Optimal Granule-Based PIs Construction for Solar Irradiance Forecast

Songjian Chai, Zhao Xu, *Senior Member, IEEE*, and Wai Kin Wong

***Abstract*—This letter proposes a novel granule computing-based framework for prediction intervals (PIs) construction of solar irradiance time series that has significant impacts on solar power production. Distinguished from most existing methods, the new framework can address both stochastic and knowledge uncertainties in constructing PIs. The proposed method has proved to be highly effective in terms of both reliability and sharpness through a real case study using measurement data obtained from Hong Kong Observatory.**

***Index Terms*—Granular neural network, prediction intervals, random vector forward link (RVFL), solar irradiance forecasting.**

I. Introduction

PREDICTING solar irradiance is essential but fairly challenging to estimate solar power production. To effectively quantify uncertainties in solar irradiance time series, many prediction interval-based approaches have been developed. The stochastic uncertainty and knowledge uncertainty are always regarded as the two main contributors to the prediction errors in a modeling system. The former describes the inherent variability of the observed values due to the natural physical phenomenon, measured error, device failure, and the like, while the latter reveals the uncertainty in knowledge transfer, such as imperfect representation of processes in a model, as well as the imperfect knowledge of the parameters associated with these processes. Most existing approaches rely on point forecasting error information to establish prediction intervals (PIs). Recently, some direct and non-parametric methods for optimal PIs construction [1] without the need of forecasting error information were widely reported. It should be highlighted that these methods focus on the stochastic errors of various time series, while the knowledge uncertainty is not considered at all. Further, the PIs are constructed based on crisp or deterministic inputs without explicitly recognizing the variability involved in the observed dataset.

Lately, information granule has emerged as a new and promising way to deal with the situations characterized by many underlying uncertainties. By granulating model inputs and parameters through, e.g., interval, rough set, and fuzzy set, an information granularity-based input-output mapping can be established, which can effectively reflect the values of various attributes in the process of knowledge transfer. Subsequently the granular outputs tailored according to such PIs at certain confidence level can be reliably constructed. In this letter, a novel granular framework for PIs construction is developed based on random vector forward link (RVFL) network, where the parameters are learned by the particle swarm optimization (PSO) technique which optimizes the performance in terms of both coverage probability and sharpness of the constructed granular outputs (i.e., PIs) with respect to the real targets. The measured solar irradiance data with 1-min resolution by Hong Kong Observatory (HKO) in 2012 are used in the case study.

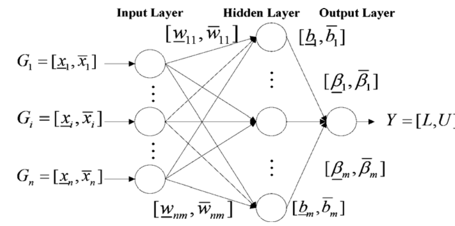

Fig. 1. Architecture of granular neural network.

II. Methodology

The interval-valued representation formalism is adopted as the granular form to quantify the stochastic uncertainty and knowledge uncertainty as it is the most simplified form to comprehend and expressed by a domain expert [2].

A. Granulation of Original Time Series

The solar irradiance time series with 1-min resolution are found highly volatile and intermittent. Nevertheless, interval-based information granule can capture this variability and abstract the complexity of such chaotic time series by simply constructing a max-min bound over certain level of temporal segments. In such a way, the original numerical time series are segmented into uniform time windows and becomes a granular time series $\boldsymbol{G}=\{[\underline{x}_i,\bar{x}_i], i=1,2,\ldots N\}$, where $\underline{x}_i$ and $\bar{x}_i$ is the minimum and maximum values over each time window.

B. Granular RVFL Network and Training Strategy

The general architecture of granular neural networks has been reported in [2] and [3], as illustrated in Fig. 1, where the traditional numeric parameters are replaced by interval values to deal with the model uncertainty.

Traditionally, the granular parameters are trained using an optimization tool aiming at minimizing the interval error function [2]. However, the quality and efficiency of the underlying optimization might be negatively affected by the large number of model parameters (double of the number of the original numeric parameters) to be optimized and the complex interval arithmetic. In order to tackle this issue, a granular neural network on the basis of RVFL network is proposed here, which could immensely benefit from the wealth of well-established learning strategies of RVFL network. The entire training process is outlined as follows. The upper bound values $\bar{x}_i$ and lower bound values $\underline{x}_i$ are extracted from the granules to train two numerical RVFL networks with the same structure and same values of input weights and bias, respectively. Hence, two sets of numerical output weights $\bar{\beta}_{ini}$ and $\underline{\beta}_{ini}$ corresponding to upper and lower dataset are determined. Afterwards, a spread percentage

Manuscript received June 17, 2015; accepted July 31, 2015. Date of publication September 15, 2015; date of current version May 02, 2016. This work was supported in part by Hong Kong RGC Theme Based Research SchemeGrants No. T23-407/13N and T23-701/14N. Paper no. PESL-00091-2015.

S. Chai and Z. Xu are with the Department of Electrical Engineering, The Hong Kong Polytechnic University, Hung Hom, Hong Kong (email: sj.chai@polyu.edu.hk; zhao.xu@polyu.edu.hk).

W. K. Wong is with Hong Kong Observatory, Hong Kong. (email:wkwong@hko.gov.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2015.2473097

vector $\boldsymbol{S} = \{[\bar{s}_i, \underline{s}_i], i = 1, 2, \ldots m\}$ is introduced, where s_i ranges in $[-100\%, 100\%]$ and is assigned to each obtained output weight value, which can be formulated as

$$[\underline{\beta}_i, \bar{\beta}_i] = [\underline{\beta}_{i,ini} + \underline{s}_i|\underline{\beta}_{i,ini}|, \bar{\beta}_{i,ini} + \bar{s}_i|\bar{\beta}_{i,ini}|]. \quad (1)$$

Finally, the whole learning process is well articulated and translated into the optimization stage only concerned with the design asset s_i. In this sense, the number of parameters to be optimized is significantly reduced and the searching space is narrowed. Furthermore, the topology of traditional neural network is augmented and simplified since the granular values of input weights and bias can be regarded as crisp values, which are the same as those in the well-trained numerical RVFL networks.

C. Construction of Optimal PIs

The granular outputs (i.e., PIs) derived from the granular RVFL are evaluated by the interval score [1] from perspectives of both reliability and sharpness.

1) Reliability: The reliability is evaluated by the value of average coverage error (ACE), which indicates the difference between the PICP (the probability of the true values lie in the PIs) and PINC (the nominal confidence level $100(1-\alpha)$ %), defined as

$$\text{ACE} = \text{PICP} - \text{PINC} \quad (2)$$

$$\text{PICP} = \frac{1}{N_t}\sum_{i=1}^{N_t} c_i \quad (3)$$

where N_t is the number of targets, and t_i is the real target, which is assumed as the mean value of each time window here. If t_i is included in the PI, $c_i = 1$; otherwise, $c_i = 0$. ACE should be diminished as small as possible to obtain higher reliability.

2) Sharpness: The sharpness is the other vital performance index reflecting the width of the resultant PIs in the evaluation process. The interval score can assess the overall skill of the constructed PIs by taking the sharpness into account, expressed as

$$Sc_i = \begin{cases} -2\alpha(U_i - L_i) - 4(L_i - t_i), \text{if } t_i < L_i \\ -2\alpha(U_i - L_i), \text{if } t_i \in [L_i, U_i] \\ -2\alpha(U_i - L_i) - 4(t_i - U_i), \text{if } t_i > U_i \end{cases}$$

$$\text{Score} = \frac{1}{N_t}\sum_{i=1}^{N_t} Sc_i. \quad (4)$$

$Y_i = [L_i, U_i]$ is the granular output with L_i and U_i as the lower and upper bound, respectively. Intuitively, higher reliability can be obtained at expense of wider PIs; hence, these two conflicting objectives should be combined together to establish the cost function in (5) [1], which is to be minimize by PSO:

$$\boldsymbol{s}_{\text{opt}} = \arg\min_{\boldsymbol{s}}(|\text{ACE}| + |\text{Score}|). \quad (5)$$

III. Experimental Results

The proposed approach is compared with our previous direct PIs construction method [1], where the inputs and model parameters are both crisp; 1-min irradiance data measured in King's Park, Hong Kong, in 2012 are used as dataset. To have a consistent forecasting duration for all days, the nighttime data is ignored. PIs with PINC of 90% are constructed over 10-min time scale by one-time-step forecasting. The study was conducted in different seasons and the whole year. The obtained results are presented in Table I and part of the PIs derived by both approaches are displayed in Fig. 2.

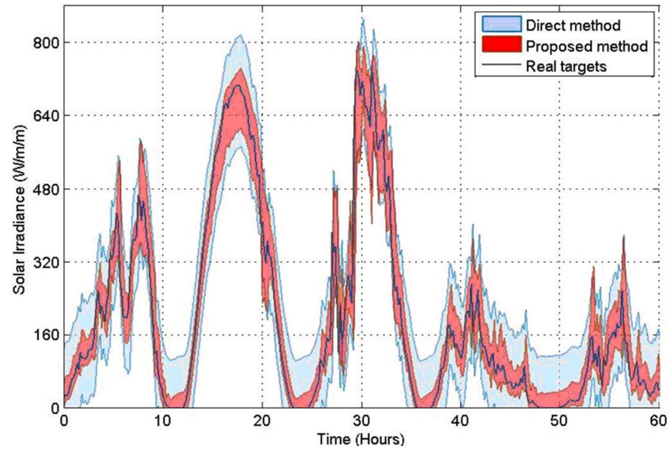

Fig. 2. Optimal PIs with PINC of 90% under 10-min time scale.

TABLE I
Comparison of Solar Irradiance Prognosis Results

Approach	Proposed method			Direct method		
Seasons	PICP	Score	PINAW	PICP	Score	PINAW
Spring	87.80%	-4.44%	20.44%	80.70%	-7.19%	19.01%
Summer	90.10%	-3.66%	19.04%	90.80%	-6.09%	28.82%
Autumn	90.90%	-2.77%	16.99%	92.30%	-4.04%	26.81%
Winter	88.60%	-2.65%	12.27%	85.50%	-3.67%	14.13%
Whole year	91.20%	-3.26%	16.94%	92.70%	-5.04%	26.57%

PINAW denotes the normalized average width of granular outputs, which is defined as

$$\text{PINAW} = \frac{1}{N_t R}\sum_{i=1}^{N_t}(U_i - L_i) \quad (6)$$

where R is the maximum range of the targets. It is clearly seen from the table that the proposed granule-based approach could significantly increase the reliability as compared to the traditional optimal direct PIs construction method, particularly in Spring and Winter, in which the irradiance shows high variability and thus gives rise to the difficulty of predicting. In other cases, the sharpness of PIs is considerably reduced while maintaining high reliability. The interval score reflects the overall performance encompassing both reliability and sharpness, the improvements are within 27%–39% with respect to the direct method depending on the cases, therefore demonstrating the effectiveness of the proposed method.

IV. Conclusion

The conceptualization of information granulation is first applied to probabilistic interval forecasting of solar irradiance. The variability of the raw irradiance time series, as well as the uncertainties in knowledge transfer, are well quantified associated with the constructed PIs. Future work is underway to further enhance the performance of the proposed approach.

References

[1] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1166–1174, May 2014.
[2] M. G. C. A. Cimino, B. Lazzerini, F. Marcelloni, and W. Pedrycz, "Genetic interval neural networks for granular data regression," *Inf. Sci.*, vol. 257, pp. 313–330, Feb. 1, 2014.
[3] M. Song and W. Pedrycz, "Granular neural networks: Concepts and development schemes," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 24, no. 4, pp. 542–553, Apr. 2013.