

Bayesian regularisation network-based models for estimating daily global solar radiation

R.YACEF

Renewable Energy Laboratory
Jijel University,
Jijel, Algeria
rimayacef@yahoo.fr

A.MELLIT

Renewable Energy Laboratory
Jijel University,
Jijel, Algeria
adelmellit@yahoo.com

Abstract— In this paper a Bayesian learning approach for the neural network (BNN) is developed to estimate daily global solar radiation (DGSR). A Bayesian framework for neural network named as automatic Relevance determination method (ARD) was employed to obtain the relative relevance of a set of atmospheric variables used for estimating DGSR. In addition, we analysed the viability of this novel technique applied to select the optimum input parameters to the neural network. A multi-layer feed-forward perceptron was trained using atmospheric data. The results show that the sun shine duration and air temperature are the most relevant input variables to the neural network.

Keywords—daily global solar radiation; neural network; bayesian neural network; atmospheric variables

I. INTRODUCTION

Global solar irradiation is considered as the most significant parameter in meteorology, solar conversion and renewable energy applications, particularly for the sizing and modelling of photovoltaic (PV) systems. Several empirical techniques [1-4] have been developed in the literature and show the possibility to find a correlation between solar irradiation and other meteorological parameters such as air temperature, sunshine duration, relative humidity, cloud cover, etc.

In the recent years, the intelligent techniques (neural network, fuzzy logic, and hybrid neural network) are widely used to solve real problems in the case where classical methods are insufficient. Solar irradiation prediction is one of the domains where the intelligent techniques were successfully used, and particularly Artificial Neural Networks (ANN) [5].

Mellit et al. [6] proposed a simplified hybrid model for generating sequences of global DGSR, which combine a neural network and Markov chain. This model is called ANN-MTM (Markov Transition Matrix). The input parameters of the proposed model are the geographical coordinates while the outputs are the DGSR. It can be used for generating sequences of solar radiation at long term and it was applied for Algeria. A correlation coefficient ranging from 90% to 92% has been obtained.

Rehman and Mohandes [7] used an ANN for the estimation of GSR for Abha city in Saudi Arabia as a function of air temperature and relative humidity. Three

different architectures were used in this case, obtained results show that NNs are capable to estimate the GSR from temperature and relative humidity in the location where only these two parameters are available.

Lauret et al., [8] proposed two-step approaches to building an ANN model to modeling the direct solar radiation. The first one deals with a probabilistic interpretation of the ANN learning by using Bayesian techniques, the second step consists of using a new sensitivity analysis based pruning method to choose the optimum ANN structure. It was found that the combination of the two approaches make the practical implementation of the Bayesian techniques more reliable.

The objective of this study is to evaluate the application of Bayesian Neural Network (BNN) for predicting (DGSR). BNN models can effectively reduce the over-fitting phenomenon while still keep the strong nonlinear approximation ability of neural networks [9]. In this work, the Bayesian framework for ANN, named as automatic relevance determination method (ARD) was employed for estimating (DGSR) and selecting the optimum input parameters of the neural network.

This paper is organized as: A database description is presented in section 2. A brief introduction to Bayesian artificial neural networks is given in Section 3. Results and discussion are given in Section 4.

II. BAYESIAN NEURAL NETWORK APPROACH

A. Principle of Bayesian NN learning: a probabilistic approach to NN learning

To improve the generalization capabilities of the conventional back-propagation algorithm, in [10,11] the authors proposed the use of Bayesian back propagation neural networks. The following briefly describes the Bayesian inference method. A more detailed review can be found in the works [10, 11]. In the Bayesian framework, the neural learning process is assigned a probabilistic interpretation. The regularized objective function is given by:

$$S(w) = \beta E_D + \alpha E_w \quad (1)$$

Where E_w are given by:

$$E_w(w) = 1/2 \sum_{i=1}^m w_i^2 \quad (2)$$

Where m is the total number of parameters in the network. α and β are termed hyper-parameters (regularization parameter). The objective, during training is to maximize the posterior distribution over the weights w , to obtain the most probable parameter values w_{MP} . The posterior distribution is then used to evaluate the predictions of the trained network for new values of the input variables. For a particular network A , trained to fit a dataset $D = \{x_i, t_i\}_{i=1}^N$, by minimizing an error function $S(w)$ given by Eq.1, Bayes' theorem can be used to estimate the posterior probability distribution $p(w/D, \alpha, \beta, A)$ for the weights as follows:

$$p(w/D, \alpha, \beta, A) = p(D/w, \beta, A) p(w/\alpha, A) / p(D/\alpha, \beta, A) \quad (3)$$

Where $p(w/\alpha, A)$ is the prior density, which represents our knowledge of the weights before any data is observed $p(D/w, \beta, A)$, is the likelihood function which represents a model for the noise process on the target data, $p(D/\alpha, \beta, A)$ is known as the evidence and is the normalization factor that guarantees that the total probability is 1. This is given by an integral over the weight space as giving in (4).

$$P(D/\alpha, \beta, A) = \int p(D/w, \beta, A) p(w, \alpha, A) \quad (4)$$

As can be observed in (4), in order to evaluate the posterior distribution, the expressions for the prior distribution and the likelihood function need to be defined. More detail can be found in [10].

B. Automatic Relevance Determination (ARD) process

With finite data set, random correlation between inputs and output are likely to occur. Conventional neural networks (even with regularization) do not set the coefficient for these irrelevant inputs to zero [10]. Thus, the irrelevant variables can deter the model performance, particularly when the variables are many and the data are few. The concept of relevance can be embodied into Bayesian learning by placing a prior Gaussian distribution on the weights. This implies that each input variable has its own prior variance parameter. Thus, with the ARD technique, a separate regularization coefficient is assigned to each input. More precisely, all weights related to the same input are controlled by the same hyper-parameter α_c . This hyper-parameter is associated with a Gaussian prior with zero mean and variance $1/\alpha_c$. After training, the weights with a large α_c (or, conversely, a small variance $1/\alpha_c$ are close to zero and have little influence on subsequent values. Consequently, one can state that the corresponding input is irrelevant and, therefore, can be eliminated. On the other hand a large variance indicates that the variable is important in explaining the data [10].

III. DATASET

The experimental dataset used in this work are available by the National Renewable Energy Laboratory (NREL). These data are DGSR, sunshine duration, air temperature and relative humidity. Fig.1 and Fig. 2 show the evolution of daily: global solar radiation, sunshine duration, air temperature and relative humidity received on a horizontal surface [12].

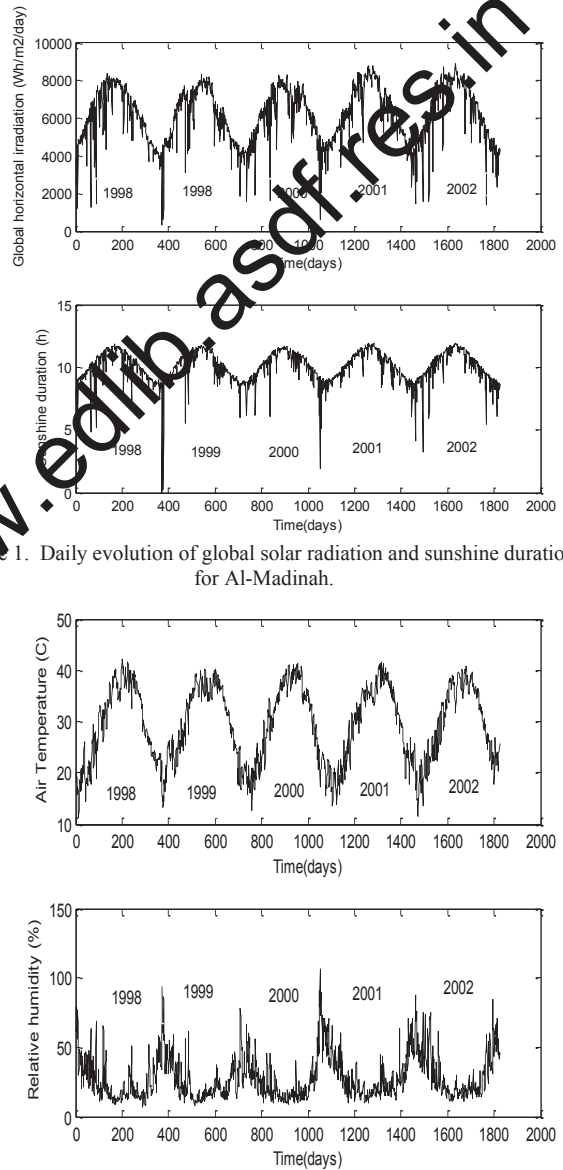


Figure 1. Daily evolution of global solar radiation and sunshine duration for Al-Madinah.

Figure 2. Daily evolution of air temperature and relative humidity for Al-Madinah.

IV. METHODOLOGY

In this study the Bayesian ARD technique is applied for selecting the more relevant input parameters affecting the global solar radiation. For this purpose, we have developed four BNN models: BNN2, BNN5, BNN7, and BNN11 by

using three input parameters: sunshine duration, air temperature and relative humidity and changing in this way the number of hidden units for each case (2, 5, 7 and 11) respectively and the number of the training cycle for all developed BNN models in order to see the influence of the number of hidden units and the training cycle on the selection of the optimum input parameters and on the estimation of the DGSR.

The accuracy of estimation was tested by calculating the mean bias error (MBE), root mean square error (RMSE) and the mean absolute error (MAE). Low values of RMSE and MAE are desirable. Positive MBE shows overestimation while negative MBE indicates underestimation. The MBE, RMSE and MAE are defined as in the following equations [12]:

$$MBE = \frac{1}{m} \sum_{i=1}^m e_i \tag{5}$$

$$RMSE = \left[\frac{1}{m} \sum_{i=1}^m e_i^2 \right]^{1/2} \tag{6}$$

$$MAE = \frac{1}{m} \sum_{i=1}^m |e_i| \tag{7}$$

Where e_i is the relative error between the measured and estimated value and m is the total number of observations. e_i is defined as follow :

$$e_i = (G_{im} - G_{ie}) / G_{ie} \times 100 \tag{8}$$

V. RESULTS AND DISCUSSION

The Bayesian ARD approach is applied to select the optimum input parameters and to modelling the DGSR. For this reason we have developed four Bayesian neural networks changing in the same time both training cycles (50 and 100), and the number of the hidden units 2,5,7, and 11 named as BNN2, BNN5, BNN7, BNN11 respectively.

The designed Bayesian models were implemented by using a specific MATLAB toolbox called BFLAB [13].

In order to find the optimal values of α as well as the optimum weight vector, the following iterative procedure was used to implement the Bayesian evidence training:

- The initial values for the hyper parameters α and β was set to 0.01 and 50 respectively.
- The weight in the network were initialised and drawn from the prior distribution defined by α .
- The weight optimisation was performed using the scale conjugate gradient algorithm to minimise the regularised error function. The total number of training cycles was set to 50 and 100. The tolerance of the weight was set to 10^{-7} .
- After each 50 and 100 training cycles the Gaussian approximation was used to compute the evidence for the hyper parameters.
- Steps 2, 3 and 4 were repeated until the hyper parameters and weights had converged.

Figure 3 shows the log-values of the hyper parameters versus training session for four multilayer perceptron with 2, 5, 7 and 11 hidden units, BNN2, BNN5, BNN7 and, BNN11 respectively. For each one 50 and 100 training cycles and the same training session equal to 50. With respect to Fig.3 the following points can be observed:

- After several training cycles for both 50 and 100 training cycles and for all developed BNN models, the hyper parameters values are constant if the number of the hidden units is low (BNN2).
- The values of the hyper parameters obtained with BNN5, BNN7, and BNN11 and trained with only 50 cycles are close to each other, providing no useful information.
- The ARD methods provide best results for the selection of the optimum input variable if the training cycle is increased.
- For all cases, the sunshine duration is the more relevant input and the next one is the air temperature.

TABEL I. STATISTICAL TEST BETWEEN THE DIFFERENT BNN MODELS FOR BOTH TRAINING CYCLES (C=50 AND 100)

Hidden Units		Performance criteria				
		RMSE%	MBE%	MAE%	R	
Training cycle	C=50	2	8.551	3.622	5.974	0.972
		5	13.895	3.942	6.372	0.971
		7	12.872	3.933	6.405	0.971
		11	12.846	3.936	6.420	0.971
	C=100	2	9.650	3.593	6.072	0.972
		5	9.642	3.592	6.071	0.972
		7	11.803	3.876	6.341	0.971
		11	8.024	0.095	4.673	0.965

Statistical tests between measured and predicted DGSR are illustrated in table I for both training cycles (50 and 100 respectively) and for the number of hidden units (2, 5, 7 and 11 respectively), it is clear that the BNN2 for 100 training cycles give the more accurate results with a height coefficient of correlation estimate to 0.972 and a minimum value of RMSE, MBE and MAE estimate to (9.650%, 3.593%, 6.072% respectively).

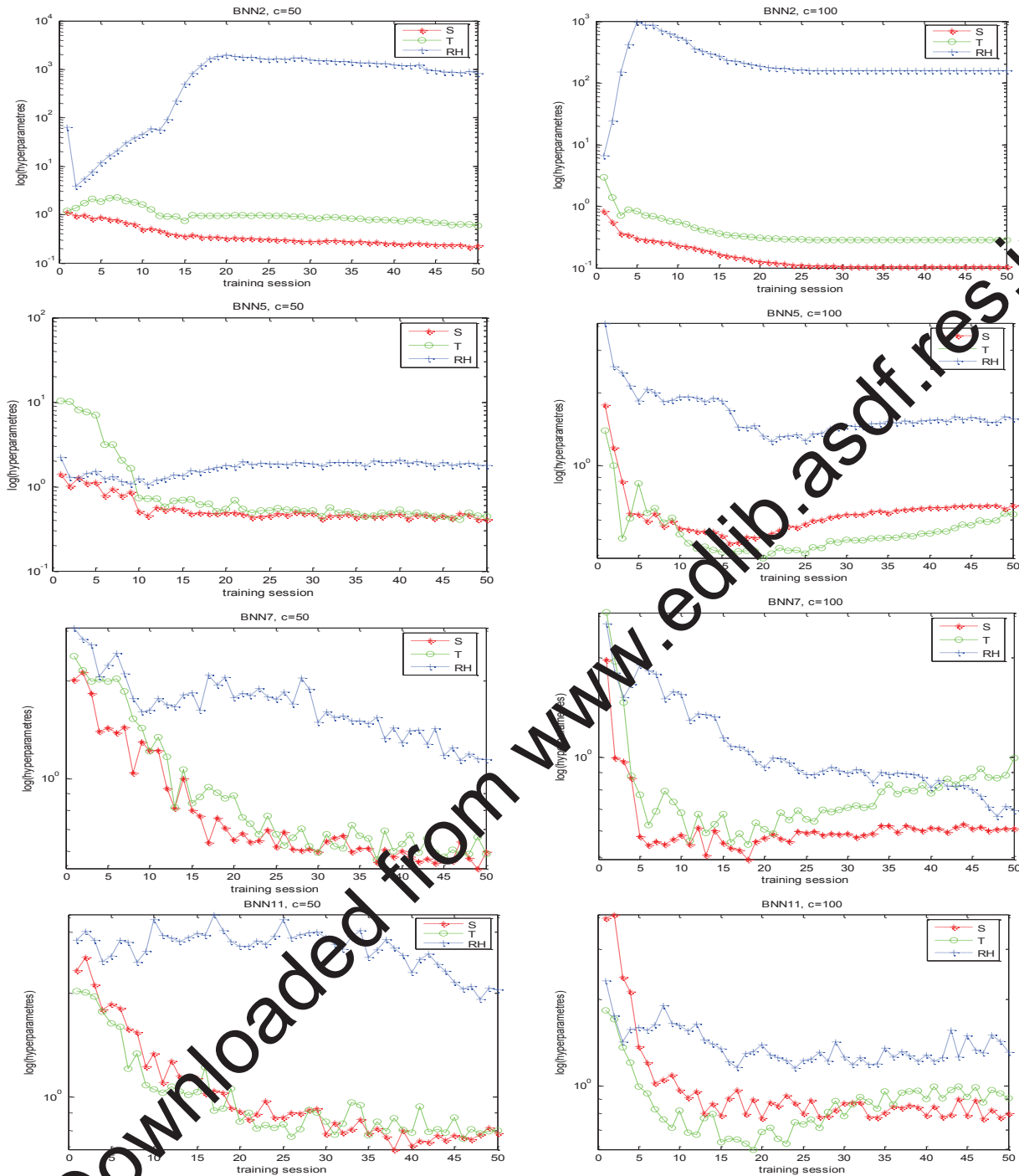


Figure 3. Log hyper-parameters value versus training session .Four BNNs with 2, 5, 7 and 11 hidden units were training using 50 and 100 training cycles

VI. CONCLUSION

The main objective of this study was to evaluate the application of the Bayesian approach for the MLP training set for predicting DGSR. The Bayesian method of the automatic relevance determination to evaluate the more relevant input parameters in modelling DGSR was used. In this sense, the relevant input variables for estimated DGSR are the sunshine duration and the next one was the air temperature. According to the RMSE, MBE and MAE the BNN with 2 hidden unit and 100 training cycles was the best one. In addition the ARD methods can provide better results in the estimation of DGSR with a low number of hidden units on comparison with a classical NN approach which is difficult to determine the number exact of hidden units for a better estimation of the DGSR. BNN could be used effectively for predicting DGSR.

REFERENCES

- [1] H.O. Menges, C. Ertekin and M.H. Sonmete, "Evaluation of global solar radiation models for Konya, Turkey". *Energy Conversion and Management*, vol.47, pp. 3149–3173, 2006.
- [2] M. Benganem, A.A. Joraid, "A multiple correlation between different solar parameters in Medina, Saudi Arabia". *Renewable Energy*, vol. 32, pp.2424–2435, 2007.
- [3] M.G. Abraha and M.J. Savage, "Comparison of estimates of daily solar radiation from air temperature range for application in crop simulations", *Agricultural and forest meteorology*, vol. 148, pp.401-416, 2008.
- [4] A .A. El-Sebaai, A .A. Al-Ghamdi, F.S. Al-Hazmi and A S . Faidah, "Estimation of global Solar radiation on horizontal surfaces in Jeddah, Saudi Arabia". *Energy Policy*, vol. 3, pp. 3645-3649, 2009.
- [5] A. Mellit and S.A. Kalogirou, "Artificial intelligence techniques for photovoltaic applications": A review, *Progress in Energy and Combustion Science*, vol. 34, pp. 547-632, 2008.
- [6] A. Mellit., M . Benganem., A . Hadj Arab. and A. Guesoum, " A simplified model for generating sequences of global radiation data for isolated sites: using artificial neural network and a library of Markov transition Matrices", *Solar Energy*, vol. 79, pp.466-482, 2005.
- [7] S. Rehman and M. Mohandes . "Artificial neural network estimation of global solar radiation using air temperature and relative humidity". *Energy Policy*, vol. 36, pp.571–576, 2009.
- [8] P. Lauret, M. David, E. Fock, A. Baudin and C. Rivier, "Bayesian and sensitivity analysis approaches to modeling the direct solar irradiance", *solar energy engineering*, vol. 128, pp.394-405, 2006.
- [9] C. Marzban. and A . Witt "A Bayesian neural network for severe-hail size prediction". *Weather & forecasting*, vol. 16, pp.600-610, 2001.
- [10] R.M. Neal, *Bayesian learning for neural networks*. Ph.D. Dissertation, University of Toronto, Toronto, Ontario, 1995.
- [11] J. Lampinen and J. Vehtari. "Bayesian approach for neural networks-review and case studies". *Neural Networks*, vol. 14, pp.257–74, 2001.
- [12] R. Yousefi, M. Benganem and A . Mellit , "Prediction of daily global solar irradiation data using Bayesian neural networks: A comparative study", *Renew Energy*, vol. 48, pp.146-154, 2012.
- [13] NABNEY, I. T. *NETLAB: Algorithms for pattern recognition*, London, Springer, 2002. <http://www.ncrg.aston.ac.uk/netlab/>.