

# SHORT-TERM SOLAR COLLECTOR POWER FORECASTING

Peder Bacher<sup>1</sup>, Henrik Madsen<sup>1</sup> and Bengt Perers<sup>2</sup>

<sup>1</sup>DTU Informatics, Richard Pedersens Plads, Building 321, DK-2800 Lyngby, Denmark

<sup>2</sup>DTU Civil Engineering, Brovej, Building 118, DK-2800 Lyngby, Denmark

## Abstract

This paper describes a new approach to online forecasting of power output from solar thermal collectors. The method is suited for online forecasting in many applications and in this paper it is applied to predict hourly values of power from a standard single glazed large area flat plate collector. The method is applied for horizons of up to 42 hours.

Solar heating systems naturally come with a hot water tank, which can be utilized for energy storage also for other energy sources. Thereby such systems can become an important part of energy systems with a large share of uncontrollable energy sources, such as wind power. In such a scenario online forecasting is a vital tool for optimal control and utilization of solar heating systems.

The method is a two-step scheme, where first a non-linear model is applied to transform the solar power into a stationary process, which then is forecasted with robust time-adaptive linear models. The approach is similar to the one by Bacher et al. (2009), but contains additional effects due to differences between solar thermal collectors and photovoltaics. Numerical weather predictions provided by Danish Meteorological Institute are used as input. The applied models adapt over time enabling tracking of changes in the system and in the surrounding conditions, such as decreasing performance due to wear and dirt, and seasonal changes such as leaves on trees. This furthermore facilitates remote monitoring and check of the system.

## 1 Introduction

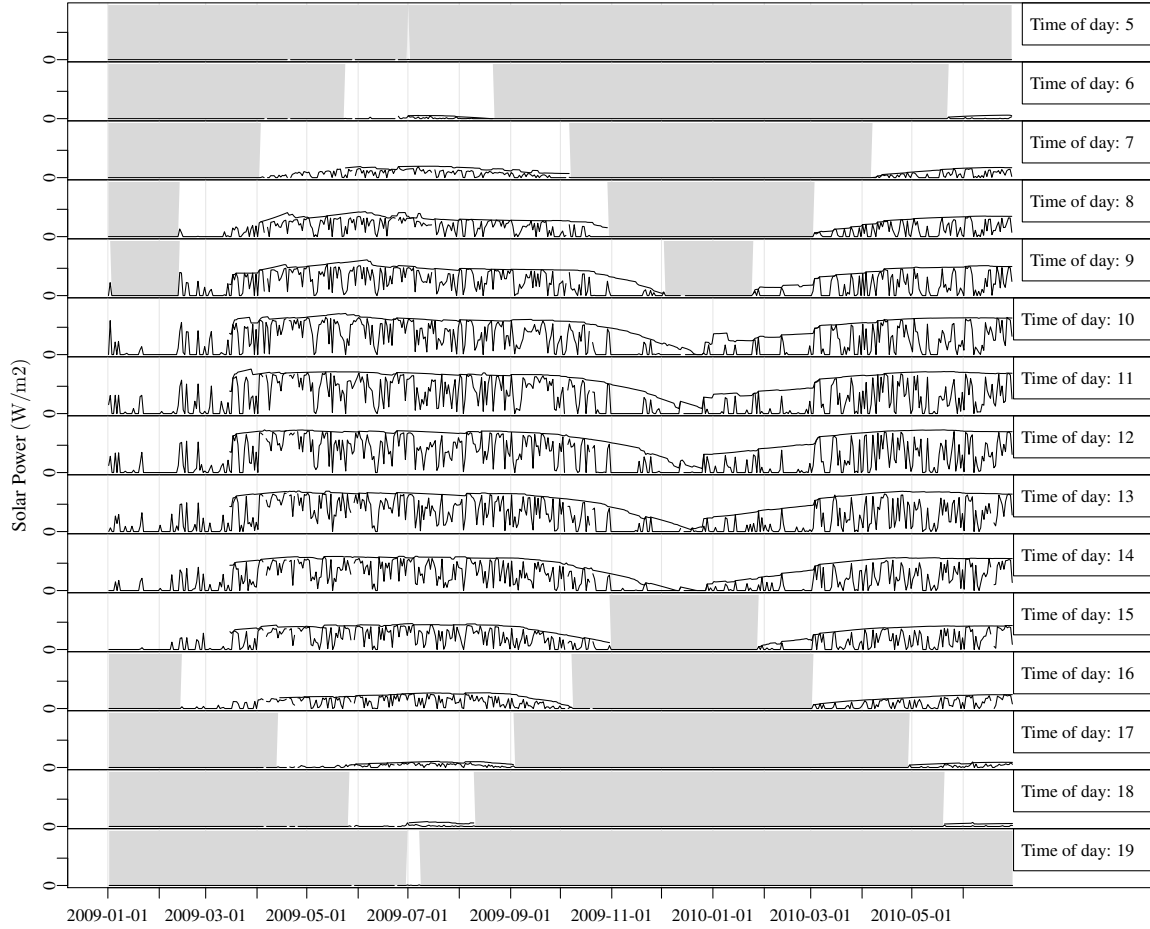
Forecasting of energy production is vital for optimization of energy systems which include wind and solar energy production. This paper describes an approach to online forecasting of power production from solar thermal collectors. In Denmark the level of wind power penetration already now gives periods with a surplus of energy and facilities to absorb this energy are needed. Solar heating systems with a hot water tank and auxiliary electrical heating can provide energy storage, which can facilitate absorption of wind energy and peak shaving, especially for levelling out diurnal energy consumption. The method is planned to be part of the control system for such heating systems (Perers et al., 2011). The study is carried out with climate data observed at a weather station at Danish Technical University. From this data, simulated hourly average values of solar thermal power is generated with a very detailed simulation model. Furthermore numerical weather predictions (NWP) provided by Danish Meteorological Institute data is used. The forecasting method is a two-step scheme, where first a statistical clear sky model is applied to transform the solar power into a more stationary process, which then is forecasted with robust time-adaptive linear models. The NWP are used as input to conditional parametric time-adaptive models to forecast the solar power. These forecasts are then transformed with the clear sky model, such that they can be applied as inputs to the linear forecasting models. Finally, a combined model, which is the most optimal for all horizons, is formed.

The paper is organized as follows. First the data and how it is preprocessed is described in a section. The next section contains an outline of the clear sky model, and this is followed by a section where all the forecasting models are described. Then an evaluation is given and the results are presented, each in a section. The second last section contains a discussion of the results and ideas for further work, and finally, the paper ends with a conclusion.

## 2 Data

The forecasting method is applied on simulated solar output power data for a flat plate collector carefully tested and modelled at DTU. A validated collector model and longterm climate data from the DTU Byg climate station was used to create realistic operating data for a solar collector during the year. The simulation model and weather data was introduced in TRNSYS 16 and the collector output power was calculated as hourly mean values. The simulation model is dynamic, such that dynamical effects - introduced when the collector starts and stops and during rapidly varying solar radiation conditions - are modelled.

In this study all time series are hourly average values and all units are implicitly per hour. Time points are set to the end of their respective sample period and all are in UTC. The units for radiation are  $\frac{W}{m^2}$  and for temperatures  $^{\circ}C$ .



**Figure 1:** The observed solar power for each hour of the day. The grayed area shows the periods not included in the modelling. The curve following the highest values of solar power is the estimated clear sky power, which is referred to in later parts of the paper.

### 2.1 Solar power

The simulated solar power time series is plotted for each hour of the day in Figure 1. A few short periods are with missing values.

### 2.2 Numerical weather predictions

The numerical weather predictions (NWP) used in the study are provided by Danish Meteorological Institute. The NWP model is DMI-HIRLAM-S05, which has a 5 kilometer grid and 40 vertical layers (Danish Meteorological Institute, 2011). NWP are updated every 6<sup>th</sup> hour and are up to a 48 hours horizon. They consist of hourly predictions of ambient temperature, and horizontal direct- and diffuse solar radiation. A couple of the considered forecasting models use the global radiation as input, which is simply the direct and the diffuse radiation summed. The scatter plots in Figure 2 shows the solar power versus the NWP for a 24 hour horizon. Clearly, the solar power is highly correlated with both the global and direct radiation, whereas the effect of diffuse and ambient temperature are not as apparent.

### 2.3 Pre-processing

On most locations on earth the solar radiation is zero at night time, hence the observed solar power is also zero. For the current dataset only periods, for a given hour of the day longer than 40 days in which the solar power is different from zero, are included. This is illustrated in Figure 1, where the non-included periods are grayed out. Furthermore a few short periods are missing from the observations. The time series of hourly observed

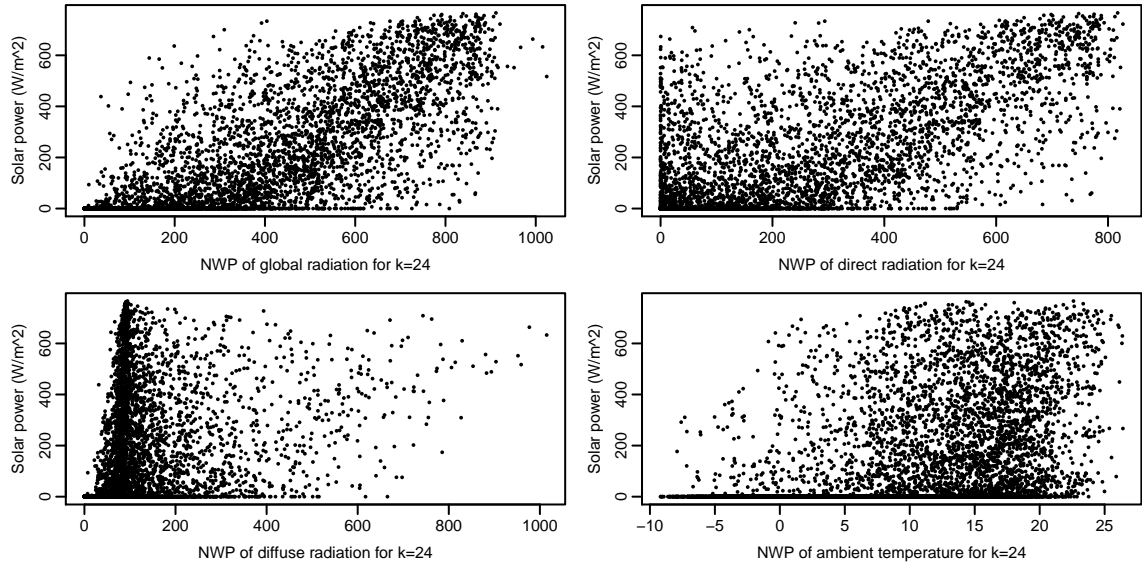


Figure 2: The observed solar power versus the NWP for  $k = 24$ , which are used as inputs to forecasting models.

solar power spanning the period from 2009-01-01 to 2010-07-01 is

$$\{P_t; t = 1, \dots, N\} \quad (1)$$

where  $N = 13104$ . The NWPs have a calculation time of 4 hours, which is taken into consideration, such that e.g. the forecast from 2009-01-01 00:00 are only available from 2009-01-01 04:00. The NWPs are pre-processed into time series of hourly values, such that the most recent available forecast  $k$  hours ahead is selected each hour. The time series are for a given  $k$ : the direct radiation

$$\{G_{t+k|t}^{b,nwp}; t = 1, \dots, N\} \quad (2)$$

the diffuse radiation

$$\{G_{t+k|t}^{d,nwp}; t = 1, \dots, N\} \quad (3)$$

and the ambient temperature

$$\{T_{t+k|t}^{a,nwp}; t = 1, \dots, N\} \quad (4)$$

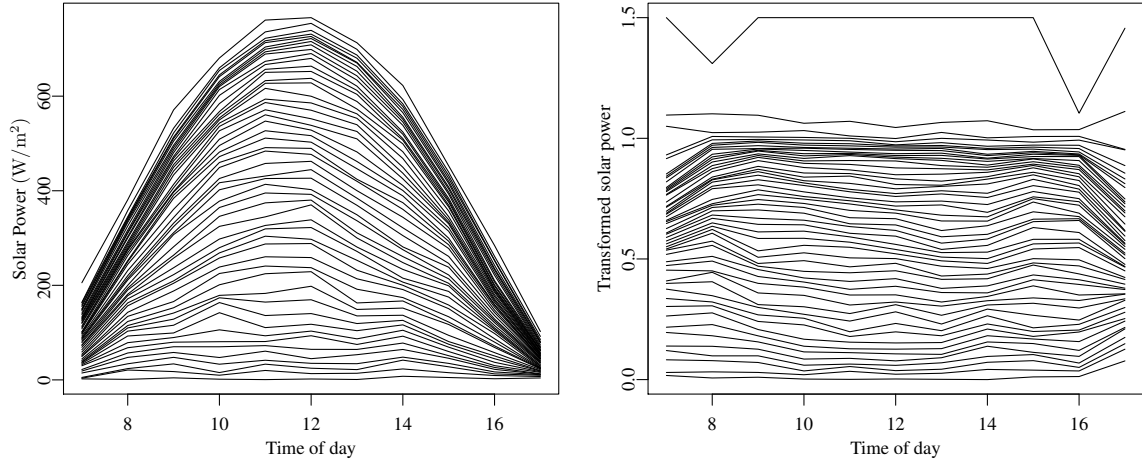
Due to the 6 hours interval the NWPs for horizons longer than 42 hours are not complete and therefore the solar power forecasting are only carried out up to 42 hours.

### 3 Clear sky model

For effective forecasting with classical linear time series methods stationarity of the process is required (Madsen, 2007). The process that generates the solar power is not stationary, which is seen by plotting quantiles of the distribution of solar power conditioned on the time of day. Such a plot is shown in Figure 3. Clearly the distribution of solar power is not independent of the time of day. The dependency can be removed by a transformation with a clear sky model

$$\tau_t = \frac{P_t}{P_t^{cs}} \quad (5)$$

where  $P_t$  is the observed solar power,  $P_t^{cs}$  is the estimated clear sky solar power, and  $\tau_t$  is the transformed solar power.



**Figure 3: The 0, 2%, ..., 100% quantiles of the distribution of the solar power and the transformed solar power conditioned on the time of day. Values above 1.5 has been clipped, which was the case for 17 values.**

### 3.1 Statistically estimated clear sky solar power

The clear sky solar power is estimated with a statistical non-linear and adaptive model. Quantile regression locally weighted in the day of year and time of day dimension is applied. In the present study this is carried out fully causal. The clear sky model is

$$P_t^{CS} = q_{0.99}(P_1, P_2, \dots, P_t, h_t, h_{tod}, h_y) \quad (6)$$

where  $q_{0.99}$  is the 99% quantile of all the solar power values up to  $t$ . The bandwidths  $h_t$ ,  $h_{tod}$ , and  $h_y$ , are in the time-, time of day-, and year-dimension, respectively. The bandwidths control how “locally” the model is fitted, i.e. a lower bandwidth puts more emphasis on data which is close in time. The local weighting function is an Epanechnikov kernel. The applied bandwidths are

$$h_t = 120 \text{ days}, \quad h_{tod} = 2 \text{ hours}, \quad h_y = 1.7 \text{ years} \quad (7)$$

which were found by visual inspection of the fitted clear sky curve. Finally, it is noted that second-order polynomials were applied in the time- and time of day-dimension to include curvature into the model. The estimate of the clear sky solar power is shown in Figure 1. From the plot it is seen that it follows the highest values of solar power quite well. Clearly, the clear sky power is most easily carried out in the periods with a high level of solar power. One advantage of the transformation is that it will automatically adapt to changes in the system, such as degraded performance or changes in the surroundings e.g. snow cover and shadowing effects. It can as well be used for monitoring of the solar system, since degraded performance from the same time of year will result in a lower clear sky solar power curve. Quantile plots of the transformed solar power conditioned on the time of day are shown in Figure 3, from which it is seen that the transformed solar power process is considerably less depend end of the time of day and therefore a much more stationary process. It is noted that further work could include physical considerations e.g. by using the air mass as an input.

## 4 Forecasting models

In this section a description of the applied forecasting models is given. The models can be divided into models without NWP as input - autoregressive (AR) models - and models with NWP as input: conditional parametric (CP) and autoregressive with exogenous inputs (ARX) models. Each model is fitted separately for each horizon, such that the same model structure is used, but the parameters are estimated separately for each horizon. In the final model, a combination of models are used to achieve the most optimal performance for all horizon.

### 4.1 Reference model

To compare the performance of prediction models, and especially when making comparisons between different studies, a common reference model is essential. The reference model for solar power used in this study is the

best performing naive predictor for a given horizon. Two naive predictors of solar power are found to be relevant. Persistence

$$p_{t+k|t} = p_t + e_{t+k}, \quad (8)$$

and diurnal persistence

$$p_{t+k|t} = p_{t-s(k)} + e_{t+k} \quad (9)$$

$$s(k) = f_{\text{spd}} + k \bmod f_{\text{spd}} \quad (10)$$

where  $f_{\text{spd}} = 24$  is the sample frequency in number of samples per day and  $s(k)$  ensures that the latest diurnal observation is used, i.e. the value which, depending on the horizon, is either 24 or 48 hours before the time point that is to be forecasted.

#### 4.2 Autoregressive models

Autoregressive (AR) models are applied to forecast the transformed solar power. These models can include either the latest available observation or the latest available diurnal observation, or both, as input. The models are fitted with  $k$ -step recursive least squares with forgetting factor (Bacher et al., 2009). The model formulated as a  $k$ -step AR model

$$\tau_{t+k|t} = m + a_1 \tau_t + a_{24} \tau_{t-s(k)} + e_{t+k} \quad (11)$$

$$s(k) = 24 + k \bmod 24 \quad (12)$$

where the function  $s(k)$  ensures that the latest observation of the diurnal component is included. It was found that depending on the horizon better performance was achieved by only using one input. Thus for short horizons (1 and 2 hours) the model without the diurnal component

$$\tau_{t+k|t} = m + a_1 \tau_t + e_{t+k} \quad (13)$$

was found to have the best performance, it is denoted *ARI*, and for longer horizons the model with only the diurnal component

$$\tau_{t+k|t} = m + a_{24} \tau_{t-s(k)} + e_{t+k} \quad (14)$$

was found to have the best performance, it is denoted *ARDiurnal*.

#### 4.3 Conditional parametric models with NWP as input

Models based on NWPs of solar radiation and ambient temperature are described in this section. It is known from physics (Perers, 1997) that the power output of a solar collector can be described by

$$P = F'(\tau\alpha)_{\text{en}} K_{\tau\alpha b}(\theta) G_{\text{b,col}} + F'(\tau\alpha)_{\text{en}} K_{\tau\alpha d} G_{\text{d,col}} - F'U_0 \left( \frac{T_o + T_i}{2} - T_a \right) \quad (15)$$

where the  $G_{\text{b,col}}$  and  $G_{\text{d,col}}$  are respectively direct and diffuse solar radiation normal to the collector plane. This is formed into a forecasting model based on NWPs by rewriting as follows. First, both the angle of incidence modifier  $K_{\tau\alpha b}(\theta)$  and the transformation of solar radiation from horizontal to the collector plane are modelled by letting the coefficients - for the radiation effects - become a function of time  $t$  and time of day  $t_{\text{tod}}$ . Furthermore, assuming that the outlet temperature is a function of the solar radiation

$$T_o = f_b(G_{\text{b,col}}) + f_d(G_{\text{d,col}}) \quad (16)$$

this give the total effect of direct radiation as a non-linear function

$$a(t, t_{\text{tod}}, G_b, G_d) G_b = F'(\tau\alpha)_{\text{en}} K_{\tau\alpha b}(\theta) G_{\text{b,col}} + F'U_0 \frac{1}{2} f_b(G_{\text{b,col}}) \quad (17)$$

and for the diffuse radiation

$$b(t, t_{\text{tod}}, G_b, G_d) G_d = F'(\tau\alpha)_{\text{en}} K_{\tau\alpha d} G_{\text{b,col}} + F'U_0 \frac{1}{2} f_d(G_{\text{d,col}}) \quad (18)$$

Finally, the effect of the ambient temperature is kept as

$$cT_a = F'U_0T_a \quad (19)$$

and by assuming a constant inlet temperature this part becomes a constant effect

$$m = F'U_0 \frac{T_i}{2} \quad (20)$$

Thus the CP model structure used for forecasting is

$$P = m + a(t, t_{\text{tod}}, G_b, G_d)G_b + b(t, t_{\text{tod}}, G_b, G_d)G_d + cT_a \quad (21)$$

Since the time-dependency and non-linearity are smooth functions in the parameters, it is modelled with conditional parametric (CP) models. The time varying effect is modelled by conditioning on  $t$  and  $t_{\text{tod}}$  - this is equivalent of a local constant effect - and the dependency of the radiation is modelled with 1-order local polynomials. A kernel method is applied, using a nearest neighbor approach to find the bandwidth of an Epanechnikov weighting function. From Equation (15) it is seen that the output can be negative if little radiation hits the collector and the ambient temperature is low. In this case the output is zero since the system stops. This effect can be seen on the plot in Figure 2. It is handled by the non-linearity of the models and by setting negative forecasts to zero.

The simplest considered conditional parametric model is

$$P_{t+k} = m + a(t, t_{\text{tod}}, G_{t+k|t}^{\text{nwp}})G_{t+k|t}^{\text{nwp}} + e_{t+k} \quad (22)$$

where  $G_{t+k|t}^{\text{nwp}}$  is the  $k$ -hour NWP of global radiation and denoted as *CPI* in the following. This second CP model has NWPs of direct and the diffuse radiation as inputs

$$P_{t+k} = m + a(t, t_{\text{tod}}, G_{t+k|t}^{\text{b,nwp}}, G_{t+k|t}^{\text{d,nwp}})G_{t+k|t}^{\text{b,nwp}} + b(t, t_{\text{tod}}, G_{t+k|t}^{\text{b,nwp}}, G_{t+k|t}^{\text{d,nwp}})G_{t+k|t}^{\text{d,nwp}} + e_{t+k} \quad (23)$$

where  $G_{t+k|t}^{\text{b,nwp}}$  is the  $k$ -hour NWP of direct radiation and  $G_{t+k|t}^{\text{d,nwp}}$  is the  $k$ -hour NWP of diffuse radiation, and denoted as *CP2*. Finally the model is expanded with NWPs of ambient temperature

$$P_{t+k} = m + a(t, t_{\text{tod}}, G_{t+k|t}^{\text{b,nwp}}, G_{t+k|t}^{\text{d,nwp}})G_{t+k|t}^{\text{b,nwp}} + b(t, t_{\text{tod}}, G_{t+k|t}^{\text{b,nwp}}, G_{t+k|t}^{\text{d,nwp}})G_{t+k|t}^{\text{d,nwp}} \quad (24)$$

$$+ c(t, t_{\text{tod}})T_{t+k|t}^{\text{a,nwp}} + e_{t+k} \quad (25)$$

where  $T_{t+k|t}^{\text{a,nwp}}$  is the  $k$ -hour NWPs of the ambient temperature and the model is denoted as *CP3*.

In the following the coefficients dependency of the time of day is elaborated on. Plots of the fitted forecasting function  $a(t, t_{\text{tod}}, G_{t+k|t}^{\text{b,nwp}}, G_{t+k|t}^{\text{d,nwp}})$  are shown in Figure 4. It is seen how the slope of the function is lower in the morning, than in the middle of the day. This is naturally caused by the higher angle of incidence in the morning, which cause less horizontal radiation to be absorbed due to reflection. Likewise for the afternoon. Finally, non-linearity in the fit is seen, which is caused by the non-negativity of the solar power (mentioned above) and varying uncertainty of the NWPs.

#### 4.4 Autoregressive model with exogenous input

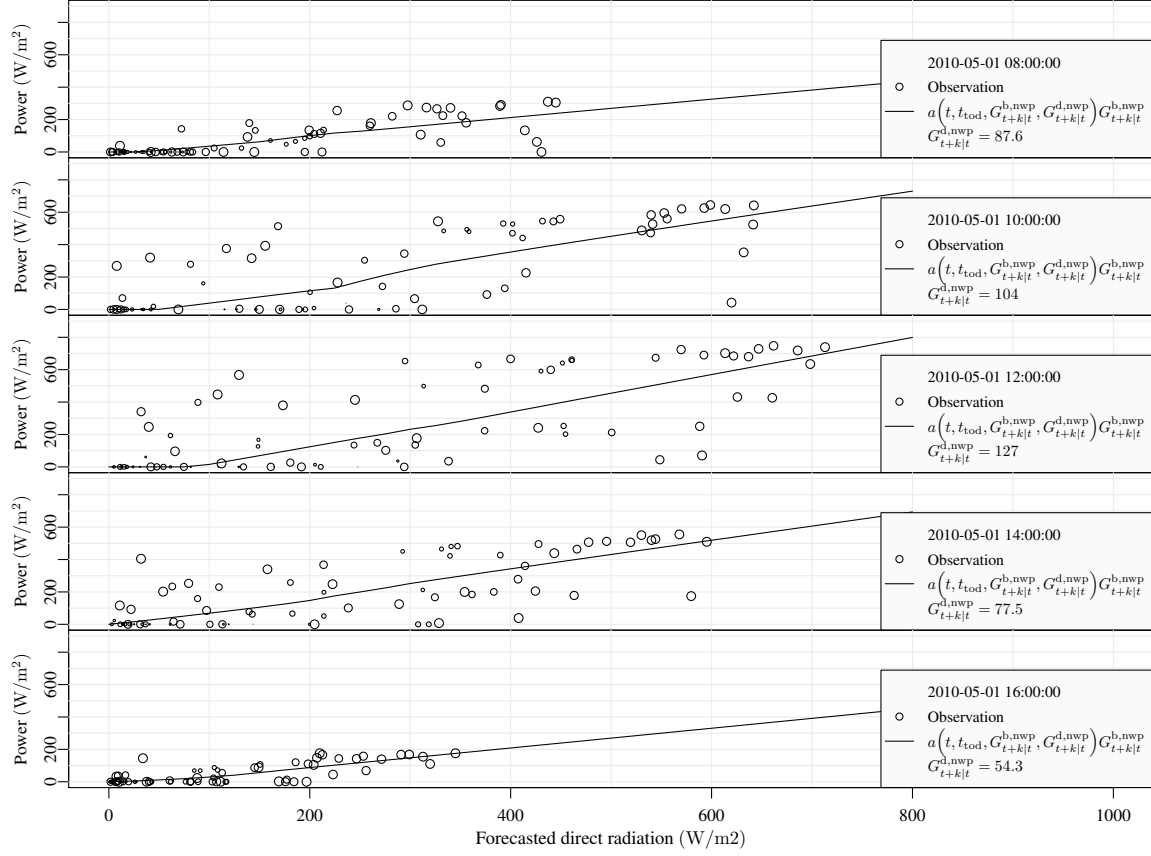
The AR model is be expanded to include the forecast of the CP models, thus combining information in past observed solar power and NWPs. The solar power forecasts from the CP is transformed with the clear sky model by

$$\hat{\tau}_{t+k|t}^{\text{nwp}} = \frac{\hat{P}_{t+k|t}^{\text{nwp}}}{P_{t-s(k)}^{\text{cs}}} \quad (26)$$

$$s(k) = f_{\text{spd}} + k \bmod f_{\text{spd}} \quad (27)$$

where  $f_{\text{spd}} = 24$  is the sample frequency in number of samples per day. This is applied as an input to the ARX model

$$\tau_{t+k|t} = m + a_1\tau_t + a_{24}\tau_{t-s(k)} + b_1\tau_{t+k|t}^{\text{nwp}} + e_{t+k} \quad (28)$$



**Figure 4:** Examples of the function fitted for forecasting of the effect of direct radiation at different times of the day the 1<sup>st</sup> of May 2010 with the CP2 model. For each observation the size of circle indicates the weighting of the observation in the CP models. Thus observations with a larger circle have more influence on the fitted function.

Again, as for the AR models, different performance is found depending on the horizon. The *ARX1* model is best for short horizons

$$\tau_{t+k|t} = m + a_1 \tau_t + b_1 \tau_{t+k|t}^{nwp} + e_{t+k} \quad (29)$$

*ARXDiurnal* for horizons up to 24 hours

$$\tau_{t+k|t} = m + a_{24} \tau_{t-s(k)} + b_1 \tau_{t+k|t}^{nwp} + e_{t+k} \quad (30)$$

and *ARX*

$$\tau_{t+k|t} = m + b_1 \tau_{t+k|t}^{nwp} + e_{t+k} \quad (31)$$

for longer horizons.

#### 4.5 Combined model

The final model is a combination of the previously described models. The model is denoted *ARXCombined*. First, missing values in forecasts from *ARX1* are replaced with forecast values from *ARXDiurnal*. These missing values are in the morning, since they were tried to be forecasted based on night values, which are zero. For horizons longer than 30 hours forecasts from *ARX* are used. Finally, any remaining missing values - which are only where the diurnal lag was not present for *ARXDiurnal* - are replaced with forecasted values from *CP2*.

## 5 Evaluation

The methods used for evaluating the prediction models are inspired by Madsen et al. (2005). They suggest a framework for evaluation of wind power forecasting. The clear sky model, RLS, and CP fitting do not use

any degrees of freedom and the dataset is therefore not divided into a training set and a test set. It is only for the optimization of the kernel bandwidths and the forgetting factor that the entire dataset is used. The period before 2009-03-15 is considered as a burn-in period and are not used when calculating the error measures.

### 5.1 Error measures

The Root Mean Square Error for the  $k$ 'th horizon is

$$RMSE_k = \left( \frac{1}{N} \sum_{t=1}^N e_{t+k}^2 \right)^{\frac{1}{2}} \quad (32)$$

where  $e_{t+k}$  is  $k$ -hourly prediction error. The  $RMSE_k$  is used as the main evaluation criterion (EC) for the performance of the models. The Normalized Root Mean Square Error is found by

$$NRMSE_k = \frac{RMSE_k}{p_{\max}} \quad (33)$$

where  $p_{\max}$  is the maximum observed solar power output. The mean value of the  $RMSE_k$  for a range of horizons

$$\overline{RMSE}_{k_{\text{start}}, k_{\text{end}}} = \frac{1}{k_{\text{end}} - k_{\text{start}} + 1} \sum_{k=k_{\text{start}}}^{k_{\text{end}}} RMSE_k \quad (34)$$

is used as a summary error measure. When comparing the performance of two models the improvement

$$I_{\text{EC}} = 100 \cdot \frac{EC_{\text{ref}} - EC}{EC_{\text{ref}}} (\%) \quad (35)$$

is used, where  $EC$  is the considered evaluation criterion. When calculating the error measures it is important to consider how to handle missing values for the solar power forecasts. The problem is handled by replacing missing forecast values with forecast values from the reference model *Ref*.

### 5.2 Completeness

In order to evaluate a model for its performance regarding missing forecast values a measure is defined, it is denoted completeness. The completeness of a forecast for horizon  $k$ , is the ratio of the total sum of solar power and the summed solar power for time points where the forecasts are not missing

$$C_k = \frac{\sum_{t=1}^N P_t I(\hat{P}_{t|t-k})}{\sum_{t=1}^N P_t} \quad (36)$$

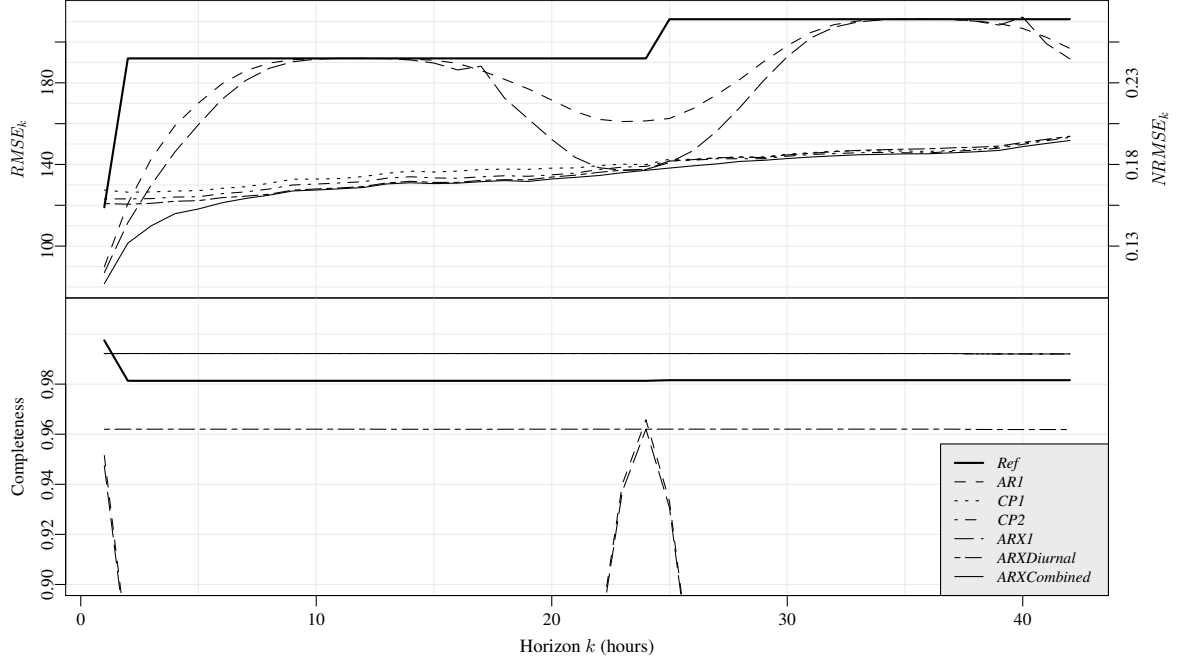
where  $I()$  is the indicator function which is 0 if  $\hat{P}_{t|t-k}$  is missing, and 1 if not. Only the included values are used, i.e. not night values.

## 6 Results

In this section the results are presented and evaluated. The  $\overline{RMSE}_{k_{\text{start}}, k_{\text{end}}}$  improvement for relevant ranges of horizons are listed in Table 1. For selected models the  $RMSE_k$  is shown in the upper plot of Figure 5 and the completeness in the lower.

Considering the improvements it is seen that most of the models perform very well on either the short horizons or the longer horizons. Starting with short horizons (1 to 2 hours), the *ARI* and *ARXI* are clearly superior, which is due to their inclusion of the most present autoregressive lag. Their performance on longer horizons are not good. The reason for this is found by considering the plot of  $RMSE_k$  and completeness. Here it is seen that the completeness of *ARI* and *ARXI* drops really quickly as the horizon increase, which cause the  $RMSE_k$  to increase and reach the reference model at the 10 hours horizon. This is simply due to missing forecast values, since for e.g. the 10 to 14 hours horizons the models use night values (which are missing) to forecast day values with.





**Figure 5:** The upper plot is  $RMSE_k$  for the forecasting models. On the right side the  $NRMSE_k$  is indicated. The lower plot is completeness  $C_k$ .

For horizons longer than three hours the best performance is seen for the models, that doesn't include the most present AR lag. The *ARDiurnal* is a clear improvement from the *AR1*, and the CP and ARX models - which include the NWP's - are superior for these horizons. An improved performance is found from *CP1* to *CP2* mainly for 3 to 24 hours horizons, whereas no clear increase in improvement is found from *CP2* to *CP3*. The CP models are slightly improved by using them as input to the ARX models, since autocorrelation of the errors are modelled.

Finally, the combined model *ARXCombined* utilizes the best parts of: *ARX1*, *ARX*, *ARXDiurnal*, and *CP2*. Especially the replacement of missing forecast values improves the performance for horizons up to 5 hours. The completeness of the combined model is as high as any of the others.

## 7 Discussion and applications

This section contains a short discussion of the results and ideas for further work, and ends with an outline of applications.

Considering the improvement achieved over the reference model the forecasting method is found to perform very well. Clearly the quality of the NWP's of solar radiation is the most influential source of error, hence improved NWP's will improve the forecasting performance. Regarding improvement of the forecasting models, the following are considered:

**Table 1:** Improvements in percent for selected ranges of horizons.

Model	$I_{RMSE_{1,2}}$	$I_{RMSE_{3,24}}$	$I_{RMSE_{25,42}}$	$I_{RMSE_{1,42}}$
<i>AR1</i>	30.8	7.1	6.1	7.8
<i>ARDiurnal</i>	-10.7	15	18	15.1
<i>CP1</i>	13.5	30	30.6	29.5
<i>CP2</i>	16.2	31.4	30.9	30.5
<i>CP3</i>	15.5	31.6	30.8	30.5
<i>ARX</i>	17.1	32	31.3	31
<i>ARX1</i>	34.4	11.6	8.7	11.4
<i>ARXDiurnal</i>	17.8	32.4	30.5	30.9
<i>ARXCombined</i>	39.3	33.3	31.5	32.8

- A thorough evaluation of the forecast errors to clarify how the models can be improved.
- Optimization of forgetting factor of the RLS has not been carried out, this will improve the performance of the AR and ARX models.
- Application of regime models to handle different aspects of forecasting for low and high radiation values.
- More optimal ways to combine the models. Utilizing a linear combination instead of the simple replacement currently carried out.
- It might be possible improve performance by including a third-stage, where modelling of the errors are carried out.

The applications for this type of solar thermal power forecasting counts the integration of solar thermal energy systems with auxiliary electrical heating into smart grid systems (Perers et al., 2011). The solar power forecasts will be used for model predictive control to optimize the operation of the system. Other applications include optimal control of large solar heating plants.

The method is furthermore well suited for monitoring the performance of solar thermal systems. Measures of the performance can be derived from the CP models, with which systems can be compared on an absolute scale. Sudden high deviation from the CP forecasting model will allow for very fast detection of failures in the system. For an individual system the change in performance over time can also be assessed by monitoring the clear sky curve for unusual behavior, and compare the change from year to year.

## 8 Conclusion

A method for forecasting of solar thermal power output is presented. It is applied to forecast hourly values for horizons up to 42 hours. The method is based on a two-stage approach, where first the solar power is normalized with a statistical clear-sky model, and secondly forecasted with time-adaptive linear time series models. Both models without and with NWP of solar radiation and ambient temperature are considered. The NWPs are included by using non-linear conditional parametric models, which are formed from prior physical knowledge. The forecast models which do not use NWPs achieve an improvement on short horizons (1 to 2 hours) in average 30% over a persistence reference model, and in average 15% on horizons up to 42 hours. Applying the NWPs an improvement around 39 % is achieved in average for short horizons and around 32% in average for longer horizons. The method can furthermore be applied to monitor and check the performance of solar thermal collectors.

## Nomenclature

$P_t$	Hourly solar thermal power, $[\text{W}/\text{m}^2]$ .
$P_t^{\text{cs}}$	Estimated clear sky solar power, $[\text{W}/\text{m}^2]$ .
$\tau_t$	Normalized solar power.
$t$	Time index, [h].
$k$	Forecast horizon index, [h].
$t_{\text{tod}}$	Time of day.
$F'(\tau\alpha)_{\text{en}}$	Zero loss efficiency of collector for direct radiation at normal incidence
$K_{\tau\alpha\text{b}}(\theta)$	Incidence angle modifier for direct radiation
$K_{\tau\alpha\text{d}}$	Incidence angle modifier for diffuse radiation
$F'U_0$	Heat loss coefficient at $(T_a - T_f) = 0$ , $[\text{W}/(\text{m}^2\text{K})]$ .
$G_{t+k t}^{\text{nwp}}$	NWP of global radiation, $[\text{W}/\text{m}^2]$ .
$G_{t+k t}^{\text{b,nwp}}$	NWP of direct solar radiation, $[\text{W}/\text{m}^2]$ .
$G_{t+k t}^{\text{d,nwp}}$	NWP of diffuse solar radiation, $[\text{W}/\text{m}^2]$ .
$T_{t+k t}^{\text{a,nwp}}$	NWP of ambient temperature, $[\text{°C}]$ .
$\hat{P}_{t+k t}^{\text{nwp}}$	$k$ -hour prediction of solar power, $[\text{W}/\text{m}^2]$ .
$\hat{\tau}_{t+k t}$	$k$ -hour prediction of normalized solar power.
$e_{t+k}$	$k$ -step prediction error.

## References

- P. Bacher, H. Madsen, and H. A. Nielsen. Online short-term solar power forecasting. *Solar Energy*, 83(10):1772–1783, 2009. ISSN 0038092x.
- Danish Meteorological Institute. DMI-HIRLAM-S05, 2011. URL [http://www.dmi.dk/eng/index/research\\_and\\_development/dmi-hirlam-2009.htm](http://www.dmi.dk/eng/index/research_and_development/dmi-hirlam-2009.htm).
- H. Madsen. *Time Series Analysis*. Chapman & Hall, 2007.
- H. Madsen, P. Pinson, G. Kariniotakis, H. A. Nielsen, and T. S. Nielsen. Standardizing the performance evaluation of shortterm wind power prediction models. *Wind Engineering*, 29(6):475, 2005. ISSN 0309524x.
- B. Perers. An improved dynamic solar collector test method for determination of non-linear optical and thermal characteristics with multiple regression. *Solar Energy*, 59(4-6):163–178, 1997. ISSN 0038092x.
- B. Perers, S. Furbo, J. Fan, E. Andersen, and Z. Chen. Solar combisystems with forecast control to increase the solar fraction and lower the auxiliary energy cost. In *ISES Solar World Congress Proceedings*, 2011.