Computational Intelligence Techniques for Solar Photovoltaic System Applications

Eleonora D'Andrea and Beatrice Lazzerini
Dipartimento di Ingegneria dell'Informazione: Elettronica, Informatica, Telecomunicazioni
University of Pisa
Largo Lucio Lazzarino 1, 56122 Pisa, Italy
eleonora.dandrea@for.unipi.it, b.lazzerini@iet.unipi.it

Abstract—**In this paper we propose a fuzzy classifier of energy production in solar photovoltaic installations based on the values of some environmental parameters. The classifier is built through a hierarchical process and is obtained by merging *basic* fuzzy models built on input domain regions increasingly smaller, as the result of the construction of appropriate grids on the input domain. The system parameters are optimized by means of a genetic algorithm.**
The interpretability of the fuzzy system helps the electric grid (e.g., smart grid) operator have fast and easy understanding of the energy production, thus allowing easier and faster decision making about electricity production and management.
The achieved results show an average correct classification rate of 97.38% with a maximum of 97.91%.

Keywords-fuzzy rule-based classifiers; genetic algorithms; pattern classification; solar photovoltaic energy

I. INTRODUCTION

Among the renewable energy sources, solar photovoltaic (PV) energy is receiving great attention as it is considered clean, free and inexhaustible. A PV installation consists of a series of solar panels which using sunlight energy generate directly usable electricity due to the PV effect. Usually several panels are connected together to form a system called PV array, to which an inverter is connected that measures the production power of that array and converts the DC power into AC power, as requested by the electrical network.

PV installations are typically used as energy sources for the electric grid. Major issues in electric grid management (in particular, in the case of smart grids) are efficiency and reliability, which require, among other things, fast and easy understanding by the grid operator of both the electricity demand and the electricity supply (energy production). With the aim of helping the grid operator to promptly make his/her decisions, we propose to model the decision process and the elements on which that process operates in linguistic terms.

More precisely, we deal with this issue as a fuzzy classification problem. We divide the values of energy production into three classes (*low*, *medium*, *high*), each modeled by a fuzzy set. We also model the environmental variables (namely, temperature and irradiation) as linguistic variables. Then we build fuzzy rules directly from data so as to associate pairs of values of the two environmental variables with a specific value of produced energy. In this way, the manager of a PV plant can gain enough information from the system so as to perform appropriate functional operations for the installation, even if the exact energy production value is not known [1].

Actually, in the last two decades fuzzy rule-based systems (FRBS) have been extensively applied to pattern classification thanks to their capability to achieve good trade-offs between accuracy and interpretability [2-3]. In particular, interpretability of a FRBS is typically measured in terms of complexity of the rule base, and depends on such factors as comprehensibility of fuzzy partitions of the domains of the involved linguistic variables, number of input variables, number of conditions in the rule antecedents, and number of fuzzy rules. In its simplest form, a fuzzy rule-based classifier is a system consisting of fuzzy if-then rules having a class label as consequent.

When designing a fuzzy classifier two main topics must be considered: fuzzy classifier identification and fuzzy parameter optimization. Major issues in fuzzy classifier identification are: i) how to choose the membership functions of linguistic variables, ii) how to generate the fuzzy rules, and iii) how to determine the output class.

In this paper we propose an easy-to-use approach to extract fuzzy rules from real-world data for PV energy approximation. We employ the fuzzy rule-based classifier `frbc` [4], which is based on the Wang and Mendel algorithm [5], for the generation of the rule base, and we adopt a hierarchical process of data analysis.

Based on this hierarchical process, the fuzzy system is obtained as a combination of *basic* fuzzy models built on input domain regions increasingly smaller, as the result of the construction of appropriate grids on the input domain. The idea is to effectively exploit the input domain space, avoiding the generation of too many, unnecessary rules. Finally, the values of a few model parameters are adjusted by means of a genetic algorithm (GA).

The paper is organized as follows: Sections II, III, IV and V, present, respectively, the experimental data, the proposed system, the experiments, and concluding remarks.

II. EXPERIMENTAL DATASET

The data used in the experiments were collected during five months (from March to July), from a PV installation, located in Italy and consisting of an array of solar panels. The input parameters taken into account are two significant environmental data, that is, *temperature* of the solar panel and *irradiation*. The first one refers to the surface of the

978-3-901882-46-3 © IFIP, 2012

panel exposed to the sunlight, while the second one is the quantity of sun radiation incident on the panel with respect to the whole surface of the panel and to all the electromagnetic spectrum frequencies. The output parameter is the *energy production* related to the PV array and measured by the associated inverter. As said earlier, the energy production numerical values have been mapped to three class labels, by operating a uniform partition on the output domain.

III. Hierarchical Approach to Fuzzy Classifiers Construction

In this Section, we describe the fuzzy system we developed to classify the energy production given the corresponding environmental input variables, that is, panel temperature and irradiation. This fuzzy system results from the union of *basic* fuzzy models built employing the fuzzy rule-based classifier `frbc`, which uses the Wang and Mendel algorithm for the rule-base generation, and adopts, as fuzzy reasoning method, an extension presented in [6] of the fuzzy classifier defined by [7].

In the following, we briefly introduce the methodology adopted to build the fuzzy system. The methodology consists of a first step, a second iterative step and a final third step. In the first step we carry out the following actions:

a. we apply the *k*-means clustering algorithm [8] separately to each input dimension;
b. we use the separation thresholds between the clusters to build a non-uniform grid in the input space, and to construct a non-uniform fuzzy partition on each input dimension consisting of two-sided Gaussian membership functions;
c. we eliminate from further consideration any grid area containing a total number of input samples below a predefined *relevance threshold*; we call each such area an *insignificant* area;
d. we identify, among the remaining areas of the grid, those containing mostly samples belonging to the same output class (majority class); each such area with a number of majority class samples greater than a given *dominance percentage* is named *univocal mapping* area;
e. we build the *first-level basic fuzzy model* by training `frbc` with an appropriate number of samples (belonging to the majority class) extracted from each univocal mapping area;
f. for each non-univocal and non-insignificant grid area, say *to-subgrid* area, we perform the second iterative step.

In the second step we perform the following actions:

g. we build a uniform hard partition on each dimension of the involved *to-subgrid* area, so as to construct a deeper-level grid of the area;
h. we eliminate from further consideration any insignificant area of the new grid, by using the relevance threshold;
i. we find the univocal mapping areas contained in the new grid and we identify the minimum (hyper)rectangle containing all these univocal mapping areas;
j. we construct a uniform fuzzy partition, consisting of Gaussian membership functions, on each dimension of the (hyper)rectangle, thus producing a fuzzy partition of the (hyper)rectangle itself; then we generate a *deeper-level basic fuzzy model* for the (hyper)rectangle; this fuzzy model is built by training `frbc` with an appropriate number of majority class samples extracted from each univocal mapping area inside the hyper(rectangle);
k. for each *to-subgrid* area of the new grid we perform again the second iterative step.

Finally, when no more grid area remains to be processed, a *deeper-level basic fuzzy model* has been built for each *to-subgrid* area at each level of analysis. So the third step is performed by building the *final fuzzy model* as the union of the first-level basic fuzzy model and all the deeper-level basic fuzzy models created during the iterative step.

More in detail, the model parameters are the relevance thresholds RT_1 and RT_2, the dominance percentages DP_1 and DP_2, the upper bound S of samples of the same class that can be extracted from the same area, the *minimum rule weight* w, and the *weight modifiers* Δw_l, where l is the level. RT_1 and DP_1 are pertinent to areas with a number of samples ≥ 20, while RT_2 and DP_2 concern areas with less than 20 samples. For each univocal mapping area A (whatever is the level) a random extraction of $K=min(70\% n_A, S)$ majority class samples is performed, with n_A equal to the number of samples falling in A, and S appropriately chosen.

Finally, the produced rule base is optimized by using w (valid for all levels) so as to reduce the complexity of the whole rule base. Furthermore, Δw_l is also adopted to enhance or inhibit the influence of the rules of a given level l.

The final values of RT_1, RT_2, DP_1, DP_2, S, w and Δw_l are optimized by applying a GA.

IV. Experimental Analysis and Results

We developed the final fuzzy model following the steps of the methodology presented above. Therefore, we performed a hierarchical process to construct the final fuzzy system by merging basic fuzzy models built on increasingly smaller input domain regions. Each *basic fuzzy model* corresponds to an input area associated with one output label.

Table I summarizes the values of the final fuzzy rule base parameters, the fuzzy inference process parameters, and the hierarchical model parameters. In particular, the hierarchical model parameters have been optimized by a real-coded GA (implemented in Matlab®), whose chromosomes represent the following genes: RT_1, RT_2, DP_1, DP_2, S, w, Δw_1, Δw_2, Δw_3, Δw_4. In particular, in order to ensure low complexity and high interpretability, we heuristically decided to allow only four hierarchical levels. This is why we have Δw_1, …, Δw_4. The range of possible values of each gene is chosen in heuristic way mainly based on the specific dataset under consideration. We adopted stochastic uniform selection, scattered crossover with probability 0.8, and uniform mutation with probability 0.01. The population consisted of 30 individuals and the maximum number of generations was 300. The fitness function was the classification error of the developed fuzzy classifier.

Figure 1 shows the scattering of the original dataset (7303 samples), and the grid obtained by applying the *k*-means clustering (*k*=3) in the first step of the methodology: 9 areas are identified on the grid (numbered from 1 to 9).

The process is repeated in a hierarchical way until no area needs to be further divided. Figure 2 depicts the hierarchical analysis up to the 2nd level.

As we can see from Figures 1 and 2, two areas were found to be *insignificant* (areas 3 and 7) and so they were discarded. The *univocal mapping* areas resulting from the application of the methodology are 1, 4, 6 and 8. Each such area is candidate to represent a possible input state for the system. On the remaining *to-subgrid* areas, the second step of the methodology is applied.

TABLE I. PARAMETERS OF THE FINAL FUZZY MODEL

Parameter name	Value
Final fuzzy rule base parameters	
Num. of fuzzy sets per input variable	69
Shape of fuzzy sets	*two-sided Gaussian, Gaussian*
Number of output classes	3
Number of rules	83
Fuzzy inference process parameters	
AND operator	*Minimum*
Implication operator	*Product*
Stress function	*No_stress*
Aggregation function	*Arithmetic mean*
Hierarchical model parameters	
Relevance Thresholds	RT_1=15, RT_2=2
Dominance percentage	DP_1=80%, DP_2=50%
Max. num. of samples extracted	S=145
Minimum rule weight	w=0.94
Weight modifiers	Δw_1=0.45, Δw_2=–0.55, Δw_3=0.15, Δw_4=0

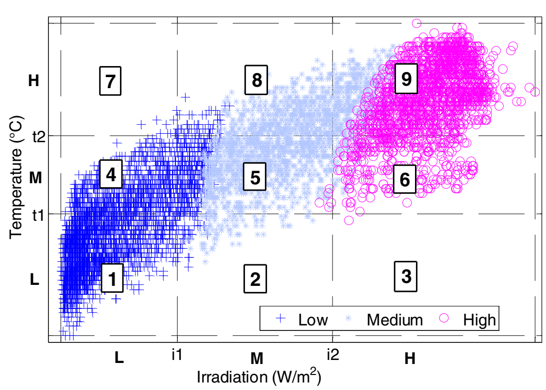

Figure 1. Partition of the input domain by means of *3*-means clustering.

In order to validate the final fuzzy classifier, we tested it on 30 different training and test sets generated from the available data by applying the proposed methodology. We achieved an average classification performance of 97.38% on the test set with a maximum correct classification performance of 97.91%.

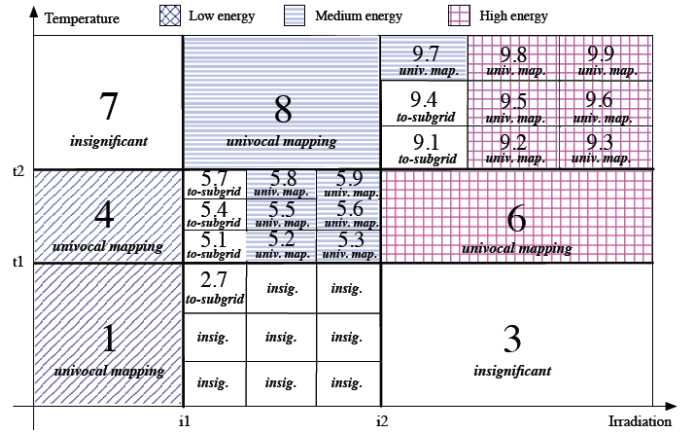

Figure 2. The hierarchical analysis partition obtained up to the 2nd level.

V. CONCLUSIONS

In this paper we have proposed an easy-to-use approach to extract fuzzy rules from real-world data for PV energy classification. We performed a hierarchical process to construct a fuzzy model by merging *basic* fuzzy models built on input domain regions increasingly smaller, as the result of the construction of appropriate grids on the input domain. The aim was to exploit the easiness of use and the interpretability of the fuzzy approach along with a methodology of input domain space analysis which builds an optimal fuzzy rule base avoiding the generation of too many rules. Finally, a GA optimized the model parameters values.

REFERENCES

[1] P.A. Gutierrez, S. Salcedo-Sanz, C. Hervas-Martinez, L. Carro-Calvo, J. Sanchez-Monedero and L. Prieto, "Evaluating nominal and ordinal classifiers for wind speed prediction from synoptic pressure patterns," International Conference on Intelligent Systems Design and Applications (ISDA), 2011, pp. 1265-1270.

[2] S.M. Fakhrahmad, A. Zare, and M. Zolghadri Jahromi, "Constructing Accurate Fuzzy Rule-Based Classification Systems Using Apriori Principles and Rule-Weighting", Intelligent Data Engineering and Automated Learning, Springer Berlin/Heidelberg, vol. 4881, 2007, pp. 547–556.

[3] H. Roubos and M. Setnes, "Compact and transparent fuzzy models and classifiers through iterative complexity reduction", IEEE Transactions on Fuzzy Systems, vol. 9, n. 4, 2001, pp. 516–524.

[4] M. Cococcioni, E. D'Andrea and B. Lazzerini, "Providing PRTools with fuzzy rule-based classifiers", IEEE World Congress on Computational Intelligence (WCCI), Barcelona, 2010, pp. 1-8.

[5] L.-X. Wang and J. M. Mendel, "Generating fuzzy rules by learning from examples", IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, n. 6, 1992, pp. 1414-1427.

[6] O. Cordón, M.J. del Jesus and F. Herrera, "A proposal on reasoning methods in fuzzy rule-based classification systems", International Journal of Approximate Reasoning, vol. 20, n. 1, 1999, pp. 21-45.

[7] L.I. Kuncheva, "On the equivalence between fuzzy and statistical classifiers", International Journal of Uncertainty Fuzziness and Knowledge-Based Systems, vol. 4, n. 3, 1996, pp. 245−253.

[8] J.B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations, Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability", Berkeley, University of California Press, vol. 1, 1967, pp. 281-297.