector require intra-day and day-ahead irradiance forecasts in hourly resolution. Photovoltaics and non-concentrating solar thermal technologies can make use of both the direct and the diffuse component and therefore need global horizontal irradiance (GHI) forecasts, while all concentrating technologies both from the photovoltaic and the thermal sector require direct normal irradiance (DNI) forecasting. They can only concentrate the irradiance normal to their concentrating plane to their focus line or point. Currently, numerical weather prediction schemes provide GHI but mostly no DNI forecasts. Therefore, empirical GHI2DNI conversions are widely used.

The European Centre for Medium-Range Weather Forecasts (ECMWF) has recently changed its output variable selection in order to include direct irradiances. This has been highly welcomed by the solar energy community needing DNI either for concentrating solar technologies or for the assessment of direct/diffuse ratios when calculating tilted irradiances for the plane of photovoltaic panels. Additionally, the MACC and MACC-II (Modelling Atmospheric Composition and Climate, http://www.gmes-atmosphere.eu/) projects are currently establishing the global and regional atmospheric environmental services delivered as

*Corresponding author: Marion Schroedter-Homscheidt, Deutsches Zentrum für Luft- und Raumfahrt (DLR), 82234 Wessling, Germany, e-mail: marion.schroedter-homscheidt@dlr.de

a component of Europe's Copernicus initiative. The MACC near-real time services provide daily analysis and 120 h forecasts of various parameters based on prognostic aerosol properties. The operational ECMWF/IFS forecast system will in the medium term profit from the MACC aerosol forecasts. Therefore, within the MACC-II project specific experiment runs were performed allowing the assessment of the performance gain of potential future capabilities making use of the aerosol capabilities developed for Copernicus and/or the provision of hourly forecasts instead of three-hourly forecasts.

Specific focus of this work is laid on the assessment of DNI forecasts. This parameter is of special interest for concentrating solar power and is not well evaluated so far, while global irradiance forecasts have been evaluated more thoroughly. Perez et al. (2013) assessed several global irradiance forecasts based on ECMWF, GFS, AEMet-HIRLAM, Skiron/GFS and WRF/GFS model runs and additional post-processing modules. Generally, the ECMWF-based forecast was performing very well or even best at most locations in Germany, Austria and Spain. The ECMWF was performing better than the GFS-driven WRF model in all locations in Europe, the US, and Canada. Mathiesen and Kleissl (2011) showed that for the US 'ECMWF is the most accurate forecast in cloudy conditions, while GFS has the best clear sky accuracy'. Having this finding in mind, the assessment of ECMWF's new aerosol capabilities is of special interest. There is a variety of other validation studies for global irradiance forecasts, but they typically deal either with the GFS model (Perez et al., 2010), Canada's GEM (Pelland et al., 2011), regional models like WRF (Weather Research and Forecasting Model, Lara-Fanego et al., 2012), the Regional Atmospheric Modeling System (RAMS, Ronzio et al., 2013) or the Australian Limited Area Prediction System (LAPS, Gregory et al., 2012). Recently, Haiden and Trentmann (2015) presented a verification of cloudiness and GHI for ECMWF in the Alpine region. For direct irradiances, there are only a few studies available: Wittmann et al. (2008) show case studies of ECMWF day-ahead forecast applied to concentrating solar thermal power plant operations. For July 2003 they find relative RMSE values of 19 % for global and 42 % for direct irradiances at a power plant's location in Southern Spain. Breitkreuz et al. (2009) evaluate the coupling of aerosol modelling with the WRF model resulting in an improvement of the day-ahead clear sky forecast while the cloudy cases are less accurate than the ECMWF forecast. Lara-Fanego et al. (2012) evaluate a GHI2DNI post-processing for WRF-model based GHI forecasts coupled to MODIS satellite retrievals for aerosols and ozone for Southern Spain. They report on relative DNI RMSE values for the first, second, and third day with only a slight increase from 61 to 62 and 63 % (no night time values, hourly values). Marquez and Coimbra (2011) report on DNI forecasts made by the help of artificial neural networks as a post-processing for numerical weather predictions provided by the US National Weather Service's (NWS) forecasting database. They find relative RMSE values between 37 and 40 % for different models and for the day ahead, while the same day has relative RMSE values between 28 and 35 % (no night time values, hourly values). Troccoli and Morcrette (2012) evaluate all three irradiance components of the ECMWF model for Australian stations, but provide results only on monthly averages. Kraas et al. (2013) evaluate the economic value of DNI day ahead forecasts based on a commercial model output statistics (MOS) system for the Spanish electricity market conditions and provide relative RMSE values between 56 and 77 % for different years at a specific power plant site in Southern Spain. They report especially on the increasing relative mean bias error as a function of DNI variability over the day.

A few studies exist on the assessment of forecast accuracy as a function of cloudiness, e.g. on the differentiation of cloud and cloud-free conditions (Lara–Fanego et al., 2012), the dependence on cloud coverage with systematic overestimations of about 100 W/m^2 in cloudy conditions (Lorenz et al., 2009), parameters like cloud index or clear sky index (Mathiesen and Kleisl, 2011) or recently Haiden and Trentmann (2015) in the Alpine region. Besides these approaches, Ronzio et al. (2013) are the first to investigate forecast accuracy as a function of cloud types like cumuliform or stratiform clouds, height level of clouds, thin ice vs. other clouds and overlapping clouds as provided by the Satellite Application Facility on support to Nowcasting and Very Short-Range Forecasting (NWCSAF).

Focusing on ECMWF's capabilities and in relation to existing studies, the following topics will be adressed below:

a) Evaluation of the ECMWF/IFS GHI and GHI2DNI for several years – allowing the comparison of different radiation schemes applied in the operational IFS over the recent years. This evaluation is extended towards different cloud situations – discriminating optically thin cirrus clouds from optically thick water/mixed phase clouds as well as overcast/broken and scattered cloud conditions.

b) Evaluation of ECMWF IFS/DIR2DNI vs. the IFS GHI2DNI allowing the assessment of the new output parameter direct irradiance (DIR) compared to the previously needed empirical GHI2DNI conversion scheme. We will see that there are drawbacks and therefore, we assess them in the next two topics.

c) Evaluation of an ECMWF's MACC-II experimental forecast providing GHI and DIR2DNI in hourly temporal resolution versus a control DIR2DNI and GHI run with three-hourly resolution. This allows the assessment of the value of an hourly forecast resolution.

d) Evaluation of an ECMWF's MACC-II experimental forecast providing GHI and DIR2DNI in hourly temporal resolution and with the new MACC aerosols. This comparison is made versus an experiment with hourly resolution, but including the standard IFS's aerosol climatology. This allows the assessment of the new MACC aerosol scheme.

A description of forecasts, ground-based, and satellite-based datasets used is given in Section 2. Section 3 provides the methodology applied, while Section 4 discusses results for several evaluation studies and Section 5 concludes the study.

2 Datasets

Several ECMWF based forecasts are under evaluation in this study. This includes (a) the operational model with its different versions depending on the year and the forecast cycle used; (b) post-processing approaches to generate direct normal irradiances from the operational IFS; and (c) experimental runs generated in the MACC-II project with updated aerosols and hourly temporal resolution.

2.1 ECMWF's operational IFS

The operational Integrated Forecast System (IFS) provides ECMWF's standard forecast including GHI forecasts ('IFS GHI'). The parameter SSRD (solar surface radiation) is the 3-hourly resolved irradiation sum of the previous 3 hours. Several radiation schemes have been used during the different years (ECMWF, 2013). Since 2 May 1989 (cycle 32) the radiation was modelled based on a delta-Eddington approximation (Fouquart and Bonnel, 1980) as described in Morcrette (1991 and 2002). Since 5 June 2007 (cycle 32r2) a new radiation scheme RRTM-SW (McRad, Morcrette et al., 2008a and 2008b) is applied on the T399 grid (0.45 °) and interpolated to a domain-averaged radiation in the T799 grid box (ECMWF, 2009). It is based on the McICA (Monte Carlo Independent Column Approximation) scheme. In cycle 36r4, November 2010, a new prognostic microphysics scheme (Forbes et al., 2011) for the parameterization of stratiform clouds has been introduced. Finally, since 18 May 2011 (cycle Cy37r2) direct irradiance has additionally been included in the output parameters. The operational IFS provides the following regular spatial resolution: 0.5 ° spatial grid before Feb 2006; an increased T799 (0.25 °) spatial resolution since Feb 2006 (cycle 30r1), and finally a T1279 (0.15 °) spatial resolution since 26 Jan 2010 (cycle 36r1). Since IFS model cycle 26r3 (7th October 2003) aerosols in the IFS are based on a monthly mean climatology following Tegen et al. (1997). The differences between the previously used and the Tegen et al. (1997) climatology are further discussed in Tompkins et al. (2005) and Rodwell and Jung (2008).

An hourly temporal forecast resolution is needed following the requirements of those electricity markets being relevant for the grid integration of solar power. The temporal interpolation from the original 3-hourly irradiation sum towards an hourly global irradiance forecast is performed following Breitkreuz (2008). Due to the strong non-linear behaviour of irradiances during the day, the explicit interpolation of irradiances is not recommended. Instead of irradiances, the clear sky index – being defined as the ratio of the global irradiance divided by the clear sky global irradiance is used. Applying a clear sky model (Hoyer-Klick et al., 2010), the clear sky index is calculated for each irradiation sum provided in the IFS forecast. The clear sky index is then interpolated linearly for each minute and averaged to hourly values. Finally, by using the hourly resolved clear sky model again, the global irradiance is calculated in hourly averaged resolution. The clear sky model inputs are based on the MACC-II reanalysis for aerosols, water vapour and ozone for 2012 cases shown here, and on MATCH aerosols (Schroedter-Homscheidt and Oumbe, 2013) for the multi-annual assessment in Section 4.1 as the MACC-II reanalysis is only available from 2003 onwards.

For the next forecast approach, an empirical conversion is used to derive direct irradiance values following the approach of Skartveit et al. (1998). These are finally transferred to the hourly averaged DNI values – named 'IFS GHI2DNI' further on - by using the cosine of the sun zenith angle. The underlying aerosols are also the standard IFS monthly mean climatology following Tegen et al. (1997) and, also the spatial resolution is the same.

Starting with 18 May 2011 (cycle Cy37r2) the ECMWF IFS was extended to provide clear-sky and total-sky direct fluxes at the surface as additional parameters – this dataset is named 'IFS DIR2DNI' in the following sections. The radiation scheme has not been changed compared to the description above, but the output variable list has been extended.

The 3-hourly direct irradiation is also interpolated to hourly averaged irradiances and also for this dataset, the aerosols in the IFS are based on Tegen et al. (1997). For all evaluations below the forecast started at 00 UTC is used to ensure that it is available at the day-ahead electricity trading closure times during the late morning or around noon depending on the country.

2.2 MACC-II experimental runs

The g3rs experiment provides forecasts for a period from 1 December 2011 to 31 December 2012. Please note that the 'g3rs' name and similar acronyms below are ECMWF internal experiment identifiers which are included here to ensure finding the experiment runs in the future. It has no further meaning. These forecasts are starting every day at 00 UTC with a forecast length of 48 hours. This experiment provides three-hourly forecast output resolution as the currently operationally available standard IFS temporal resolution.

It includes both global and direct irradiance parameters; but includes only the standard IFS aerosol climatology. The g3rs experiment is based on meteorology and aerosols from the MACC reanalysis (fbov experiment) and provides irradiance forecasts at a TL255 (0.7 °) resolution – which is slightly less spatially resolved than the radiation scheme being applied in the IFS on the T399 (0.45 °) grid. It also uses the IFS model cycle Cy38r2 which went operational in the IFS on 25 June 2013. The forecasts are retrieved and therefore interpolated on a 0.15×0.15 ° grid as today's standard IFS spatial resolution. It serves as reference for the other MACC-II experiments g3q4 and g3o4 which have enhanced temporal resolution and aerosol capabilities.

The g3q4 experiment provides forecasts for the same period, but with hourly output resolution instead of the standard IFS' 3-hourly temporal output resolution. All other characteristics are the same as described above.

The g3o4 experiment provides hourly forecast resolution instead of the standard IFS' 3-hourly temporal output resolution. Also, it includes the GEMS aerosol scheme as implemented in the MACC reanalysis instead of the standard IFS aerosol climatology. A detailed description of the ECMWF forecast and analysis model including aerosol processes is given in Morcrette et al. (2009) and Benedetti et al. (2009). The initial package of ECMWF physical parameterisations dedicated to aerosol processes mainly follows the aerosol treatment in the LOA/LMD-Z model (Boucher et al., 2002; Reddy et al., 2005). Five types of tropospheric aerosols are considered: sea salt, dust, organic and black carbon and sulphate aerosols. The aerosols are fully coupled with the meteorology. Prognostic aerosols of natural origin, such as mineral dust and sea-salt, are described using three size bins. Emissions of natural aerosols depend on model parameters (surface winds among others). Anthropogenic emissions are specified using current emission inventories. Biomass burning emissions are taken from the Global Fire Assimilation System (GFAS) inventory (Kaiser et al., 2012). Moderate Resolution Imaging Spectroradiometer (MODIS) satellite observations from the dark target approach are routinely assimilated over dark surface regions (e.g. excluding deserts) in a 4D-Var framework as described by Benedetti et al. (2009). It provides mass concentrations and aerosol optical depth for five species being dust, sea-salt, black carbon, organic carbon, and sulphates. This aerosol scheme has been further described and validated against AERONET ground measurements by Cesnulyte et al. (2014). Within this experiment run, the direct aerosol effects are included, while indirect aerosol effects are switched off as they are still under development. All other model configuration parameters are the same as in g3rs and g3q4 IFS runs.

2.3 Ground-based observations

In this study, the focus is laid on the verification of DNI forecasts – therefore, only ground-based observations providing the direct and the global component are of interest. Additionally, direct irradiance measurements are highly sensitive to daily cleaning of the sensors and require a rigorous data quality control. Therefore, the study focuses on measurements from the Baseline Surface Radiation Network (BSRN, Ohmura et al., 1998) being a part of the World Climate Research Programme (WCRP) and the EnerMENA network. BSRN stations provide ground-based datasets of GHI, DNI, and the diffuse hemispherical irradiance (DHI) component in a 1-min temporal resolution. The quality control procedures are described in a report of Long and Dutton (2012). Only years with data coverage over all seasons have been taken into account.

EnerMENA is a project being fully named 'Towards a Sustainable Implementation of Solar Thermal Power in the MENA Region). MENA represents the Mediterranean Europe and Northern Africa regions. It includes a dedicated ground measurement program 'enerMENA meteo network' applying ventilated CMP21 Secondary Standard Kipp & Zonen pyranometers and Kipp & Zonen CHP1 First Class pyrheliometer instruments at various locations in Algeria, Egypt, Jordan, Morocco, and Tunisia. For this study and its year 2012, quality controlled data from two stations in Jordan and Tunesia operated by University of Jordan and Centre de Recherché et Technologies de l'énergie (CRTEn) were already available.

Additionally, DLR provides 1-min measurements from the Plataforma Solar de Almeria site in Southern Spain. The station is equipped with First Class pyrheliometers (ISO 9060, 1990) to measure DNI and Secondary Standard Pyranometers (ISO 9060, 1990) mounted on a solar tracker with sun sensor. Data gaps in the data set from these sensors are filled with Rotating Shadowband Irradiometer measurements. Further information in the instrumentation can be found in Geuder et al. (2009). This data has been quality controlled following the recommendations of the MESOR project (Beyer et al., 2008) which is an extension to the BSRN quality control standards with special focus on the solar energy sector.

In this study we use the stations given in Table 1. All measurements have been averaged to hourly time series with the hour being defined from minute 1 of the previous hour to minute zero of the hour being named (Beyer et al., 2008). Anyhow, perfect agreement with station observations and model runs at coarse resolution is always challenging, especially for stations such as TAM (DZ) due to the strong orography of the area. As this station is the only one existing in large areas of Northern Africa, it is kept in the evaluation, but all results have always to be treated very carefully due to its specific location. Due to the mountainous characteristics of the location IZA (E) the same restriction applies due to a strong orography inside the forecast grid box applies.

Table 1: Details of ground stations and their data availability.

Name	ID	Code	Lat. (°)	Lon (°)	Elev (m)	Period
DLR/PSA	1	PSA (E)	37.091	−2.358	492	2002–2012
Cabauw (BSRN)	2	CAB (NL)	51.971	4.927	0	2006–2012
Camborne (BSRN)	3	CAM (UK)	50.217	−5.317	88	2002–2006
Carpentras (BSRN)	4	CAR (F)	44.083	5.059	100	2002–2012
Cener (BSRN)	5	CEN (E)	42.816	−1.601	471	2010–2012
Izana (BSRN)	6	IZA (E)	28.309	−16.499	2373	2010–2012
Lerwick (BSRN)	7	LER (UK)	60.133	−1.183	84	2002–2006
Lindenberg (BSRN)	8	LIN (D)	52.210	14.122	125	2002–2006
Maan (EnerMena)	9	MAA (JO)	30.172	35.818	1012	2012
Palaiseau (BSRN)	10	PAL (F)	48.713	2.208	156	2005–2006
Payerne (BSRN)	11	PAY (CH)	46.815	6.944	491	2002–2010
Sede Boquer (BSRN)	12	SBO (IL)	30.905	34.782	500	2003–2011
Tamanrasset (BSRN)	13	TAM (DZ)	22.780	5.510	1385	2003–2012
Tataouine (EnerMENA)	14	TAT (TN)	32.974	10.485	210	2012
Toravere (BSRN)	15	TOR (EE)	58.254	26.462	70	2002–2012

2.4 Satellite-based cloud type information

The APOLLO (AVHRR Processing scheme Over cLouds, Land and Ocean; Kriebel et al., 1989, 2003) algorithm was originally developed to exploit data from the AVHRR sensors aboard the polar orbiting series of NOAA satellites, in order to estimate the properties of clouds. It has been adapted to process images of the SEVIRI (Spinning Enhanced Visible and Infrared Imager) instrument aboard the series of Meteosat Second Generation satellites (APOLLO/SEV). APOLLO/SEV provides quantities related to clouds for each pixel (3 km at nadir) and every 15 min. They include a mask with values being cloud-free or cloudy, cloud optical depth, cloud fraction and cloud types according to the height level of the cloud. Optically thick water or mixed phase clouds are separated in low, medium, or high level clouds. These are separated according to the cloud top temperature. Layer boundaries are set to 700 hPa and 400 hPa and associated with the 700 and 400 hPa temperatures taken from standard atmospheres. The thin clouds type exclusively contains thin pure ice phase clouds with no thick clouds underneath.

In a further post-processing several hourly cloud information parameters are created: If more than 1 MSG slots from minutes 15, 30, 45 of the previous and 00 of the actual hour have thin ice clouds, the hour is set as a 'thin ice cloud case'. Otherwise the hour is treated as a 'water/mixed phase' cloud case in the cloudy case. The hour is classified as cloud free if there is no cloud at all found in the 4 slots. This selection scheme is very sensitive to thin ice clouds, but this is intended as we focus our study on direct normal irradiances, which are very sensitive to this cases.

Additionally to the single pixel results, the surrounding 29 × 29 pixel window is evaluated in order to understand the medium scale cloud situation and to discriminate between overcast/broken, scattered and isolated cloud fields. This is done independently from the actual pixel conditions – assuming that a clear pixel can be in a cloud scattered situation affecting the whole radiative field in the vicinity. If there are more than 10 individual cloud elements in the surroundings of a cloudy pixel, the situation is classified as 'scattered' unless the total cloud fraction in the surroundings is above 80 %, which classifies the case as 'broken/overcast'. If there are less than 10 cloud elements, the situation is classified as 'broken/overcast' unless there is a high number of more than 175 cloud/no cloud changes from pixel to pixel in any direction, which again results in a 'scattered' case. Also, any cloud/no cloud change from the central pixel to one of the four direct neighbours always results in a 'scattered' classification.

If there are more than one broken/overcast cloud situations in the respective hour, the case is treated as broken/overcast. Also, if there are more than one scattered cloud situations in the hour, the case is treated as scattered cloud situation. It may happen that an hour is treated both in the broken/overcast class and the scattered cloud class later on if the individual MSG slots indicate both classes within the hour. In such cases a discrimination of the complete hour between broken/overcast and scattered would be artificial, therefore it is not performed.

3 Methodology

Statistical parameters for the comparison of hourly values include biases, root mean square error (RMSE), and linear correlation coefficients following the definitions in Beyer et al. (2008) and Hoyer-Klick et al. (2010). Relative values are normalized by the observational mean values. Both relative and absolute values are reported to allow inter-comparison with other studies. Especially, RMSE is seen as the most important statistical metric for the application of electricity grid management issues. This is reflecting the fact that in this application area small errors are of less importance compared to large errors affecting the electricity grid's security.

The persistence approach used in our case is mostly a two day persistence. This is a specific feature required by the day-ahead electricity market as realised today e.g. in Spain. Typically, any day-ahead forecast has to be provided to the market operator during the late morning hours. This restricts the persistence to the values obtained on the previous day as the best estimate for tomorrow without using any weather forecast.

In order to generate results being comparable against other studies, we also include the standard one-day persistence in the results. This also reflects the expectation that some electricity markets in some countries may be organised with closure times allowing the usage of standard persistence as poor man's forecast. Also, intra-day markets typically allow the use of shorter range forecasts due to either continuous or repeated closure times during the day. Nevertheless, this is not the focus of our paper, having in mind that intra-day markets typically make use of nowcasting schemes either from satellites or from rapid-update cycling numerical weather prediction. Both technologies are not in the focus of this study.

All parameters are available for each forecast hour up to 48 hours, but only daytime values are evaluated following the MESOR standard (Beyer et al., 2008). All results are derived at the location of ground-based observations using the arithmetic average of values at the 4 nearest neighbouring ECMWF grid points.

A compressed evaluation of both biases and RMSEs is favourable for the comparison of different forecasts runs. Besides the usual plots on RMSE over different years or the hour of the day, we apply target plots following ideas as introduced by Jolliff et al. (2009) and currently being widely adapted in the air quality community. It sets bias and centred root mean square errors (CRMSE) in relation to each other. The CRMSE is the standard deviation of the differences (named '$\sigma_{\text{(mod-obs)}}$' in the figures) between the reference and the forecasted irradiance dataset. A movement towards the centre of this plot can be interpreted as an overall improvement in RMSE (given in W/m^2). The distance to the horizontal zero axis reflects the absolute bias in W/m^2. In case the forecasted amplitude (as described by the standard deviation σ_{mod} of all modelled values) is less than the observed amplitude at the station (as described by the standard deviation σ_{obs} of all observed values) the sign on the horizontal axis is chosen as negative, otherwise it is set to be positive.

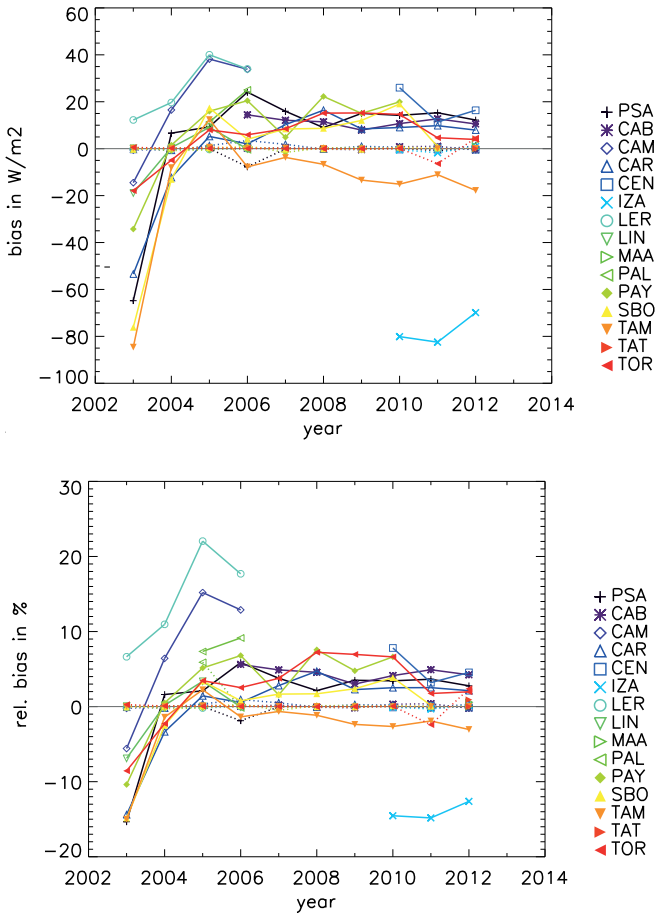

Figure 1: Day ahead hourly GHI forecast verification – annual absolute and relative bias of ECMWF/IFS (line) together with two-day persistence (dotted).

4 Results

4.1 Evaluation of the ECMWF IFS GHI and GHI2DNI for 2003–2012

First of all, a time series of basic statistics over the different years and IFS model versions is computed versus the two day persistence (Figs. 1 to 3 for GHI and 3 to 5 for DNI). Generally, the change in the radiation scheme in 2007 does not result in a major improvement of statistical parameters. For TAM (DZ) a negative bias in GHI is even increased, while for other stations the positive GHI bias remains similar. For DNI, the bias of most stations is even increasing after 2007. There is no improvement visible in the RMSE at the GHI and other statistical parameters do also not indicate a major improvement in DNI based on the radiation code change in 2007 (Morcrette et al., 2008a and b). Only a slight improvement at many stations in the linear correlation coefficient of DNI might be seen.

Fig. 1 supports the finding of several authors being reviewed in Mathiesen et al. (2013) and summarized in the statement that regardless of the model, global irradiance NWP forecasts are generally positively biased. Here, the same is found for all years later than 2005 and with exception of the station TAM (DZ). The results are also very much in line with the 23 W/m^2 of average overestimation found for the Southern Plains Atmospheric Radiation Measurement (ARM) site in the US for 2004–2009 in all-sky conditions (Ahlgrimm and Forbes, 2012). Also, Haiden and Trentmann (2015) report on a general under-prediction of cloudiness for the ECMWF/IFS, which results in an over-prediction of global irradiances most likely. They also report on an improvement in cloud forecast skill in the Alpine region after the introduction of the new prog-

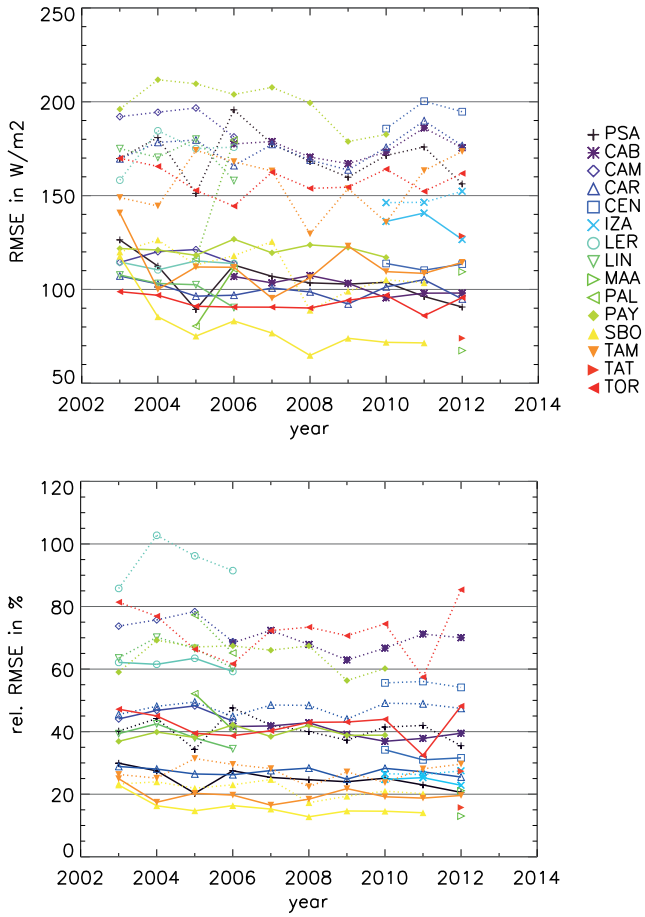

Figure 2: Day ahead hourly GHI forecast verification – annual absolute and relative RMSE of ECMWF/IFS (line) together with two-day persistence (dotted).

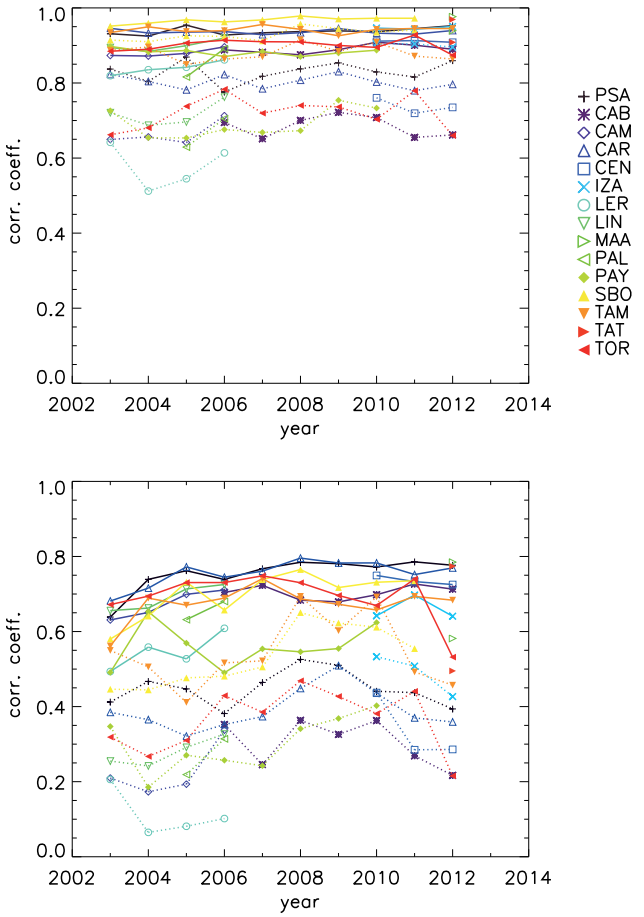

Figure 3: Day ahead hourly GHI (upper panel) and DNI (lower panel) forecast verification – annual correlation coefficients of ECMWF/IFS (line) together with two-day persistence (dotted).

nostic microphysics scheme (Forbes et al., 2011) into the ECMWF/IFS in November 2010. This can be confirmed by a bias reduction in GHI at all stations besides CEN (E) and TAM (DZ) in this study, while for DNI the effect is not visible as clear as for GHI.

For the DNI annual biases, the year 2003 is very eye-catching with low biases down $-250\,W/m^2$ for SBO (IL), PSA (E), TAM (DZ), CAR (F), PAY (CH), LIN (D), TOR (EE), and CAM (UK) (ordered from largest to smaller values). Also, in the GHI verification, a strong underestimation with biases of up to $-80\,W/m^2$ is found. A deeper insight in monthly resolved statistics shows large negative DNI bias values for months before October 2003 for these stations (e.g. Fig. 6 for SBO (IL)). The visual interpretation of hourly time series reveals that especially in clear sky days the maximum DNI value is underestimated systematically. Referring to ECMWF (2013), the new IFS model cycle 26r3 was introduced on 7th October 2003 – including the Tegen et al. (1997) aerosol climatology in the model. Having in mind that SBO (IL), PSA (E) and TAM (DZ) are the stations being most affected by dust aerosols, it is not surprising that this change occurs (see also Rodwell and Jung, 2008). Obviously, the change towards cycle 26r3 was positive for the IFS's verification results.

Nevertheless, for TAM (DZ) mostly clear sky cases are still underestimated – a result pointing into the direction of deficiencies in the aerosol climatology. Also, both for GHI and DNI the station IZA (E) shows large negative biases for all years. Anyhow, perfect agreement with station observations and model runs at coarse resolution cannot be expected especially for stations such as TAM (DZ) or IZA (E) as discussed above. Also for both, GHI and DNI at the station LER (UK), low correlation coefficients and large relative RMSE values are found for all years. A visual interpretation reveals that cloudy days are often forecasted as typical clear sky days. Partly, those days show a reduced GHI, but with values being still larger than the observations. Also, CAM (UK) has large positive biases for 2004 to 2006. Visual interpretation of time series also reveals that cloudy cases with low and medium height clouds are often forecasted as cloud free. The station PAY (CH) shows a rather low correlation coefficient. Also it is found, that clear sky days are often forecasted as cloudy, while cloudy days are forecasted as clear sky. While other stations show

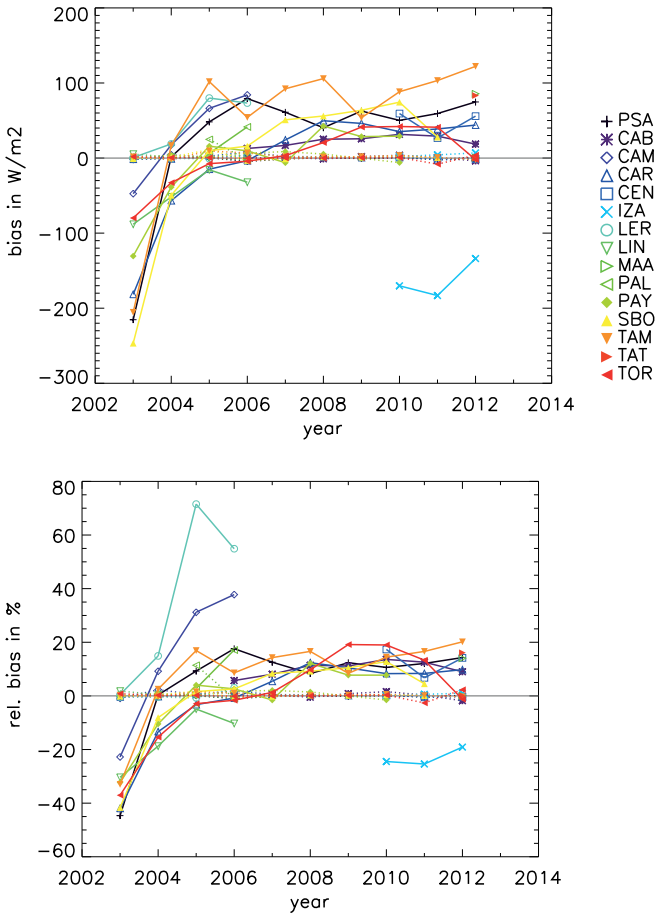

Figure 4: Day ahead hourly DNI forecast verification – annual absolute and relative bias of ECMWF/IFS (line) together with two-day persistence (dotted).

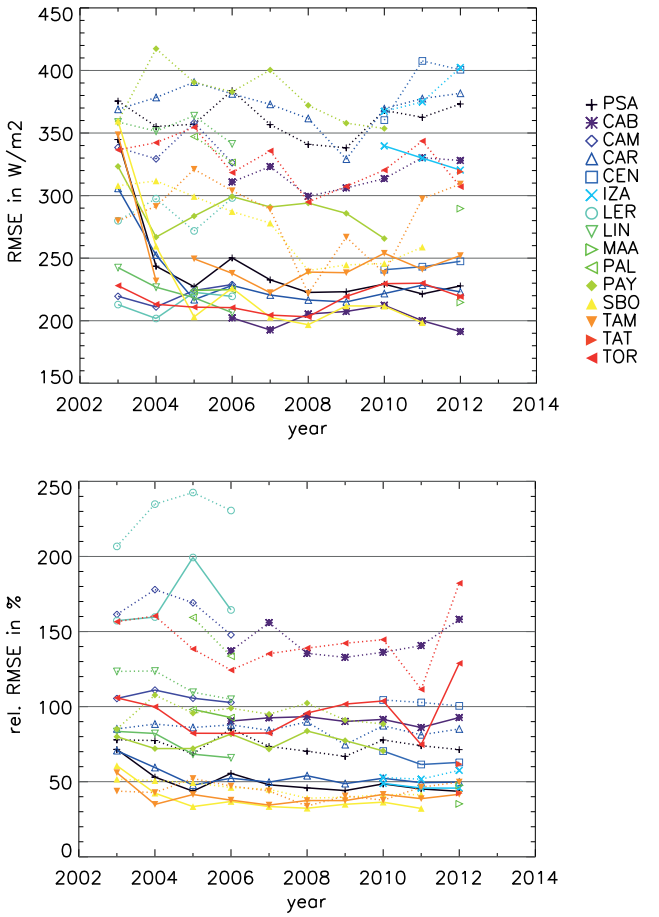

Figure 5: Day ahead hourly DNI forecast verification – annual absolute and relative RMSE of ECMWF/IFS (line) together with two-day persistence (dotted).

discrepancies in situations with quickly varying irradiance values from hour to hour, PAY (CH) is a station with many complete forecast failures of the overall situation in a day. In case of solar technologies with thermal or battery storages, this error characteristic is very important as the overall daily sum of irradiation is also badly forecasted in such conditions.

Overall, Figs. 1 to 6 are showing two-day persistence forecasts as reference as discussed above and reflecting today's day-ahead electricity market requirements. Generally, it can be noted, that the bias of the two-day persistence approach is much smaller compared to ECMWF forecast biases, but both RMSE and correlation coefficients are much worse than those of the ECMWF forecast. So, overall, ECMWF has a positive forecast skill for all stations and both GHI and DNI.

Besides today's day-ahead electricity market requirements, the day zero forecast verification is also of interest for intra-day trading. Also, the internal requirements of power plant operators and electricity grid operators may ask for day zero forecasts. Therefore, we concentrate now on the verification of day-ahead forecasts versus the actual day forecasts ('day zero') as shown in Figs. 7 and 8. Each forecast run is started at 00 UTC. Therefore, the day zero forecast includes the hours 1 to 24 while the day-ahead forecast includes hours 25 to 48. Lines indicate the results for the day-ahead verification results, while dotted lines give the results for the day zero verification. Both verifications are made against the persistence, being a 2-day persistence in the day ahead approach and the classical 1-day persistence in the day zero verification.

Biases in GHI seem to be rather similar, with small differences in both directions for some years and stations. The GHI's RMSE of day zero is systematically below the RMSE of the day ahead as expected. For SBO (IL) and TAM (DZ) the relative RMSE of day zero is nearly the same than the day ahead or smaller by few percent only. For other stations differences of up to 10 % are reached frequently, while both EnerMENA stations Maan and TAT (TN) show a clear difference of more than 20 %. Perez et al. (2013) found a 'composite' all site averaged relative RMSE of 42 % for day zero (their 'day 1') and of 44 % for the day ahead in 2007/2008 for Central Europe. Compared to Fig. 7 this is the maximum value found in our analysis – being found only in TOR (EE), CAB (NL) and PAY (CH) being also Central or Northern European stations. For Spain, they found values of 22 and 24 % which is in line with our results for the PSA (E) station.

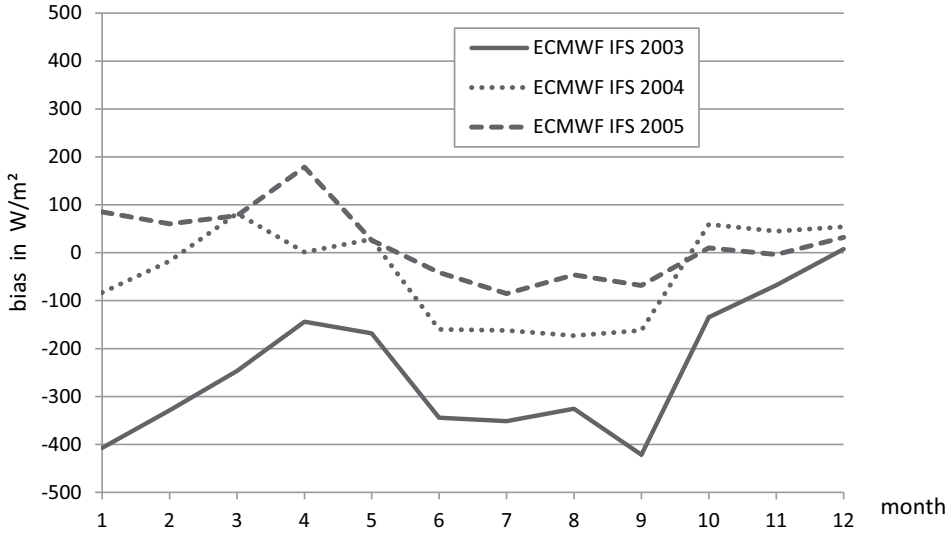

Figure 6: Day ahead hourly DNI forecast verification – monthly resolved absolute bias of ECMWF/IFS for the station SBO (IL) for 2003, 2004, and 2005.

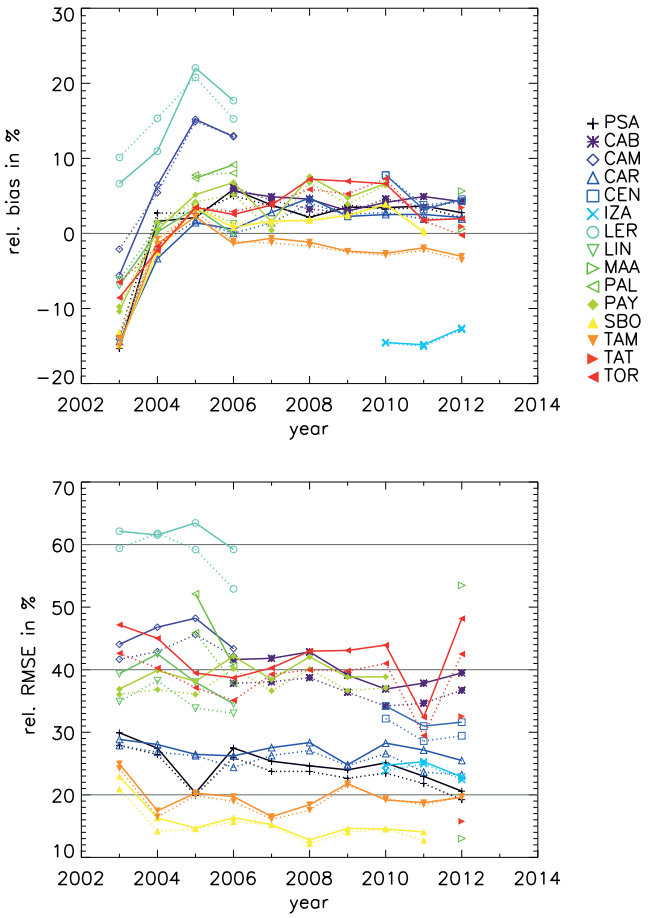

Figure 7: Day ahead (lines) and day zero (dotted) hourly ECMWF/ IFS GHI forecast verification – annual relative bias and RMSE.

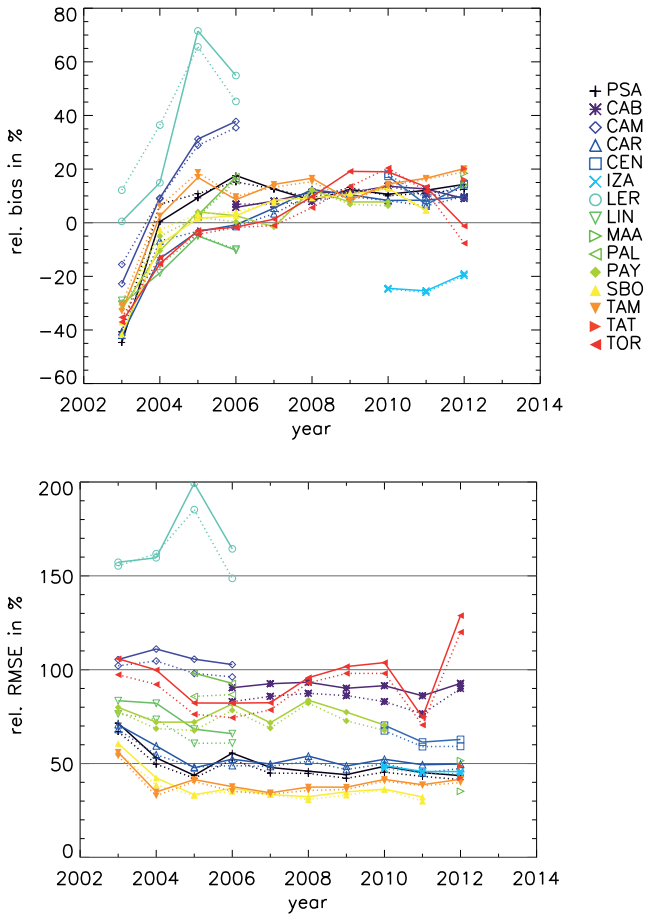

Figure 8: Day ahead (lines) and day zero (dotted) hourly ECMWF/ IFS DNI forecast verification – annual relative bias and RMSE.

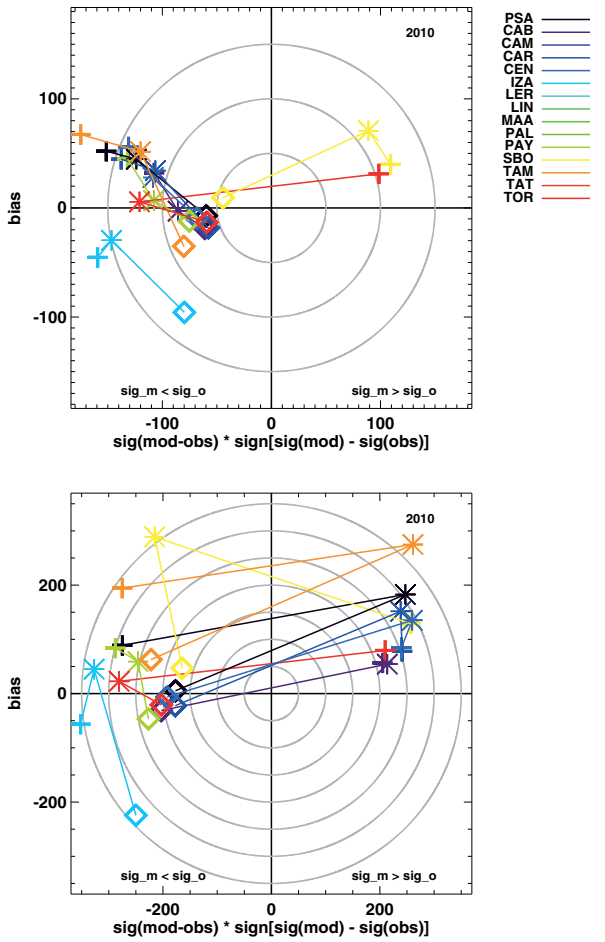

Figure 9: Day ahead hourly ECMWF/IFS GHI (upper) and DNI (lower panel) forecast verification separated into classes 'cloud free' (◊), 'other water/mixed phase clouds' (+), and 'thin ice clouds' (*). Please note changing axis scales.

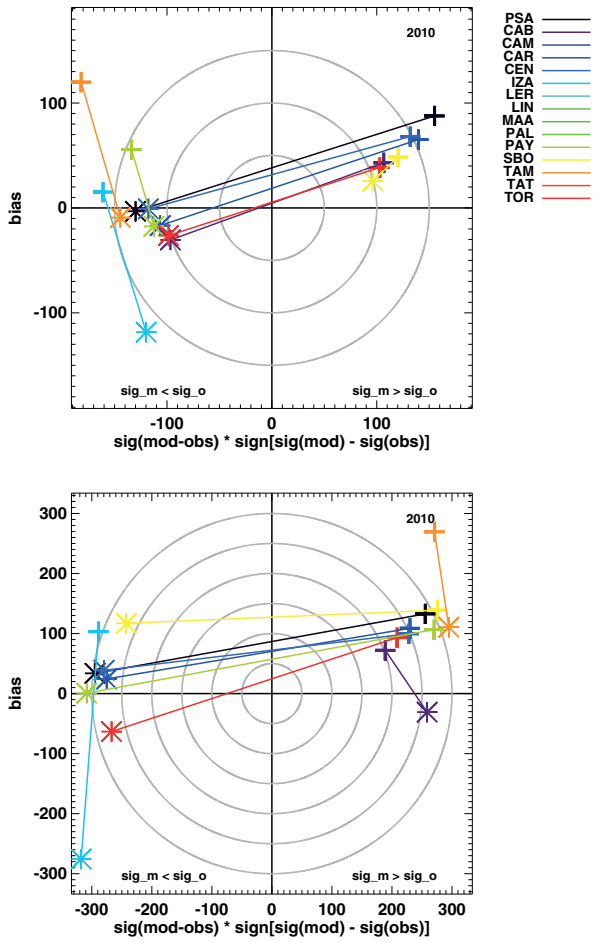

Figure 10: Day ahead hourly ECMWF/IFS GHI (upper) and DNI (lower panel) forecast verification for water/mixed phase clouds being separated into classes 'overcast/broken' (+) and 'scattered clouds' (*). Please note changing axis scales.

For DNI biases the same is found – small variations in both directions for some years and stations, but no clear trend. The DNI's RMSE of day zero is systematically below the day ahead RMSE, but the differences are smaller than for the GHI case. Differences in relative RMSE up to 5 % are found for most years and stations and reach approximately 10 % for the EnerMENA stations.

Based on APOLLO cloud types a discrimination between cloud free, thin ice clouds and other cloud cases can be made. Fig. 9 gives the result for the year 2010 – the year with the most recent radiation scheme being in place and the most ground stations being available. As we have not seen any dedicated differences between the years 2008 to 2012 in the above verification, we focus on a single year from now on – being either 2010 with the maximum of available ground observations or 2012 as having the experimental MACC-II runs available.

Cloud free situations tend to have small, but negative biases in the GHI – with exception of the station IZA (E), which is probably due to the mountainous location of the ground station compared to the area averaged forecast value. Also, TAM (DZ) shows a larger negative bias for cloud free situations.

Several stations (TOR (EE), PAY (CH), CAB (NL), CAR (F), and CEN (E)) show low biases around zero also for the thin ice cloud cases, while other stations as PSA (E), SBO (IL), and TAM (DZ) show large positive biases for thin ice clouds. For all stations, the biases are large for the water and mixed phase cloud conditions. This is in line with the results of Ronzio et al. (2013) who found for 2011 the worst GHI forecast performance for mid and high level thick clouds for three Italian stations and all forecasts being assessed (IFS/ECMWF, GFS/NCEP, LAMI, RAMS). Doing the same analysis for 2011, we find this behavior for all stations, but TOR (EE) and SBO (IL), where the RMSE for thin ice clouds is slightly larger than for other clouds. This can be seen in Fig. 10 if remembering that in target plots the RMSE is visible as distance to the center.

For most stations, the forecasted amplitude is below the observed amplitude of irradiance values, resulting in symbols being located on the left hand side of the target plot. Thin ice clouds for the PSA (E) station and

water/mixed-phase clouds for TOR (EE) and SBO (IL) are the only cases where the forecasted amplitude is higher than the observed amplitude.

For DNI, cloud free situations are characterized by low biases, but still large RMSE values (Fig. 9 – lower panel). IZA (E) is an exemption with its specific location as discussed above. Besides PAY (CH) and IZA (E), all other stations show the largest RMSE values for thin ice clouds. Biases are also highest for thin ice clouds at the stations CAR (F), CEN (E), PSA (E), SBO (IL), and TAM (DZ). For CAB (NL), the biases of both cloud classes are similar, while for PAY (CH) and TOR (EE) the water/mixed-phase class is biased more.

The forecast tends to overestimate the DNI in thin ice cloud conditions. The visual interpretation of time series reveals frequent cases with typical clear sky DNI patterns in the forecast, while the observed DNI is varying around 30 to 70 % values of the typical clear sky DNI. This indicates frequent cases where the existence of thin ice clouds is not forecasted or their optical depth is much too low.

There is no clear finding about the modelled amplitude versus the observed amplitude for the different cloud classes. It can be noted that for most stations (besides PAY (CH) and IZA (E)) the two cloud classes are not on the same half of the plot – indicating differences in the modelled amplitude in the two cloud classes, but the results are heterogeneous for all stations. Some thin ice cloud symbols can be found in the right half of the target plot, others on the left half. The same applies for water/mixed phase clouds.

Based on APOLLO cloud masks, also a distinction between overcast/broken clouds and scattered clouds is made. Broken clouds are understood as large cloud systems with small gaps in between – being closer to overcast conditions than to scattered cloud systems. Scattered cloud systems on the other hand are typical shallow cumulus cloud systems. Fig. 10 show all cases having a water or mixed phase cloud as the ground station's pixel and broken/overcast (+) or scattered (*) cloud conditions in the surroundings of 29×29 pixels.

For GHI, all stations (besides IZA (E)) show clearly larger RMSE for broken/overcast conditions than for scattered cloud situations. This is not obvious and against many user expectations assuming scattered cloud fields (as e.g. cumulus clouds) as the reason for large RMSE values. Actually it is found, that broken/overcast cloud conditions always (with exception of IZA (E)) show larger biases in GHI and therefore the overall larger RMSE values in broken/overcast conditions is caused. This is in line with Ahlgrimm and Forbes (2012) reporting that deep and low classified clouds 'explain just over a quarter of the accumulated bias'. The findings of Haiden and Trentmann (2015) on the seasonality with low winter time forecast skills in stratus cloud situations can be seen in our study e.g. at CAB (NL) and PAY (CH) as well.

For DNI, the situation is different – scattered clouds result in larger RMSE values for all stations besides SBO (IL) and TAM (DZ), but the differences are not as strong as for GHI. Nevertheless, for all stations (besides IZA (E)) the scattered cloud situations have smaller biases in DNI as found for GHI. This is pointing to the fact that the larger RMSE of the scattered cloud class is driven by a larger scatter between forecasts and observations – as being expected for scattered cloud fields.

Also, it is remarkable, that for PSA (E), CAB (NL), CAR (F), CEN (E), SBO (IL), and TOR (EE), the modeled variability in GHI is larger than the observed for broken/overcast situations, while this is never the case for scattered cloud conditions for any of the stations.

For several stations a different behavior in the forecasted amplitude of both GHI and DNI values versus the observed amplitude is found for the two cloud classes – resulting in one station's mark in the target plot on the one side and the other mark on the other side of the target plot. For GHI, the stations PSA (E), CAB (NL), CAR (F), CEN (E), SBO (IL) and TOR (EE) have larger forecasted amplitudes than observed. For DNI, the stations PSA (E), CAR (F), CEN (E), SBO (IL), and TOR (EE) show the same effect – broken/overcast conditions have larger forecasted amplitudes than observed, while scattered clouds have smaller forecasted amplitudes than observed. Actually, finding that scattered clouds have smaller forecasted GHI amplitudes is not surprising as all three-dimensional scattering at clouds is not treated in the IFS radiation scheme and is known to result in larger irradiance values. But the fact that broken/overcast clouds have larger forecasted amplitudes than observed is not obvious.

4.2 Evaluation of ECMWF IFS DIR2DNI vs. the IFS GHI2DNI

The empirical transformation from GHI to DNI based on Skartveit et al. (1998) has been discussed frequently as being derived only from the station in Bergen, Norway, and being perhaps only valid for the northern regions. Therefore, having the announcement of direct irradiances as explicit model output parameter was very much welcomed by the solar energy community.

Fig. 11 provides the comparison of both day ahead approaches (GHI2DNI with stars, DIR2DNI with diamonds) vs. the two day persistence (+ symbols) for the nine stations available in 2012. A clear shift of the bias is seen. The GHI2DNI approach results in mainly positive biases (with exception of TOR (EE)), but the use of the DIR2DNI also results in biases, only with the opposite sign. For some stations the bias is reduced (CEN (E), PSA (E), MAA (JO), TAT (TN), TAM (DZ)). It is very favorable that especially the locations being most relevant for concentrating technologies as in Spain, Jordan and Northern Africa are being improved by the new direct irradiance output. But for the more Northern stations it is even worse in absolute values (CAB (NL), CAR (F), TOR (EE)). The RMSE is reduced for CEN (E), MAA (JO), Tatouine,

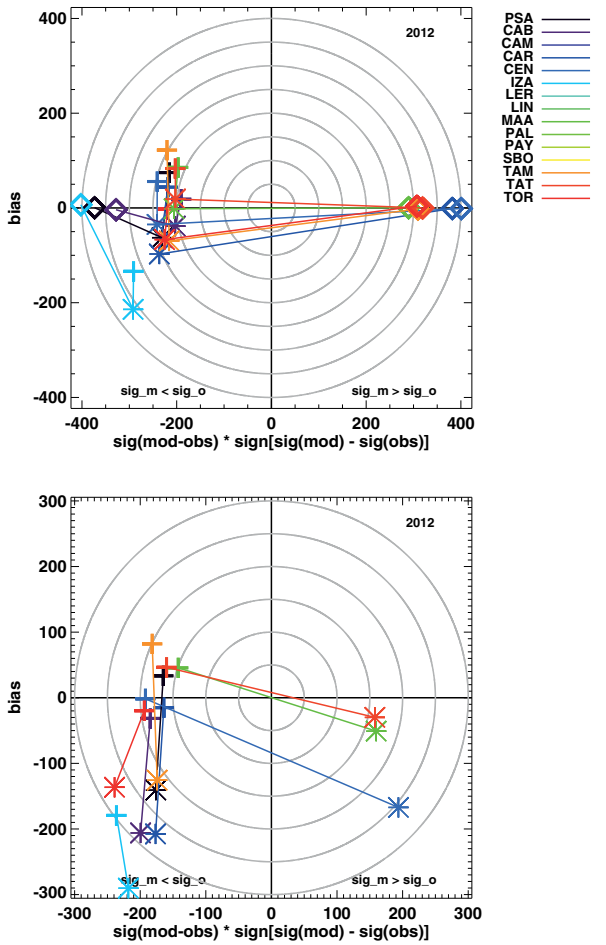

Figure 11: Day ahead hourly ECMWF/IFS DNI forecast verification. The upper panel shows all cases for the two day persistence (◊), 'GHI2DNI' (+), and 'DIR2DNI' (*), while the lower panel includes only cloud free cases for 'GHI2DNI' (+) and 'DIR2DNI' (*). Please note changing axis scales.

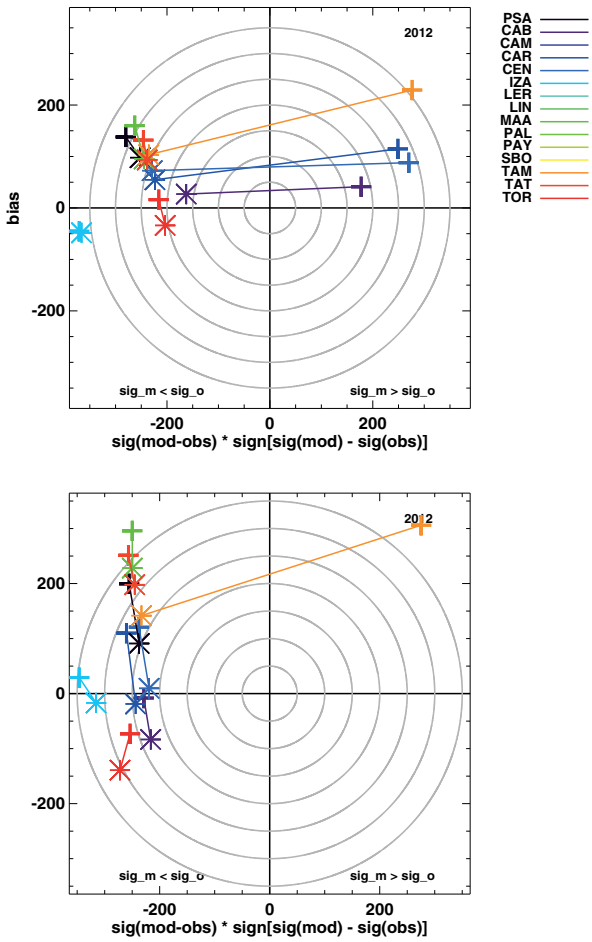

Figure 12: Day ahead hourly ECMWF/IFS DNI forecast verification in water/mixed phase (upper) and thin ice cloud (lower panel) cases for 'GHI2DNI' (+) and 'DIR2DNI' (*) approaches.

and TAM (DZ), but increased for the other stations. The same effect is found for the day zero – the RMSE values are generally smaller on day zero, but the bias shifting remains the same.

A further analysis reveals that cloudy cases are forecasted much better by the DIR2DNI approach, but large biases are introduced in clear sky cases (Fig. 11 lower panel and Fig. 12, + for GHI2DNI and * for DIR2DNI). For all stations the biases are negative for the DIR2DNI approach and being larger (with exception of MAA (JO) and TAT (TN)) in their absolute value for the clear sky cases. For thin ice cloud cases a strong reduction in biases and RMSE is found for all stations besides CAB (NL) and TOR (EE) (where negative biases turn more negative). For water/mixed phase clouds also a reduction in biases and RMSE is found for all stations besides TOR (EE) - for TOR (EE) the RMSE is also reduced, but the bias turns from a small positive to a slightly larger negative bias.

4.3 Evaluation of MACC-II GHI and DIR2DNI in 3-hourly vs. hourly output resolution

In the previous chapter the standard ECMWF IFS model runs have been assessed. Previous work (e.g. Perez et al., 2013) has shown that ECMWF is performing very well compared to spatially higher resolved models. But it is still needed to post-process ECMWF's model output from the 3-hourly standard output frequency to an hourly forecast. Therefore, we evaluate further the question if a higher, hourly temporal output resolution of the forecast would be beneficial.

This section compares forecasts from the experimental run 'g3q4' which outputs DIR and GHI hourly against results from the control run (experiment name 'g3rs') for which DIR and GHI are output three-hourly and interpolated to hourly. It has to be noted that the internal model resolution in time and space is kept constant. Results are compared against the control run (experiment name 'g3rs').

Overall, only a small impact is found (Table 2). For GHI, biases are slightly smaller for CAR (F), IZA (E),

Table 2: Day ahead hourly GHI and DNI forecast verification for three-hourly forecasts interpolated to hourly (INT, from exeriment 'g3rs') and for hourly forecasts (OUT, from experiment 'g3q4').

station	GHI	GHI	DNI	DNI
	Bias [W/m^2] INT/OUT	RMSE [W/m^2] INT/OUT	Bias [W/m^2] INT/OUT	RMSE [W/m^2] INT/OUT
PSA (E)	−2/0	94/96	−86/−82	256/258
CAB (NL)	14/14	107/111	−34/−31	217/219
CAR (F)	−5/−3	97/100	−120/−116	267/268
CEN (E)	8/10	118/120	−58/−55	259/260
IZA (E)	−115/−112	190/188	−315/−312	487/489
MAA (JO)	−25/−22	81/82	−61/−57	234/238
TAM (DZ)	−25/−21	117/117	−82/−76	231/234
TAT (TN)	−3/−1	78/80	4/7	212/217
TOR (EE)	6/6	98/101	−67/−64	239/239

MAA (JO), PSA (E), TAM (DZ), and TAT (TN), but nearly unchanged in CAB (NL) and TOR (EE) and slightly larger at the CEN (E) station. On the other side, RMSE values do slightly increase. For DNI, negative biases tend to be smaller in the hourly resolved run, but RMSE remains the same or even slightly increases. Overall, a finer temporal resolution output seems to be not needed for irradiance forecasts as long as the spatial resolution remains coarse as in global models. This is probably in line with findings of e.g. Girodo (2006) that an increase in spatial resolution does not necessarily improve the forecast's RMSE. This is generally being interpreted as the lack of today's cloud parameterization schemes to accurately describe individual clouds. Here we illustrate that this even holds for the 1 hour vs. the 3 hourly temporal output resolutions.

Nevertheless, the study was also motivated by the expectation that the hourly output resolution might be helpful to cover the strong temporal irradiance gradient in morning and afternoon hours better. Overall, the impact is rather small while the results for GHI forecasts show a slightly larger positive impact than those for DNI (Figs. 13 and 14, interpolated forecasts are dotted, hourly model output in lines). For GHI the improvement in morning and afternoon hours can be found. Positive biases are reduced in CAB (NL), TOR (EE), CEN (E), CAR (F) and PSA (E) in the morning. In the afternoon, relative bias values have lower values, but as for most stations they are already negative in the g3rs run, the absolute value of the bias is increased. Only in TAM (DZ) improved afternoon hours are found as well. Relative RMSE values are reduced in morning hours for some stations, but increase at noon and in afternoon hours.

Contrary to the GHI findings, the DNI RMSE shows nearly no changes in morning or afternoon hours for both forecast days (Fig. 14). The 3-hourly-interpolated forecast even shows smaller relative RMSE for TOR (EE) during noon and early afternoon hours – an effect being slightly visible also in CAB (NL) and CEN (E). At other stations there are nearly no differences. Relative biases are mostly slightly reduced for the hourly compared to the three-hourly forecast. For most stations this is favorable as the three-hourly forecast has a strong negative bias. But for CAB (NL), TAT (TN) and TOR (EE) (on the day ahead) their positive biases are enlarged as well during noon hours.

4.4 Evaluation of MACC-II GHI and DIR2DNI based on an aerosol climatology vs. interactive prognostic aerosols

In Sections 4.1 and 4.2 standard ECMWF IFS model runs have been assessed. We have seen that cloudy cases are treated better if an explicit direct irradiance output is used, but clear sky biases have been introduced. Recent developments at ECMWF in GEMS and MACC/MACC-II projects have introduced an advanced aerosol scheme. Therefore, we evaluate further the question if an explicit aerosol scheme is advantageous. This section is based on the experimental run 'g3o4' using the new aerosol scheme, explicit direct irradiance output and also an hourly output frequency. Results are compared against the hourly resolved 'old-aerosol' run with also an explicit direct irradiance output (experiment name 'g3q4', as evaluated in Section 4.3). It has been shown in Section 4.3 that the impact of hourly vs. 3-hourly output frequency is negligible – therefore, results as given in Tables 3 and 4 should be applicable also to the operationally used 3-hourly output frequency.

It has to be noted, that the aerosol-radiation interaction has been switched on only for the direct aerosol effect, as the parameterization for indirect effects is still not in its final stage. Direct effects are understood as the absorption and scattering by the aerosols themselves, whereas indirect effects include changes in cloud droplet size and hence cloud reflectivity due to the changed quantity of aerosols as cloud condensation nuclei, as well as changes in cloud lifetime due to changes in the droplet size (Haywood and Boucher, 2000). These direct aerosol effects also include the coupling of the aerosol layer and dust mobilization due to changes in the surface temperature. This effect typically results in variations in the thermal stability in the boundary layer and therefore the wind speed close to the ground.

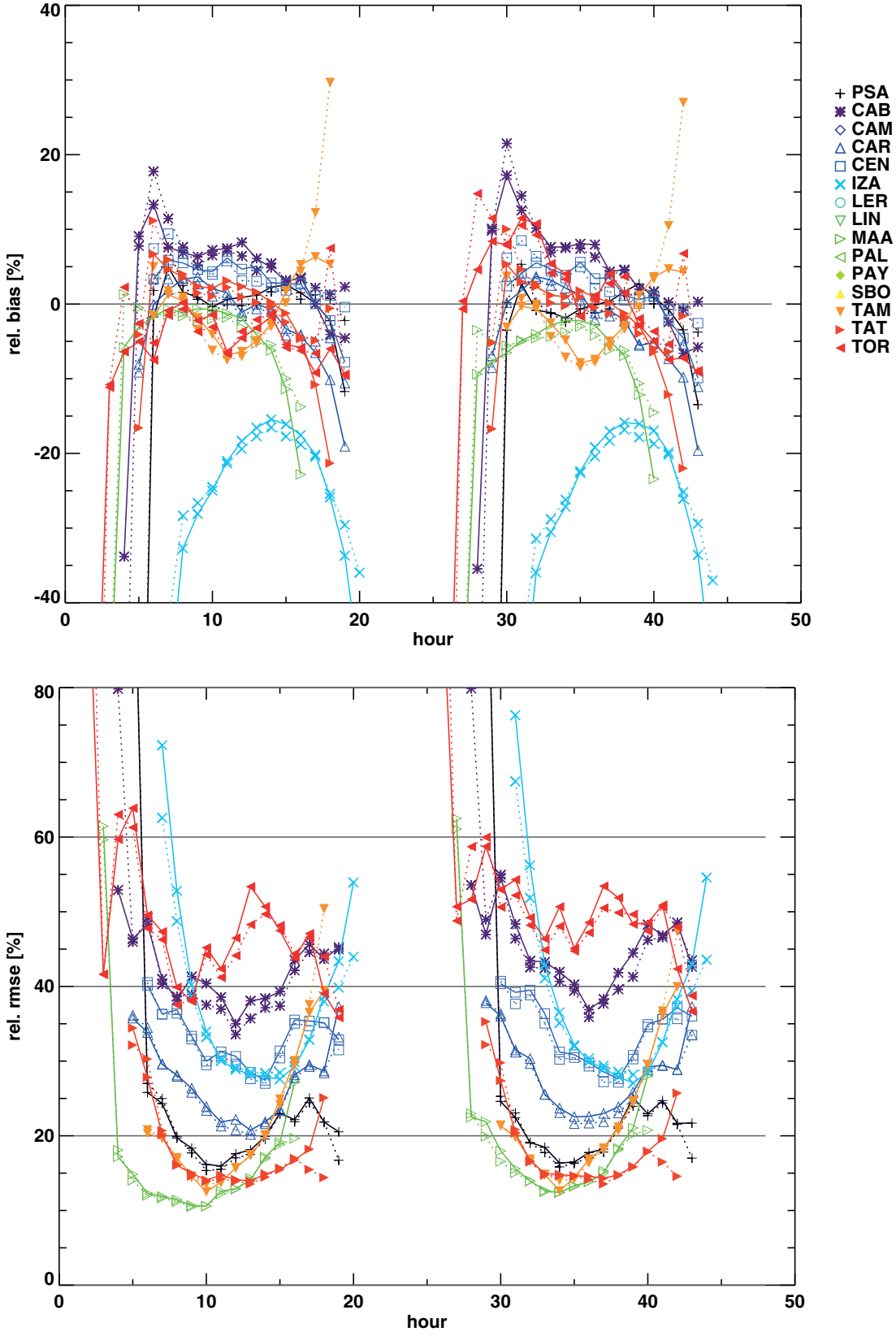

Figure 13: Relative RMSE and biases for day zero (< 25 h) and day ahead (25–48 h) hourly GHI forecasts – both for 3-hourly forecasts interpolated to hourly forecasts (dotted, control run) and forecasts with hourly direct model output (line, experiment run).

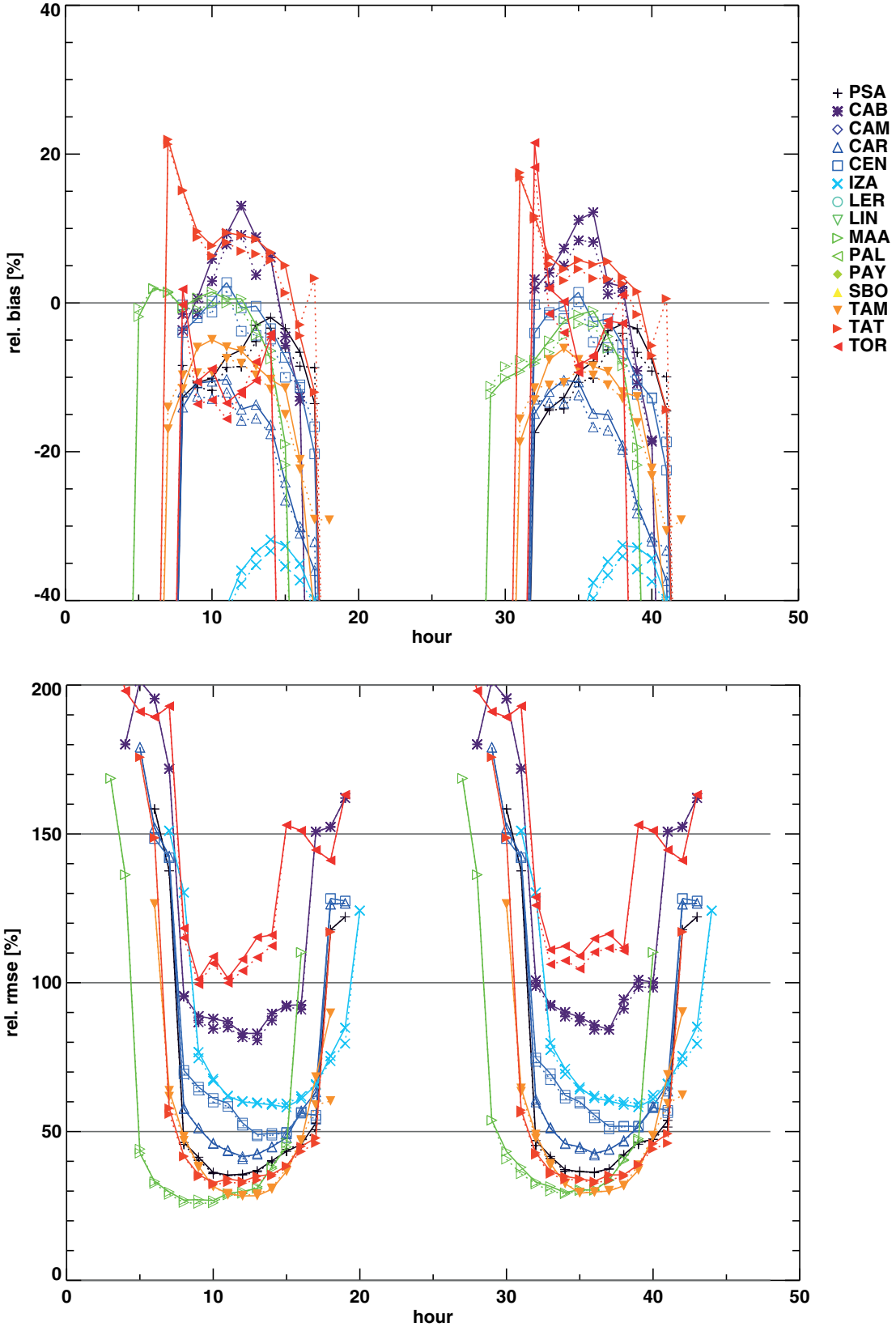

Figure 14: Relative RMSE and biases for day zero (< 25 h) and day ahead (25–48 h) hourly DNI forecasts – both for 3-hourly forecasts interpolated to hourly forecasts (dotted, control run) and forecasts with hourly direct model output (line, experiment run).

Table 3: Day ahead hourly GHI forecast verification with the control run 'g3q4' using an aerosol climatology and the new prognostic aerosol run 'g3o4' for all skies, cloud free and water/mixed phase cloudy cases. Values being >4 W/m^2 smaller are marked in bold.

Station	Bias [W/m^2]			RMSE [W/m^2]		
	all control/new	cloud free control/new	cloudy control/new	all control/new	cloud free control/new	cloudy control/new
PSA (E)	0/4	−23/−**18**	58/58	96/97	63/62	156/157
CAB (NL)	**14**/26	−30/−**16**	**37**/47	110/113	65/**59**	**121**/126
CAR (F)	−**3**/10	−29/−**13**	**43**/54	100/102	63/60	141/145
CEN (E)	**9**/14	−38/−**33**	51/54	120/122	82/82	149/151
IZA (E)	−112/−111	−143/−142	−42/−40	188/183	185/**178**	201/199
MAA (JO)	−22/−**13**	−36/−**22**	7/7	82/**76**	60/**53**	123/**116**
TAM (DZ)	−**21**/−28	−54/−58	70/**58**	**117**/122	**93**/103	179/**173**
TAT (TN)	−**1**/11	−19/−**4**	34/38	80/79	48/46	122/122
TOR (EE)	**6**/13	−20/−**16**	**19**/26	101/102	65/63	106/108

Table 4: Day ahead hourly DNI forecast verification with the control run 'g3q4' using an aerosol climatology and the new prognostic aerosol run 'g3o4' for all skies, cloud free and water/mixed phase cloudy cases. Values being >4 W/m^2 smaller are marked in bold.

Station	Bias [W/m^2]			RMSE [W/m^2]		
	all control/new	cloud free control/new	cloudy control/new	all control/new	cloud free control/new	cloudy control/new
PSA (E)	−82/−67	−169/−**151**	**107**/112	258/**248**	251/**234**	279/279
CAB (NL)	−31/−9	−211/−182	38/56	219/218	293/272	182 /191
CAR (F)	−116/−**79**	−226/−**181**	**44**/66	268/**253**	290/**260**	230/239
CEN (E)	−55/−**40**	−211/−**192**	**75**/85	260/259	292/**282**	239/243
IZA (E)	−312/−314	−417/−421	−78/−75	489/**479**	511/**497**	456/453
MAA (JO)	−57/−**38**	−112/−**84**	49/51	238/**214**	215/**187**	281/263
TAM (DZ)	−**76**/−102	−**133**/−157	106/**79**	234/**225**	226/222	265/236
TAT (TN)	**7**/39	−47/−**9**	**94**/119	217/**204**	178/**157**	272/267
TOR (EE)	−64/−**51**	−146/−**138**	−30/−**16**	239/**242**	278/277	213/218

For GHI, both biases and the RMSE are the same or even increased by the introduction of the new aerosol scheme (with exception of IZA (E) and MAA (JO)), while for DNI biases are reduced (with exception of IZA (E), TAT (TN) and TAM (DZ)) and for some stations (CAR (F), IZA (E), MAA (JO), PSA (E), TAM (DZ), TAT (TN), and TOR (EE)) also RMSE is reduced (Tables 3 and 4). So, overall, there is no improvement for GHI and only a partial improvement for DNI visible if looking at all forecast conditions.

A further evaluation distinguishes again between cloud free and water/mixed phase clouds in order to assess the new aerosol scheme's impact by separating direct and indirect aerosol effect implicitly. As the new aerosol scheme is affecting the boundary layer, thin ice clouds are not evaluated separately. Therefore, the term 'cloudy' in Tables 3 and 4 implies a classification as water/mixed phase cloud case. With exception of TAM (DZ) an improvement for clear sky cases both in GHI and DNI can be seen, as expected from the inclusion of the aerosol direct effect. For water/mixed phase clouds both increased or rather similar biases and RMSE values are found in both GHI and DNI. Values being different less than 5 W/m^2 are assumed as to 'similar'. Overall, the introduction of the aerosol scheme results in an improvement in the direct effect of aerosols on the irradiance parameters, but this is reduced for DNI and even reverted for GHI possibly due to the lack of a proper parameterization for the indirect aerosols effect in the water/mixed phase cloud cases.

5 Conclusions

This study assesses the performance of solar surface irradiance 48 hour forecasts as provided by the ECWMF IFS operational system starting the forecasts at 00 UTC and additional experimental runs with enhanced aerosol representation and increased temporal output resolution. A focus is laid on direct normal irradiance forecasts being needed by concentrating solar technologies and poorly evaluated so far. Nevertheless, global irradiance forecasts are also evaluated having non-concentrating photovoltaic technologies in mind.

Satellite-based cloud type information allow the assessment of performances in different cloud conditions in a direct manner – without needing any inverse assumption about which irradiance patterns are reflecting which cloud conditions.

It was found that the inclusion of the new aerosol climatology in October 2003 improved both the GHI and DNI forecasts remarkably, while the change towards

a new radiation scheme in 2007 had only minor and partly unfavorable impacts on the performance indicators. Following Morcrette et al. (2008b) the change towards the McRad scheme was motivated mainly by improvements in the radiative heating rate vertical profiles and therefore the cloud/radiation interaction. Impacts of McRad have been assessed by monitoring the behavior of meteorological parameters as geopotential, temperature, humidity, horizontal wind and vertical velocity in different heights, longwave cloud forcing as seen from satellites, or sea surface temperatures; but not against ground based surface irradiance measurements as in this study.

Here, we evaluate both two- and one-day persistence forecasts as reference. Two-day persistence forecasts are needed due to timelines in day-ahead electricity grid markets with closure times during the late morning/noon as e.g. in Spain or Germany. Nevertheless, results are given for both in order to ensure the comparability to other studies, to allow the assessment of global numerical weather prediction for the intra-day markets.

Cloud free situations show small, but negative biases in GHI, while positive biases are larger for water and mixed phase clouds compared to thin ice clouds. For DNI, cloud free situations show low biases, but still large RMSE values. Largest bias and RMSE values are found for thin ice clouds which are most likely not being forecasted in all cases.

For GHI, larger RMSE values are found for broken/overcast conditions than for scattered cloud fields. This is caused by large GHI biases in broken/overcast situations. For DNI, the findings are opposite with larger RMSE values and smaller biases for scattered clouds compared to overcast/broken cloud situations.

Generally, both GHI and DNI are forecasted with a larger variability than found in observations for broken/overcast situations, while in scattered situations the variability being forecasted is smaller than observed. The latter is expected for GHI due to the three-dimensional cloud effects resulting in overshooting irradiances of GHI with GHI being larger than clear sky values. But the larger modeled variability in broken/overcast situations can not be explained without further modeling experiments.

The introduction of direct irradiances as an output parameter in the operational IFS version has favorable impact on the accuracy in regions as Spain, Jordan and Northern Africa. This is very important as these are the regions where concentrating solar technologies are most likely to be deployed. Nevertheless, in more Northern conditions this new direct irradiance output did not result in a general performance improvement compared to the widely used Skartveit et al. (1998) global to direct irradiance conversion scheme. Cloudy situations and especially thin ice cloud cases are forecasted much better, but large biases are introduced in clear sky cases. This might affect any direct to diffuse conversion schemes applied for photovoltaic system modeling.

If applying the MACC aerosol scheme in experimental IFS runs in hourly temporal output resolution, an improvement in DNI biases is found for most stations. But overall, the impact of the hourly resolution is low both in GHI and DNI. Also, an improvement in the strong temporal gradients during morning and afternoon can only partially be seen. The positive effect of spatial averaging compared to any modelling in higher spatial resolution has been discussed by other authors for irradiances already. For the temporal resolution, the same effect has now been shown at least on the global modeling scale.

When applying an interactive prognostic aerosol scheme, GHI and DNI biases are reduced especially in cloud free cases being dominated by direct aerosol effects. On the other hand, for the water/mixed phase clouds larger biases and RMSE values are found, even if there is no direct aerosol cloud interaction implemented. But in the run with prognostic aerosols, they are allowed to interact with the radiation in a direct way. This is different from specifying aerosol climatologies which are static and do not change daily in response to the other meteorological fields. Due to this, even if the effect of aerosols on clouds is not explicitly taken into account, a modification of the radiation field due to the interactive aerosols will result in a change in temperature and to a lesser degree other atmospheric fields. This, in turn, will affect the cloud fields which also respond to local and large scale meteorological forcings.

Overall, this indicates the need to implement also the indirect aerosol effects, while the direct aerosol effects and the interaction with changed dust mobilization due to changes in thermal boundary layer stability are already shown to be beneficial.

Overall, the coupling of the operational ECMWF IFS with an aerosol scheme as the GEMS scheme used currently in the MACC reanalysis or any other further enhanced aerosol scheme is advantageous for the DNI as we see smaller biases. But still with this addition, there are still drawbacks due to the decrease in forecast performance for cloudy situations which points to the need for further research in indirect aerosol-radiation effects. Especially, for users from the photovoltaic sector with their usage of GHI only, the negative impact on water/mixed phase cloud cases in both DNI and GHI and the overall GHI accuracy is seen as critical.

Acknowledgements

We thank the numerous operators of the BSRN network stations for providing their measurements. Also, we would like to acknowledge the efforts made by the enerMENA project funded by the German Federal Foreign Office and the partners at University of Jordan and Tunisian Research and Technology Center of Energy. Additionally, we thank our colleague S. Wilbert for providing PSA (E) measurements and J.-J. Morcrette from ECMWF for preparing and running MACC-II experimental runs.

This work has been supported by EU-project Monitoring Atmospheric Composition and Climate (MACC, MACC-II, and MACC-III) under the European Union Seventh Framework Programme, grant agreement number 283576 and the Horizon 2020 Programme, grant agreement number 633080.

References

Ahlgrimm, M., R. Forbes, 2012: The Impact of Low Clouds on Surface Shortwave Radiation in the ECMWF Model. – Mon. Wea. Rev. **140**, 3783–3795. DOI: 10.1175/MWR-D-11-00316.1.

Benedetti, A., J.-J. Morcrette, O. Boucher, A. Dethof, R.J. Engelen, M. Fisher, H. Flentje, N. Huneeus, L. Jones, J.W. Kaiser, S. Kinne, A. Mangold, M. Razinger, A.J. Simmons, M. Suttie, 2009: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation. – J. Geophys. Res. **114**, D13205. DOI: 10.1029/2008JD011115.

Beyer, H.G., J.P. Martinez, M. Suri, J.L. Torres, E. Lorenz, C. Hoyer-Klick, P. Ineichen, 2008: MESOR project deliverable report D 1.1.1 Handbook on Benchmarking, Management and Exploitation of Solar Resource Knowledge. – MESOR CA – Contract No. 038665.

Boucher, O., M. Pham, C. Ventkataraman, 2002: Simulation of the atmospheric sulfur cycle in the LMD GCM: model description, model evaluation, and global and European budgets. – Note **23**, Institut Pierre-Simon Laplace, Paris, 26 pp, http://www.ipsl.jussieu.fr/poles/Modelisation/NotesSciences.htm.

Breitkreuz, H., 2008: Solare Strahlungsvorhersagen für energiewirtschaftliche Anwendungen – Der Einfluss von Aerosolen auf das solare Strahlungsangebot in Europa. – DLR Forschungsbericht 2008-24.

Breitkreuz, H., M. Schroedter-Homscheidt, T. Holzer-Popp, S. Dech, 2009: Short-range direct and diffuse irradiance forecasts for solar energy applications based on aerosol chemical transport and numerical weather modeling. – J. Appl. Meteor. Climate **48**, 1766–1779.

Cesnulyte, V., A.V. Lindfors, M.R.A. Pitkänen, K.E.J. Lehtinen, J.-J. Morcrette, A. Arola, 2014: Comparing ECMWF AOD with AERONET observations at visible and UV wavelengths. Atmos. Chem. Phys., 14, 593–608, doi: 10.5194/acp-14-593-2014.

ECMWF, 2009: IFS Documentation, Cy33r1, Part VI: physical processes.

ECMWF, 2013: Atmospheric model identification numbers. – http://www.ecmwf.int/products/data/technical/model_id/index.html.

Forbes, R., A.M. Tompkins, A. Untch, 2011: A new prognostic bulk microphysics scheme for the IRS. – ECMWF Tech. Memorandum, No **649**, 28 pp.

Fouquart, Y., B. Bonnel, 1980: Computations of solar heating of the earth's atmosphere: a new parameterization. – Beitr. Phys. Atmos. **53**, 35–62.

Geuder, N., N. Janotte, S. Wilbert, 2009: Precise Measurements of Solar Beam Irradiance through Improved Sensor Calibration. – Proceedings of SolarPACES, Berlin, Germany, http://solarpaces2009.org.

Gregory, P.A., L.J. Rikus, J.D. Kepert, 2012: Testing and Diagnosing the Ability of the Bureau of Meteorology's Numerical Weather Prediction Systems to Support Prediction of Solar Energy Production. – J. Appl. Meteor. Climate **51**, 1577–1601. DOI:10175/JAMC-D-10-05027.1.

Girodo, M., 2006: Strahlungsvorhersage auf der Basis numerischer Wettermodelle. – PhD thesis, University of Oldenburg, Germany.

Haiden, T., J. Trentmann, 2015: Verification of cloudiness and radiation forecasts in the greater Alpine region. – Meteorol. Z., **25**, 3–15, DOI:10.1127/metz/2015/0630.

Haywood, I., O. Boucher, 2000: Estimates of the direct and indirect radiative forcing due to tropospheric aerosols: A review. – Rev. Geophys. **38**, 513–542. DOI: 10.1029/1999RG000078.

Hoyer-Klick, C., M. Lefèvre, A. Oumbe, M. Schroedter-Homscheidt, L. Wald, 2010: User's Guide to the SODA and SOLEMI Services, MACC project report D_R-RAD_2.2._1 and _2, version v1.0. – Public report accessible via http://www.gmes-atmosphere.eu.

ISO 9060, 1990: Solar Energy – Specification and Classification of Instruments for Measuring Hemispherical Solar and Direct Solar Radiation. Standard of the International Organization for Standardization (ISO). – International Organization for Standardization.

Jolliff, J.K., J.C. Kindle, I. Shulman, B. Penta, M.A.M. Friedrichs, R. Helber, R.A. , Arnone, 2009: Summary diagrams for coupled hydrodynamic-ecosystem model skill assessment. – J. Marine Syst. **76**, 64–82.

Kaiser, J.W., A. Heil, M.O. Andreae, A. Benedetti, N. Chubarova, L. Jones, J. -J. Morcrette, M. Razinger, M.G. Schultz, M. Suttie, G.R. van der Werf, 2012: Biomass burning emissions estimated with a global fire assimilation system based on observed fire radiative power. – Biogeosci. **9**, 527–554. DOI:10.5194/bg-9-527-2012.

Kraas, B., M. Schroedter-Homscheidt, R. Madlener, 2013: Economic merits of a state-of-the-art concentrating solar power forecasting system for participation in the Spanish electricity market. – Sol. Energy **93**, 244–255.

Kriebel, K.T., R.W. Saunders, G. Gesell, 1989: Optical properties of clouds derived from fully cloudy AVHRR pixels. – Beitr. Phys. Atmos. **3**, 165–171.

Kriebel, K.T., G. Gesell, M. Kästner, H. Mannstein, 2003: The cloud analysis tool APOLLO: Improvements and validations. – Int. J. Remote Sens. **12**, 2389–2408. DOI: 10.1080/01431160210163065.

Lara-Fanego, V., J.A. Ruiz-Arias, D. Pozo-Vazquez, F.J. Santos-Alamillos, J. Tovar-Pescador, 2012: Evaluation of the WRF model solar irradiance forecasts in Andalusia (southern Spain). – Sol. Energy **86**, 2200–2217.

Long, C.N., E.G. Dutton, 2012: Bsrn Global Network recommended QC tests, V2.0. – document accessible via http://www.bsrn.awi.de/fileadmin/user_upload/Home/Publications/BSRN_recommended_QC_tests_V2.pdf.

Lorenz, E., J. Hurka, D. Heinmann, H. -G. Beyer, 2009: Irradiance forecasting for the power prediction of grid-connected photovoltaic systems. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2, 1, 2–10.

Marquez, R., C.F.M. Coimbra, 2011: Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. – Sol. Energy **85**, 746–756.

Mathiesen, P., J. Kleissl, 2011: Evaluation of numerical weather prediction for intra-day solar forecasting in the United States. – Sol. Energy **85**, 967–977.

Mathiesen, P., C. Collier, J. Kleissl, 2013: A high-resolution, cloud-assimilating numerical weather prediction model for solar irradiance forecasting. – Sol. Energy **92**, 47–61.

Morcrette, J.-J., 1991: Radiation and Cloud Radiative Properties in the European Centre for Medium Range Weather Forecasts Forecasting System. – J. Geophys. Res. **96**, D5, 9121–9132.

Morcrette, J.-J., 2002: Chapter on 'Radiation Transfer'. – Meteorological Training Course Lecture Series of ECMWF.

Morcrette, J.-J., G. Modzynski, M. Leutbecher, 2008a: A Reduced Radiation Grid for the ECMWF Integrated Forecasting System. – Mon. Wea. Rev. **136**, 4760–4772, DOI: 10.1175/2008MWR2590.1.

Morcrette, J.-J., H.W. Barker, J.N.S. Cole, M.J. Iacono, R. Pincus, 2008b: Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System. – Mon. Wea. Rev. **136**, 4773–4798. DOI: 10.1175/2008MWR2363.1.

Morcrette, J.-J., O. Boucher, L. Jones, D. Salmond, P. Bechtold, A. Beljaars, A. , Benedetti, A. , Bonet, J.W. Kaiser, M. Razinger, M. , Schulz, S. Serrar, A.J. Simmons, M. Sofiev, M. , Suttie, A.M. Tompkins, A. Untch, 2009: Aerosol analysis and forecast in the ECMWF Integrated Forecast System. Part I: Forward modelling. – J. Geophys. Res. **114**, D06206. DOI: 10.1029/2008JD011235.

Ohmura, A., E.G. Dutton, B. Forgan, C. Fröhlich, H. Gilgen, H. Hegner, A. Heimo, G. König-Langlo, B. McArthur, G. Müller, R. Philipona, R. , Pinker, C.H. Whitlock, K. Dehne, 1998: Baseline Surface Radiation Network (BSRN/WCRP): New Precision Radiometry for Climate Research. – Bull. Amer. Meteor. Soc. **79**, 2115–2136.

Pelland, S., G. Galanis, G. Kallos, 2011: Solar and photovoltaic forecasting through post-processing of the Global Environmental Multiscale numerical weather prediction model. – Progress in Photovoltaics: Research and Applications. DOI: 10.1002/pip.1180.

Perez, R., S. Kivalov, J. Schlemmer, Jr. K. Hemker, D. Renne, T.E. Hoff, 2010: Validation of short and medium term operational solar radiation forecasts in the US. – Sol. Energy **84**, 2161–2172.

Perez, R., E. Lorenz, S. Pelland, M. Beauharnois, G. van Knowe, K. Hemker Jr., D. Heinemann, J. Remund, S.C. Müller, W. TraunmÜLler, G. Steinmauer, D. Pozo, J.A. Ruiz-Arias, V. Lara-Fanego, L. Ramirez-Santigosa, M. Gaston-Romero, L.M. Pomares, 2013: Comparison of numerical weather prediction solar irradiance forecasts in the US, Canada and Europe. – Sol. Energy **94**, 305–326.

Reddy, M.S., O. Boucher, N. Bellouin, M. Schulz, Y. Balkanski, J. -L. Dufresne, M. Pham, 2005: Estimates of global multi-component aerosol optical depth and direct radiative perturbation in the Laboratoire de Météorologie Dynamique general circulation model. – J. Geophys. Res. **110**, D10S16. DOI:10.1029/2004JD004757.

Rodwell, M.J., T. Jung, 2008: Understanding the local and global impacts of model physics changes: an aerosol example. Quart. J. Roy. Meteor. Soc. **134**, 1479–1497. DOI: 10.1002/qj.298.

Ronzio, D.A., E. Collino, P. Bonelli, 2013: A survey on different radiative and cloud schemes for the solar radiation modeling. – Sol. Energy **98**, 153–166.

Schroedter-Homscheidt, M., A. Oumbe, 2013: Validation of an hourly resolved global aerosol model in answer to solar electricity generation information needs. – Atmos. Chem. Phys. **13**, 3777–3791. DOI:10.5194/acp-13-3777-2013.

Skartveit, A., J.A. Olseth, M.E. Tuft, 1998: An Hourly Diffuse Fraction Model with Correction for Variability and Surface Albedo. – Sol. Energy **63**, 173–183.

Tegen, I., P. Hoorig, M. Chin, I. Fung, D. Jacob, J. Penner, 1997: Contribution of different aerosol species to the global aerosol extinction optical thickness: Estimates from model results. – J. Geophys. Res. **102**, D20, 23895–23915.

Tompkins, A.M, C. Cardinali, J. -J. Morcrette, M.J. Rodwell, 2005: Influence of aerosol climatology on forecasts of the African Easterly Jet. – Geophys. Res. Lett. **32**, L10801. DOI:10.1029/2004GL022189.

Troccoli, A., J.-J. Morcrette, 2012: Forecast assessment of surface solar radiation over Australia, World Renewable Energy Conference, Denver, CO, USA, 2012. – Available at http://ases6.conference-services.net/resources/252/2859/pdf/SOLAR2012_0280_full%20paper.pdf.

Wittmann, M., H. Breitkreuz, M. Schroedter-Homscheidt, M. Eck, 2008: Case Studies on the Use of Solar Irradiance Forecast for Optimized Operation Strategies of Solar Thermal Power Plants. – IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing **1**, 1, 18–27.