2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

Support Vector Regression of Multiple Predictive Models of Downward Short-wave Radiation

Pavel Krömer, Petr Musílek, Emil Pelikán, Pavel Krč, Pavel Juruš, and Kryštof Eben

Abstract— **Accurate forecasts of weather conditions are of the utmost importance for the management and operation of renewable energy sources with intermittent (stochastic) production. With the growing amount of intermittent energy sources, the need for precise weather predictions increases. Production of energy from renewable power sources, such as wind and solar, can be predicted using numerical weather prediction models. These models can provide high-resolution, localized forecast of wind speed and solar irradiation. However, different instances of numerical weather prediction models may provide different forecasts, depending on their properties and parameterizations. To alleviate this problem, it is possible to employ multiple models and to combine their outputs to obtain more accurate localized forecasts. This work uses the machine-learning tool of Support Vector Regression to amalgamate downward short-wave radiation forecasts of several numerical weather prediction models. Results of SVR-based multi-model forecasts of irradiation at a large set of locations show a significant improvement of prediction accuracy.**

I. Introduction

Photovoltaic (PV) electricity generation has been rapidly increasing in many countries around the world. Installed capacity over the last decade reached at least 96.5 GW and, at the end of 2012 (IEA PVPS 2013), PV could produce around 115 TWh of electricity on a yearly basis. This represents 0.6% of the world electricity consumption in 2012. PV also represents 20% of the world's installed capacity of renewables, excluding hydropower. Some countries have the ability to produce more than 5% of their electricity production (e.g. Italy and Germany). Accelerated development of new PV industry can be expected in several major countries outside Europe: in China, India and in other Asian countries, and also in the Middle East, South and Central America, and in Africa.

However, the outputs of PV systems are very volatile and they exhibit variability at all timescales from seconds to years. The forecasting of PV output is very difficult and new methods and approaches are rapidly being developed and tested. Among them, numerical weather prediction (NWP) plays a key role in the short term prediction horizon from several hours up to several days ahead. However, these models are very sensitive to initial and boundary conditions provided as the input to the forecasting process, along with digital elevation models and gridded characteristics of the Earth surface. One way to improve solar and PV forecasting is to combine forecasts from different NWP models. In this article, we use Support Vector Regression (SVR) techniques to combine outputs of four different NWP models to obtain a more accurate forecasts of downward short-wave radiation for a specific location.

This article is organized in five sections. Section II provides an introduction to PV power forecasting process and outlines methods used for solar irradiance prediction at different time horizons. Other studies related to forecasting of energy-related weather phenomena are summarized in Section III. The next section IV briefly introduces SVR, the machine learning method used to combine multiple NWP models. The models, data and procedures used for experiments are described in section V. The last section VI provides major conclusions and suggests several possibilities for future work.

II. Background

Forecasting of PV power production is typically based on the prediction of solar irradiance. The entire forecasting chain usually involves: prediction of global horizontal solar irradiance $\rightarrow$ transformation of these predictions to irradiance on a slanted solar panel $\rightarrow$ transformation of the obtained irradiance on slanted solar panel to prediction of PV power production. All the steps of this tool chain are critical for a successful prediction.

The first step contributes a substantial part of forecast error into final prediction due to the complexity of atmospheric dynamics in general and the complexity of physical processes related to optical properties of atmosphere in particular. Therefore a lot of attention is given to proper prediction of horizontal solar irradiance.

Approaches used for the forecasting of global horizontal irradiance include methods of mathematical statistics, artificial intelligence and machine learning, and methods of numerical weather prediction. Their performance depends mainly on forecasting horizon and the type of underlying data. Several or all these methods can be combined to form an ensemble of predictors.

Time series modeling usually outperforms other techniques for short prediction horizons (typically up to 2 hours). Methods used for this type of predictions can be based on statistical time series methods (comparison of several

Pavel Krömer and Petr Musilek are with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton AB T6G 2V4, Canada (email: {pavel.kromer, petr.musilek}@ualberta.ca). Pavel and Petr are also with VŠB-Technical University of Ostrava, Ostrava, Czech Republic. Emil Pelikán, Pavel Krč, Pavel Juruš, and Kryštof Eben are with the Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 2, 182 07 Prague 8, Czech Republic (email: {emil, krc, jurus, eben}@cs.cas.cz)

This work was supported by the Natural Science and Engineering Research Council (NSERC) of Canada, the Helmholtz-Alberta Initiative, by project LD12009 of the Ministry of Education of the Czech Republic, and by project SGS No. SP2014/110, VŠB - Technical University of Ostrava.

time series approaches can be found, e.g., in [1]) or on the artificial neural networks or other methods of computational intelligence (CI). Reviews of CI methods for solar radiation prediction can be found in [2] and in [3]. Methods based on the extrapolation of cloud motions can be advantageous for horizons ranging from several minutes up to 6 hours. Underlying data for the prediction are either satellite data or images from ground-based sky imagery. Examples of the work based on cloud motion extrapolation can be found in [4], [5] and [6].

NWP models are used for horizons from hours to days ahead, although the forecast skill decreases gradually for horizons longer than 3 days; the longest still useful forecast horizon is somewhere around 15 days ahead. Widely used choices for solar irradiance forecasting are the NWP models MM5 [7] and WRF [8]. Both models offer multiple options for the parameterizations of cloud microphysics, planetary boundary layer processes and for shortwave and longwave radiative transfer in the atmosphere. Evaluation of the skill of numerical weather prediction model with respect to global horizontal irradiance forecast is discussed e.g. in [9], and [10] (together with performance of different parameterizations).

Longer horizons cannot be reliably forecast by numerical weather prediction and various statistical approaches based on historical data are used. Typical source of the data for long-term forecasts (weeks, months or even years ahead) are historical measurements (in-situ or satellite, see e.g. [11]) or meteorological reanalyses (e.g. [12]).

III. Related work

An accurate prediction of weather phenomena such as solar irradiance and wind power is a key factor for efficient PV and wind electricity generation in the context of power consumption forecasting, infrastructure and maintenance planning, grid connectivity etc. Forecasting models proposed in the literature predict either solar irradiance and wind speed or directly the power generated by PV and wind power plants. The need for accurate predictions has led to an extensive research in this area. This section provides a brief overview of recent applications of ANNs, SMVs, and related (hybrid) techniques to weather phenomena predictions. A motivation and an overview of various approaches (i.e. persistent, physical, statistical, and machine learning-based approaches) to wind power forecasting was given in 2010 by Lydia and Kumar [13]. The machine learning techniques surveyed in the work included ANNs and SVMs.

A. SVM-based approaches

Illuminance prediction by Support Vector Regression (SVR) was proposed by Bellocchio et al. in 2011 [14]. The study showed that the proposed approach outperformed several naïve predictors. However, the data used by the authors for forecast accuracy comparison contained comprehensive information of previous illuminance (e.g. illuminance one hour ago, average illuminance etc.). A machine learning-based method for the creation of site-specific power generation prediction models for small-scale installations (e.g. smart homes) has been proposed by Sharma et al. [15]. The study in fact predicted site-specific solar irradiance in Wm^{-2} which was then used to compute power output. Among the models compared in scope of the study, a SVM with RBF kernel function performed best when processing 7 weather attributes (day, temperature, dew point, wind speed, sky coverage, precipitation, and humidity) on input and predicting the irradiance for the next 3 hours.

B. ANN-based approaches

Another short-term solar irradiance prediction method utilizing time delay neural network (TDNN) is due to Ji and Chan [16]. The authors used a combination of Autoregressive and Moving Average (ARMA) and TDNN to predict hourly solar irradiance in Singapore. It was shown that a hybrid combination of ARMA and TDNN performs for the conditions in Singapore better than individual models. An application of a simple multilayer feed forward neural network to PV power plant output prediction was investigated in 2012 by Prokop et al. [17]. The prediction by ANN was compared to prediction obtained by Adaptive Network-based Fuzzy Inference System (ANFIS) which performed better. Zarnani et al. [18] proposed a new SVM based method for the prediction of a different weather phenomena - ice accretion on electric power lines. The SVR-based model investigated in the study was compared to a number of other ice accretion forecasting techniques and it was concluded that it performs better than the other state-of-the-art methods. A series of evaluations covering a “single-storm” scenario (both training and test data comes from single storm) and ”multiple-storm” scenario (training data is from one storm, test data comes from another storm) was performed to support the findings.

C. Hybrid approaches

A number of studies focusing on the prediction of solar irradiance, wind speed, or renewable power production was published in 2013. Voyant et al. [19] proposed a hybrid methodology for short-term global irradiance prediction in the Mediterranean. The proposed algorithm combined ARMA and ANN and evaluated a number of ARMA/ANN configurations on data sets obtained from weather stations from the south of France. The best model found by the study selected particular predictor on the basis of the clear sky index and slightly outperformed other predictors. However, as noted by the authors, the prediction accuracy improvement obtained by the hybrid model was not overwhelming. Mellit et al. [20] used two ANNs (one for cloudy days, one for clear days) for PV power plant output prediction. The study showed that the ANNs perform better than other models (e.g. linear and polynomial regression, analytical model, one-diode model).

Chen et al. [21] used a combination of Fuzzy System and ANN for the prediction of energy produced by a grid-connected PV system. In the proposed method, a number of ANNs were trained to forecast solar energy at different sky conditions (e.g. cloud coverage). In the prediction phase, the fuzzy system selected particular ANN on the basis of momentary cloud coverage and the ANN predicted the power output. The hybrid method was compared to statistical methods and individual fuzzy and ANN-based predictors and found better.

Yan et al. [22] used in 2013 an optimized Relevance Vector Machine (RVM) for the prediction of wind power values and range at certain confidence level. The proposed method focused on large-scale wind farms and used self-organizing feature mapping to cluster the turbines into groups for which models were developed. The models were used to correct NWP errors in each month (i.e. model for each month was created for each turbine group). The parameters of RVM models in the study were optimized by Particle Swarm Optimization and compared to SVM and Genetic Algorithm/ANN based prediction models. It was concluded that the short-term prediction error of the optimized RVM was lowest, but the training time of SVM was shorter.

Another study dealing with ANNs in order to predict PV power plant output is due to Zhu and Yao [23]. The neural network in this work was trained (using the backpropagation algorithm) to process some NWP information (solar radiation intensity, temperature, humidity) and provide accurate short-term power output forecasts. A SVR-based PV power output predictor was a part of recent study by Prokop et al. [24]. The authors compared a simple ϵ-SVR based model with a method based on artificial evolution of fuzzy rules and concluded that both methods provide acceptable predictions.

A simple method for low-cost micro-forecasts was developed by Mammoli et al. [25]. The proposed system obtained images of sky above certain location and processed them using an ART-type neural network (Lateral Adaptive Priming Adaptive Resonance Theory) to provide a one-minute-ahead predictions of the power produced by the PV system. An emphasis was put on the simplicity and low-cost of the whole system rather on the comparison with other state-of-the-art predictors.

Bouzerdoum et al. [26] used a hybrid combination of Seasonal Auto-Regressive Integrated Moving Average (SARIMA) and SVM for short term forecasts of power output of a small PV power plant. It was shown that the hybrid method was better (in terms of prediction accuracy) than both individual methods. In [27], Qu et al. used a combination of ARMA and SVM to correct NWP errors of wind speed forecasting. The ability of SVM to correct wind speed forecasts from raw NWP data and from data corrected by ARMA was compared and it was concluded that the latter yielded better corrections (i.e. more accurate forecasts).

Another recent hybrid model for wind and PV power production forecasting is due to Quan et al. [28]. The algorithm used a hybrid evolutionary-neuro-fuzzy system to provide accurate hourly forecasts of the power produced by PV and wind components of the system. It was shown, among others, that the wind speed and henceforth wind power was harder to predict.

Ishak et al. [29] used a number of techniques to correct the errors in wind speed prediction provided by the MM5 numerical model. The authors used multiple linear regression, non-linear regression, ANN, and SVM to reduce error in MM5 wind speed forecast on the basis of other parameters estimated by the MM5 (i.e. temperature, relative humidity, solar radiation, and rainfall) and evaluated them on data obtained on a single location in Somerset, Southwest of England. It was shown that SVM performs better than ANN and simpler models. The simple (regression) models, however, were easier to implement and parametrize.

IV. Support vector regression

Support vector regression (SVR) is an extension of support vector machines (SVM), a family of popular supervised machine learning tools based on statistical learning theory originally proposed by Vapnik [30], [31].

A. SVM algorithm

SVM were designed to find an optimal directed hyperplane separating two non-overlapping classes of data with the help of support vectors (i.e. the points in the data closest to the separating hyperplane) [31]. However, later extensions enabled the SVM to learn and classify multiple classes of data, overlapping classes, and noisy data by the introduction of slack variables ξ_i, ξ_i^* that enable soft-margin classifiers [32], [31].

The SVM uses a linear separating hyperplane to construct a classifier with maximum margin by the means of constrained non-linear optimization [30]. Data that is not linearly separable can be processed by the SVM with the help of kernel substitution, i.e. a translation of input data to a high-dimensional feature space where it might be linearly separable [33], [31]. The SVM combines both, success in practical applications and well-established theory.

The basic SVM for binary classification aims to learn a decision function [31]

$$f(\mathbf{x}) = \mathrm{sign}(\mathbf{w} \cdot \mathbf{x} + b), \tag{1}$$

where $\cdot$ is dot product, $\mathbf{x}$ is the set of input data vectors (points) $x_1, x_2, \ldots, x_m$, $f(\mathbf{x})$ is the vector of corresponding labels $y_1, y_2, \ldots, y_m$, subject to $y_i = \pm 1$, sign is the signum function, and $\mathbf{w}$ is the vector of weights. In a geometrical representation, the hyperplanes $\mathbf{w} \cdot \mathbf{x} + b = 1$ and $\mathbf{w} \cdot \mathbf{x} + b = -1$ are called canonical hyperplanes and the area between them margin band. Maximizing the margin (i.e. finding optimal hyperplane) involves maximization of the function

$$W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \tag{2}$$

subject to

$$\alpha_i \geq 0, \qquad \sum_{i=1}^{m} \alpha_i y_i = 0, \tag{3}$$

where K is a kernel used for mapping of input data to high-dimensional feature space (kernel substitution) and α_i, α_j are

Lagrange multipliers. Bias b is given by [31]

$$b = -\frac{1}{2}\left[\max_{\{i|y_i=-1\}}\Big(\sum_{j=1}^{m}\alpha_j y_j K(x_i, y_i)\Big) + \min_{\{i|y_i=1\}}\Big(\sum_{j=1}^{m}\alpha_j y_j K(x_i, y_i)\Big)\right]. \tag{4}$$

B. SVR algorithm

In contrast with SVM, SVR aims to learn the mapping of data to real-valued labels [31], [32], [34]. The ϵ-SVR algorithm aims at learning a function $f(x_i)$ that has at most ϵ deviation from corresponding y_i [34]

$$f(\mathbf{x}) = \mathbf{w}\cdot\mathbf{x} + b \tag{5}$$

and leads to maximization of linear or quadratic ϵ-insensitive loss function. The linear ϵ-insensitive loss function is given by [31]

$$W(a, a^*) = \sum_{i=1}^{m} y_i(\alpha_i - \alpha_i^*) - \epsilon\sum_{i=1}^{m}(\alpha_i + \alpha_i^*) - \frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) \tag{6}$$

subject to

$$\sum_{i=1}^{m}\alpha_i = \sum_{i=1}^{m}\alpha_i^*, \quad \alpha_i, \alpha_i^* \in [0, C] \tag{7}$$

The ϵ-SVR can be visualised as a tube around hypothesis function which outlines training errors from valid training points.

V. Experimental Results

A. Models and Data

Four different numerical weather prediction (NWP) model versions have been chosen based on the previous experience that each of the configurations, and selection of physical parameterizations can give widely differing forecasts of cloudiness and solar radiation. Performance of particular NWP configuration depends on a type of weather situation and each NWP model can be better that the others based on the season of year, atmospheric circulation type, vertical stability of the atmosphere etc.

Following model versions and configurations have been used:

- *Model 1 (MM5_36):* MM5 version 3.6, RRTM radiation scheme, Grell cumulus parameterization, 26 vertical levels
- *Model 2 (MM5_37):* MM5 version 3.7, RRTM radiation scheme, Grell cumulus parameterization, 31 vertical levels
- *Model 3 (WRF_22):* WRF version 2.2, Dudhia radiation scheme, Kain-Fritsch cumulus parameterization, 39 vertical levels
- *Model 4 (WRF_34):* WRF version 3.4, Goddard radiation scheme, Grell cumulus parameterization, 39 vertical levels

TABLE I

ϵ-SVR parameters used for irradiance predictions correction.

Parameter	Value
algorithm	ϵ-SVR
kernel	radial basis function $e^{-\gamma\|u-v\|^2}$
loss function parameter ϵ	$1e^{-6}$
termination tolerance p	$1e^{-6}$
cost parameter C	2500
kernel function parameter γ	1

Each of the models was run for the same area and time period. Two nested domains have been used covering central Europe in 27 km and 9 km resolution. Simulation period was 8 months - from May to December 2011.

The simulation has been reproduced in genuine forecast regime. Predictions were initialized from NCEP GFS global model and predictions have been evaluated for the 24-48 hours forecast horizon.

Measurements of horizontal solar irradiance originate from the network of 51 stations operated by the Czech Hydrometeorological Institute. Compared quantity corresponds to average hourly irradiance (in Wm^{-2}) and its model equivalent.

B. Experiments

The ϵ-SVR algorithm was used to learn models for accurate correction of downward short-wave radiation predictions provided by the NWP models. The data for each station was divided into training (first 50% of the data) and test (second 50% of the data) subset. Both the test and training data for each station spanned multiple full days of model forecasts (e.g. traning data from AKOBA station covered the period from 06/03/2011 5:00AM to 08/04/2011 11:00 AM and the test data covered the period from 08/04/11 12:00 PM to 12/31/11 3:00 PM). A separate correction model was developed for each station by ϵ-SVR.

Each correction model was learned from the training subset and evaluated on the full data set (i.e. training and test data set) in order to provide a forecast that can be compared with the NWP models that always spanned the full data set. The parameters of the ϵ-SVR algorithm, shown in table I, were set based on initial trials and previous experience with the method.

C. Results Description

The results of the conducted experiments are shown in table II and table III respectively. Table II shows the average absolute mean error of forecasts provided by each method for every weather station. The SVR-based correction is compared with raw (uncorrected) NWP forecasts and forecasts with corrected bias (see [35] for details on irradiance forecast bias correction). It can be seen that the SVR-corrected forecasts were in all cases more accurate than both the original forecasts provided by the NWP models alone and bias-corrected forecasts.

TABLE II

MEAN ABSOLUTE ERROR OF IRRADIANCE PREDICTIONS BY EACH METHOD FOR EACH STATION [Wm^{-2}]. LOWER IS BETTER.

Station	ϵ-SVR	Forecast model (uncorrected)				Forecast model (bias corrected)			
		MM5_36	MM5_37	WRF_22	WRF_34	MM5_36	MM5_37	WRF_22	WRF_34
AKOBA	91.190	119.439	115.841	109.636	129.971	92.0211	107.224	82.8138	88.6244
ALIBA	85.2342	119.453	107.848	98.265	111.582	86.9746	100.651	78.5866	84.3815
AREPA	78.8115	102.503	104.493	110.895	139.761	90.5438	104.913	84.4662	91.1163
ASMIA	93.1683	144.863	128.092	105.899	115.708	92.5371	103.668	84.036	89.4018
ASUCA	63.1851	96.4012	103.259	113.17	142.832	93.1717	109.95	85.6494	93.9905
BMISA	68.8534	96.7447	97.5955	103.795	128.158	97.3625	109.163	85.3875	85.4922
CHVOA	77.9054	108.267	101.725	107.695	115.371	100.582	110.701	83.7971	91.1282
EPAUA	73.5852	108.93	96.4919	95.092	112.376	89.4	99.7781	78.5812	81.2312
HHKBA	72.6831	116.694	99.3801	87.6625	107.618	92.9176	101.511	79.4897	82.5607
HKRYA	105.022	140.574	125.211	115.494	119.53	90.4256	100.272	85.4038	85.7559
HSERA	76.732	98.3612	96.6656	105.85	129.141	86.2947	102.064	83.8034	90.7001
HTRMA	92.1769	137.45	118.121	104.832	114.036	92.343	99.1507	84.7813	86.4406
JJIHA	75.2275	109.714	103.129	95.4972	109.362	93.3919	106.518	83.3158	87.5967
KCHMA	95.378	154.36	135.411	110.03	109.222	97.4519	110.522	92.204	97.0807
KKVMA	82.3831	113.398	102.991	95.9768	117.991	87.3257	99.7186	79.9588	88.4805
KPRBA	76.4059	106.483	96.5602	95.5595	104.446	86.4606	102.733	89.5983	101.261
KSOMA	85.7751	122.811	109.483	99.072	108.768	88.7333	102.897	84.3487	92.7993
LCLMA	80.459	111.544	101.04	99.6344	118.901	85.059	100.361	82.7129	87.8219
LFRUA	68.6459	101.125	89.0805	87.6492	101.328	84.9175	98.6977	74.8309	86.6284
LLIMA	73.7993	94.4371	97.4398	107.471	130.498	89.1884	105.183	86.6645	95.3464
LSOUA	85.0755	117.591	102.453	102.166	109.631	88.6277	100.531	89.3338	100.69
MJESA	82.5898	105.537	102.507	101.755	122.068	87.915	101.686	78.4237	89.2933
MPRRA	78.0429	103.117	102.477	98.3812	114.551	92.8973	106.82	77.1358	77.7275
MPSTA	68.7543	95.7844	97.3943	112.17	141.555	96.572	105.108	83.4614	84.6737
PPLVA	85.6755	112.784	108.762	102.196	118.8	86.1388	106.882	84.2071	85.6631
SBERA	59.7994	89.6569	107.764	135.7	169.75	91.0222	111.368	93.0668	100.89
SKLMA	76.9878	110.798	102.372	91.1876	114.032	85.2268	102.876	75.3562	81.7533
SMBOA	93.3806	147.002	136.964	109.31	112.631	90.1176	108.88	88.2464	89.181
SONRA	66.626	92.7768	101.064	115.305	144.587	90.0134	108.677	88.3429	94.5302
TBKRA	89.7243	109.319	109.93	116.358	135.592	97.3676	106.557	87.2725	97.3874
TKARA	98.1242	113.018	117.759	128.239	153.463	96.5927	112.573	93.7163	96.9476
TOFFA	66.281	95.6009	94.4563	95.4963	122.39	91.3276	103.316	81.164	85.1841
TOVKA	79.3879	113.349	100.175	90.9064	107	89.9584	98.2172	76.3266	81.7606
TSTDA	68.6323	97.0063	95.1463	95.1218	126.169	90.7928	104.808	80.8334	87.945
UCHMA	89.6388	126.067	112.835	102.236	120.468	87.4439	95.8244	74.2407	80.4997
UDCMA	67.4403	87.8871	92.8778	106.848	135.657	84.7573	101.157	86.3081	91.2823
UKRUA	71.1658	93.2403	101.304	111.62	147.2	83.5942	103.016	82.9274	88.3344
ULOMA	83.7468	109.617	102.124	101.786	131.136	87.3762	97.0025	75.264	83.306
ULTTA	74.3828	104.243	100.361	95.5	123.805	84.9657	103.459	78.0534	86.2746
UMOMA	63.6703	96.7313	104.455	118.143	161.881	95.674	109.703	83.8999	95.0267
URVHA	91.4067	154.489	126.806	106.95	105.418	94.2099	97.5649	80.9421	85.9354
USNZA	92.5786	130.978	119.461	109.183	113.263	86.4077	99.1341	85.4203	85.5433
UTPMA	70.8883	93.0521	98.9284	114.627	152.578	85.2432	100.157	79.0647	86.4485
UTUSA	87.2568	129.926	113.176	98.3457	110.858	87.4221	97.5307	74.3886	79.6691
UULKA	77.5282	125.159	107.109	93.4715	103.128	87.042	99.6318	74.2541	75.6839
UULMA	67.4045	98.509	100.45	106.644	137.703	89.3182	103.306	78.5477	83.0732
UVALA	77.7562	115.175	103.451	94.6575	107.237	88.7364	101.634	80.827	85.46
XMOLA	45.3201	147.5	193.539	235.372	270.07	133.227	174.521	141.843	134.73
XMOUA	74.283	102.615	97.513	105.283	121.964	95.7186	108.698	94.2111	103.601
ZSNVA	86.846	118.89	121.546	137.75	146.016	92.0469	108.182	85.765	87.4938
ZZLNA	102.501	120.788	115.394	117.957	132.635	96.8513	109.342	84.6694	85.83

TABLE III

MEAN ABSOLUTE ERROR OF IRRADIANCE PREDICTIONS BY EACH METHOD FOR ALL STATIONS [Wm^{-2}]. LOWER IS BETTER.

	Forecast model (uncorrected)				Forecast model (bias corrected)			
ϵ-SVR	MM5_36	MM5_37	WRF_22	WRF_34	MM5_36	MM5_37	WRF_22	WRF_34
79.7937	112.249	106.525	105.222	124.238	90.4021	103.838	82.9613	88.5457

The mean absolute error of irradiance prediction for all stations is shown in table III. Again, it can be seen that the SVR-corrected forecasts yielded lower mean absolute error than individual NWP models.

Student's paired t-test [36] at significance level $\alpha = 0.05$ was performed to compare the average error of predictions by ϵ-SVR and forecast models for every weather station. The test confirmed that the differences between predictions by ϵ-SVR and each forecast model (both uncorrected and corrected) are statistically significant with p-value lower than 0.001 in all cases.

VI. CONCLUSIONS

The results of the numerical experiments suggest that support vector regression can be used to combine and correct downward short-wave radiation predictions provided by multiple NWP models, and achieve a significantly lower mean absolute prediction error. The mean absolute error of ϵ-SVR corrected forecasts is approx. 29% lower than the mean absolute error of MM5_36 forecasts (12% lower than bias-corrected MM5_36 forecasts), 25% lower than the mean absolute error of MM5_37 forecasts (23% lower than bias corrected MM5_37), 24% lower than the mean absolute error of WRF_22 forecasts (4% lower than bias-corrected WRF_22 forecasts), and more than 35% lower than the error of WRF_34 forecasts (10% lower than bias-corrected WRF_34 forecasts).

The research presented in this work will continue in several directions. Other SVR variants as well as other machine-learning meta-heuristics can be evaluated for their use in multi-model irradiation prediction. Moreover, their results can be compared with other recently proposed heuristic methods for the same task (e.g. those by Eben et al. [35]).

REFERENCES

[1] G. Reikard, "Predicting solar radiation at high resolutions: A comparison of time series forecasts," *Solar Energy*, vol. 83, no. 3, pp. 342 – 349, 2009.

[2] A. Sfetsos and A. Coonick, "Univariate and multivariate forecasting of hourly solar radiation with artificial intelligence techniques," *Solar Energy*, vol. 68, no. 2, pp. 169 – 178, 2000.

[3] A. Mellit, "Artificial intelligence technique for modelling and forecasting of solar radiation data: a review," *Int. J. Artif. Intell. Soft Comput.*, vol. 1, no. 1, pp. 52–76, Nov. 2008.

[4] A. Hammer, D. Heinemann, E. Lorenz, and B. Lückehe, "Short-term forecasting of solar radiation: a statistical approach using satellite data," *Solar Energy*, vol. 67, no. 1-3, pp. 139 – 150, 1999.

[5] W. P. Menzel, "Cloud Tracking with Satellite Imagery: From the Pioneering Work of Ted Fujita to the Present." *Bulletin of the American Meteorological Society*, vol. 82, pp. 33–48, Jan. 2001.

[6] A. Hammer, D. Heinemann, C. Hoyer, R. Kuhlemann, E. Lorenz, R. Müller, and H. G. Beyer, "Solar energy assessment using remote sensing technologies," *Remote Sensing of Environment*, vol. 86, no. 3, pp. 423 – 432, 2003.

[7] G. Grell, J. Dudhia, and D. Stauffer, *A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5)*, ser. NCAR technical notes. Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, 1994.

[8] W. C. Skamarock, J. B. K. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang, W. Wang, and J. G. Powers, *A Description of the Advanced Research WRF Version 3*, ncar/tn-47 ed. Boulder, CO: NCAR, 2008, no. June. [Online]. Available: http://www.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf

[9] R. J. Zamora, E. G. Dutton, M. Trainer, S. A. McKeen, J. M. Wilczak, and H. Yu-Tai, "The accuracy of solar irradiance calculations used in mesoscale numerical weather prediction." *Monthly Weather Review*, vol. 133, no. 4, pp. 783 – 792, 2005.

[10] J. A. Ruiz-Arias, D. Pozo-Vázquez, N. Sánchez-Sánchez, J. P. Montávez, A. Hayas-Barru, and J. Tovar-Pescador, "Evaluation of two MM5-PBL parameterizations for solar radiation and temperature estimation in the South-Eastern area of the Iberian Peninsula," *Nuovo Cimento C Geophysics Space Physics C*, vol. 31, pp. 825–842, Sep. 2008.

[11] P. Blanc, B. Gschwind, M. Lefévre, and L. Wald, "The helioclim project: Surface solar irradiance data for climate applications," *Remote Sensing*, vol. 3, no. 2, pp. 343–361, 2011.

[12] P. Juruš, K. Eben, J. Resler, P. Krč, I. Kasanický, E. Pelikán, M. Brabec, and J. Hošek, "Estimating climatological variability of solar energy production," *Solar Energy*, vol. 98, Part C, no. 0, pp. 255 – 264, 2013.

[13] M. Lydia and S. Kumar, "A comprehensive overview on wind power forecasting," in *IPEC, 2010 Conference Proceedings*, 2010, pp. 268–273.

[14] F. Bellocchio, S. Ferrari, M. Lazzaroni, L. Cristaldi, M. Rossi, T. Poli, and R. Paolini, "Illuminance prediction through svm regression," in *Environmental Energy and Structural Monitoring Systems (EESMS), 2011 IEEE Workshop on*, 2011, pp. 1–5.

[15] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy, "Predicting solar generation from weather forecasts using machine learning," in *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, 2011, pp. 528–533.

[16] W. Ji and K. C. Chee, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN," *Solar Energy*, vol. 85, no. 5, pp. 808 – 817, 2011.

[17] L. Prokop, S. Mišák, T. Novosád, P. Krömer, J. Platoš, and V. Snášel, "Photovoltaic power plant output estimation by neural networks and fuzzy inference," in *Intelligent Data Engineering and Automated Learning - IDEAL 2012*, ser. Lecture Notes in Computer Science, H. Yin, J. A. Costa, and G. Barreto, Eds. Springer Berlin Heidelberg, 2012, vol. 7435, pp. 810–817.

[18] A. Zarnani, P. Musilek, X. Shi, X. Ke, H. He, and R. Greiner, "Learning to predict ice accretion on electric power lines," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 3, pp. 609 – 617, 2012.

[19] C. Voyant, M. Muselli, C. Paoli, and M.-L. Nivet, "Hybrid methodology for hourly global radiation forecasting in mediterranean area," *Renewable Energy*, vol. 53, no. 0, pp. 1 – 11, 2013.

[20] A. Mellit, S. Sağlam, and S. Kalogirou, "Artificial neural network-based model for estimating the produced power of a photovoltaic module," *Renewable Energy*, vol. 60, no. 0, pp. 71 – 78, 2013.

[21] S. Chen, H. Gooi, and M. Wang, "Solar radiation forecast based on fuzzy logic and neural networks," *Renewable Energy*, vol. 60, no. 0, pp. 195 – 201, 2013.

[22] J. Yan, Y. Liu, S. Han, and M. Qiu, "Wind power grouping forecasts and its uncertainty analysis using optimized relevance vector machine," *Renewable and Sustainable Energy Reviews*, vol. 27, no. 0, pp. 613 – 621, 2013.

[23] H. Zhu and J. Yao, “A power prediction method for photovoltaic power station based on neutral network using numerical weather information,” *Advanced Materials Research*, vol. 724 - 725, pp. 3 – 9, 2013.
[24] L. Prokop, S. Misak, V. Snasel, J. Platos, and P. Kromer, “Supervised learning of photovoltaic power plant output prediction models,” *Neural Network World*, vol. 23, no. 4, pp. 321–338, 2013.
[25] A. Mammoli, A. Menicucci, T. Caudell, A. Ellis, S. Willard, and J. Simmins, “Low-cost solar micro-forecasts for pv smoothing,” in *Technologies for Sustainability (SusTech), 2013 1st IEEE Conference on*, 2013, pp. 238–243.
[26] M. Bouzerdoum, A. Mellit, and A. M. Pavan, “A hybrid model (sarima-svm) for short-term power forecasting of a small-scale grid-connected photovoltaic plant,” *Solar Energy*, vol. 98, Part C, no. 0, pp. 226 – 235, 2013.
[27] G. Qu, J. Mei, and D. He, “Short-term wind power forecasting based on numerical weather prediction adjustment,” in *Industrial Informatics (INDIN), 2013 11th IEEE International Conference on*, 2013, pp. 453–457.
[28] D. M. Quan, E. Ogliari, F. Grimaccia, S. Leva, and M. Mussetta, “Hybrid model for hourly forecast of photovoltaic and wind power,” in *Fuzzy Systems (FUZZ), 2013 IEEE International Conference on*, 2013, pp. 1–6.
[29] A. Ishak, R. Remesan, P. Srivastava, T. Islam, and D. Han, “Error correction modelling of wind speed through hydro-meteorological parameters and mesoscale model: A hybrid approach,” *Water Resources Management*, vol. 27, no. 1, pp. 1–23, 2013.
[30] V. Kecman, “Support vector machines an introduction,” in *Support Vector Machines: Theory and Applications*, ser. Studies in Fuzziness and Soft Computing, L. Wang, Ed. Springer Berlin / Heidelberg, 2005, vol. 177, pp. 605–605.
[31] C. Campbell and Y. Ying, “Learning with support vector machines,” *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 5, no. 1, pp. 1–95, 2011.
[32] L. H. Hamel, *Knowledge Discovery with Support Vector Machines*. New York, NY, USA: Wiley-Interscience, 2009.
[33] R. Herbrich, *Learning Kernel Classifiers: Theory and Algorithms (Adaptive Computation and Machine Learning)*. The MIT Press, Dec. 2001.
[34] A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, Aug. 2004.
[35] K. Eben, J. Resler, P. Krc, P. Jurus, and E. Pelikan, “Multi-model prediction of downward short-wave radiation,” in *12th EMS / 9th ECAC, EMS Annual Meeting EMS2012*, vol. 9, 2012.
[36] G. Box, J. Hunter, and W. Hunter, *Statistics for experimenters: design, innovation, and discovery*, ser. Wiley series in probability and statistics. Wiley-Interscience, 2005.