Steady-State Inference in Performance Tests of Refrigeration Compressors Using ANN Ensemble

Cesar Alberto Penz, Carlos Alberto Flesch, João Paulo Rossetto

Steady-State Inference in Performance Tests of Refrigeration Compressors Using ANN Ensemble

Cesar Alberto Penz[1,*], Carlos Alberto Flesch[1], João Paulo Rossetto[2]

[1]Federal University of Santa Catarina, Mechanical Engineering Dept., Florianópolis, Brazil
`cep@labmetro.ufsc.br, carlos.flesch@ufsc.br`
[2]Whirlpool S.A., Embraco Cooling Solutions, Joinville, Brazil
`joao.p.rossetto@embraco.com`

Abstract. The use of ensembles is an important way to ensure reliability of ANN-based tools. This article presents the development of an ANNs ensemble to infer steady-state in performance tests related to refrigeration compressors. Practical results are presented about the performance of individual ANNs and the ensemble. ANNs accuracy and diversity are discussed and evaluated.

Keywords: neural networks ensembles; reliability improvement; steady-state inference; refrigeration compressor.

1 Introduction

The use of neural networks ensembles is typical in many industrial and scientific applications. The combination of a set of networks regards as a way to manage the limitations of individual networks. It is known that neural networks are susceptible to errors, but the neural networks can be combined in a way to reduce the effects of these individual errors [1,2,3,4,5,6,7].

Reference [8], entitled "Neural networks and the bias/variance dilemma", deals with the performance of nonparametric inference tools. An example of this type of inference tool is the multilayer perceptron neural network. The reference [8] presents these errors as being comprised of two parts – bias and variance – and compares the features of responses obtained with nonparametric tools (model-free) to the responses obtained with model-based approaches. According to [8] neural networks may present significant portion of variance since they may have great amount of parameters to be obtained in the training phase. In contrast, model-based solutions are difficult to implement for complex problems and can cause high bias. Additionally, it is highlighted that a high bias level is achieved when ANNs are trained with low amount of samples or the training process is driven to overfitting (poor generalization). The concepts and developments presented in the reference [8] are supported by a series of studies using neural networks ensembles for inference in complex scientific and industrial applications [5,6], [10,11,12,13,14,15,16]. These studies emphasize that the use of neural networks ensembles result on better balance between bias and variance.

An efficient ANN ensemble is composed by networks that are as accurate and diversified as possible [12][14,15][17,18]. The accuracy of the neural networks can be

achieved by using appropriate training sets and administrating the training process to avoid, as far as possible, the overfitting [12][14,15].

The warranty of diversity of a set of neural networks is directly related to how the ensemble will manage the variance issue [6][14,15][17]. An important feature for ensuring a diverse set of ANNs is the concurrency of errors. A set of networks which share low amount of concurrent errors tends to meet the diversity criterion [19,20,21].

Some ways to obtain potentially different networks are: different training algorithms; different architectures (number of layers and neurons, connectivity between layers and neurons, different activation functions); different training sets; different sets of initial weights for training. Among these ways, the simultaneous use of different training sets and different sets of initial weights has shown great potential for generating diversity [17][20,21].

This paper presents the development of a neural network ensemble for the inference of steady-state condition in performance tests of refrigeration compressors. It is shown the evolution from complete ignorance when considering neural networks individually to a high reliable state when neural networks are considered to work together. Analysis of correlation, true positive and false positive observations are made to ensure accuracy and diversity for the ANNs ensemble.

2   Performance test of refrigeration compressors

The performance tests are run on special rigs, which operate as a refrigeration circuit with many controlled variables. In addition, it is possible to measure quantities that are not generally monitored in refrigeration systems.

Refrigerating capacity is measured in watts and is calculated as the product of the mass flow rate of refrigerant fluid through the compressor and the difference between the specific enthalpy of the refrigerant fluid at two points in the refrigeration circuit [22]. According to the enthalpy determination points that are chosen, it is possible to obtain different interpretations for the refrigerating capacity [22]. The most common is the heat removal from the cooled environment. The simplest method for determining the mass flow rate in a refrigeration circuit is to measure it directly using mass flow meters that are available on the market. Equipments that satisfy the measurement uncertainties required by international standards can be easily found. However, the measurement of refrigerating capacity needs to be carried out using two different and independent methods [22]. Typically, in order to ensure the methods independence, one is based on the direct measurement of the mass flow rate and the other is based on the heat balance inside a calorimeter.

Fig. 1 shows a simplified schematic of the refrigeration circuit and instrumentation. In this case the calorimeter is comprised in the area defined by the dashed line, where there is an evaporator, a heating resistor, and four temperature measurements transducers.

A typical characteristic of these tests is the long time required to achieve stable conditions, under which the measurements – used for the calculation of the final results – can be performed. In general, performance tests last 2½ hours. The criteria for

stability are defined for the most important quantities that are measured during the test. For each quantity there is a percentage or an absolute variation limit which must be calculated based at the average of the measured values of the last period of 1 hour. So, if a specific quantity stays within the respective variation limit during 1 hour it is said to be stable. For example, the compressor shell temperature has an absolute variation limit of ±3°C then if the final shell temperature is to be +55°C this quantity is said to be stable when it overtakes +52°C. The issue is that the final values for many quantities are just known at the end of the test, indeed it is not possible to know when the quantity enters the stability region.
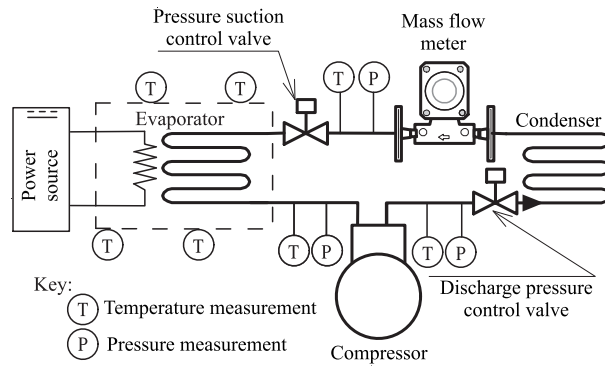

Fig. 1. Simplified schematic of the test rig

The refrigerating capacity obtained using the mass flow meter method is analyzed in this article. Fig. 2 presents a set of four typical examples of refrigerating capacity data obtained during performance tests. These tests were run on compressors of different models produced by the same manufacturer. Since the refrigeration capacities have different values – from 50 W to 200 W – the normalized data is presented for better visualization of the dynamic behavior.

The steady-state region for the refrigerating capacity is characterized when it remains within ±2% of the final value. This final value is obtained through the mean of the measurements taken during the last hour of the test.

Figs. 3 and 4 present other two important quantities that are measured during a performance test: compressor suction pressure and the compressor shell temperature. These quantities will be used as inputs of the neural networks that compose the ensemble to infer the steady-state.

An example of refrigerating capacity data and the respective desired ANN output is presented in Fig. 5. As the refrigerating capacity enters the ±2% range the ANN output should rise from 0 to 1. It is important to highlight that, even being interested in the refrigerating capacity, its final value can just be known at the end of the test. As a consequence, the ±2% stability range can only be known at the end of the test. Accordingly, other quantities must be used to support de steady-state inference. Supposing a true positive inference one could start the computation of the test result and early stop the test.
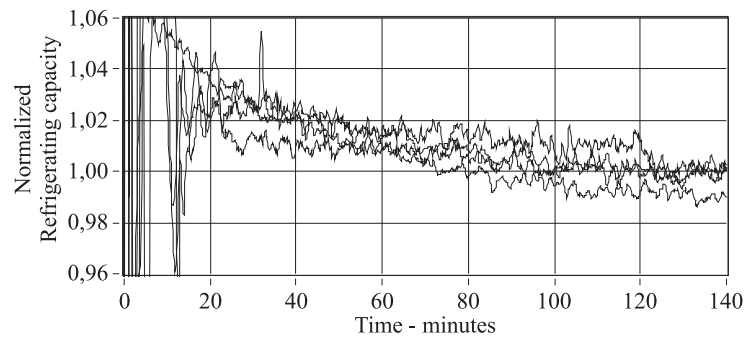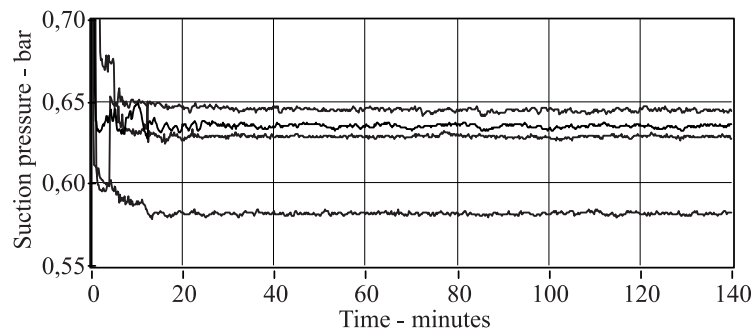
Fig. 2. Normalized refrigerating capacity data
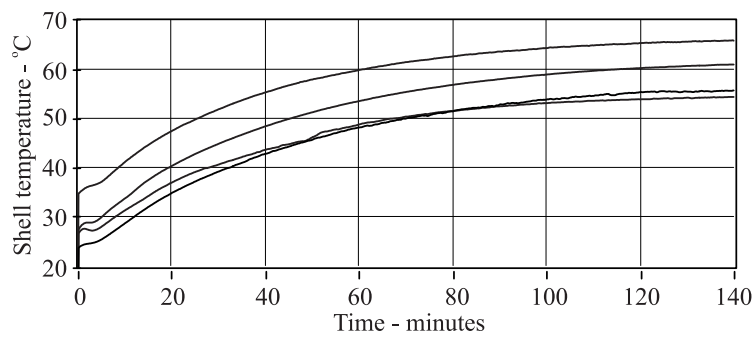

Fig. 3. Compressor suction pressure

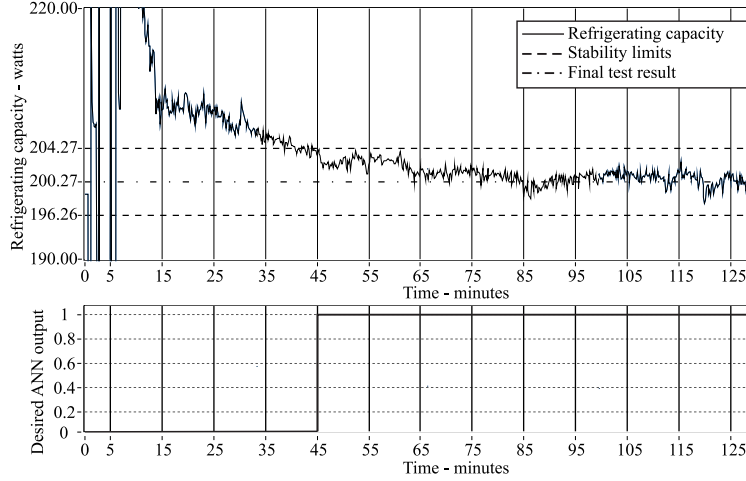**Fig. 4.** Compressor shell temperature

Fig. 5. Refrigeration capacity and desired ANN output

These types of data are used in Section 3 for describing the neural network used to steady-state inference of the refrigeration capacity. The next section briefly presents the concepts used to develop the neural network.

3 Neural network ensemble

The multilayer perceptron neural network trained with backpropagation algorithm is used for the steady-state inference. Despite this type of neural network is not an intrinsically dynamic network, it is used for this application. The dynamical modeling is done by the inputs that were chosen to compose the network.

The ANN inputs are average values of refrigerating capacity, suction pressure and shell temperature (three inputs); lagged average values of these quantities (four inputs – two for the refrigerating capacity and one for each one of the other quantities); difference between actual and lagged average values (eight inputs); and standard deviation of refrigerating capacity (one input). The sample size for averages and standard deviation is 15 with sampling frequency of 0.1 Hz. The number of inputs is 16 for a 3-hidden layer network (15-10-10 neurons) with hyperbolic tangent activation functions. The output layer is comprised of one neuron with logistic activation function. This architecture resulted from an extensive evaluation heuristic and literature insights. The performances of individual neural networks are presented in this section.

The training set was composed by 50 examples of performance tests obtained from different compressors. The bootstrap method was used to compose the individual ANNs training sets (40 examples randomly selected) and the remaining examples were used for on-line validation (early stop to avoid overfitting). As the bootstrap was run with replacement, the on-line validation set could have more than 10 examples. The neural network toolbox of Matlab® was used to train a set of 30 ANNs. After

that, the neural networks were tested with a set of 210 examples which does not include the 50 examples used for training. The number of ANNs that was trained is based on the achieved performance. It is shown ahead that the ensemble comprised of 30 ANNs is suitable for this application when used with a confirmation time. The literature presents several applications that indicate a number between 20 and 200 for the ensemble members to provide proper results [2][6][12][14,15].

Fig. 6 presents the probability of not indication the steady-state for each one of the 30 ANNs when the test data set is used. It can be seen that the probability varies from almost zero to 0.40, which means that some ANNs indicate steady-state condition to almost all the 210 examples and some that do not indicate on almost half of the examples.

Fig. 7 presents the probabilities of true positive and false positive indications for the test data set. One can note that most of the probabilities are around 0.5, which indicates a high degree of ignorance when the neural networks are observed individually. This fact could make the specialist regret on using this neural networks for the steady-state identification. Another point that should be noticed is that ANNs with high probability of not indication have true positive probability higher than false positive probability. This means that if these neural networks infer the steady-state they act mostly on a true positive way.
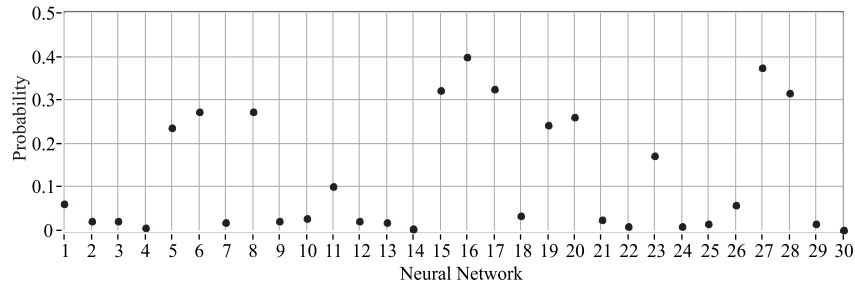

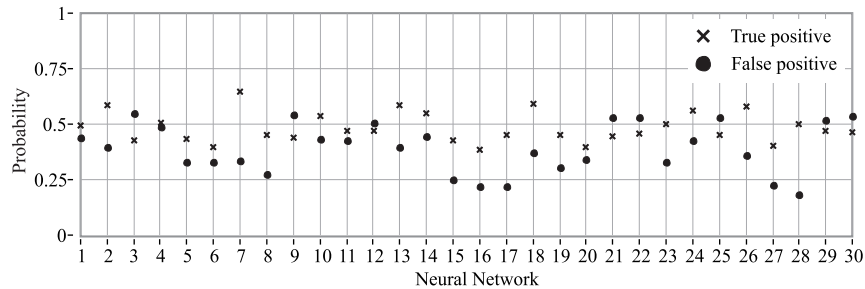**Fig. 6.** Probability of no indication of steady-state

Fig. 7. Probability of indication of steady-state

To evaluate the adequacy and the diversity of the 30-ANNs ensemble, Figs. 8 and 9 present the correlations of true positive and false positive indications. These correlations values were obtained for the test data set and for all the 30 ANNs.

The histograms show that there are no equal networks or that behave like so in the ensemble. Most of the ANNs presented correlation from weak to moderate, that is, they share equal levels of accurate and inaccurate indications. This information indicates that the ensemble has an acceptable level of diversity.
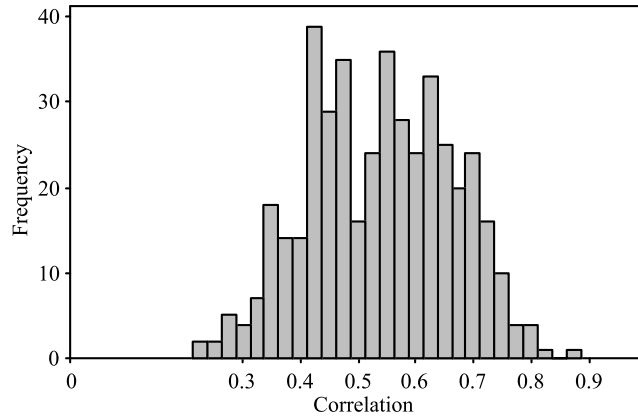

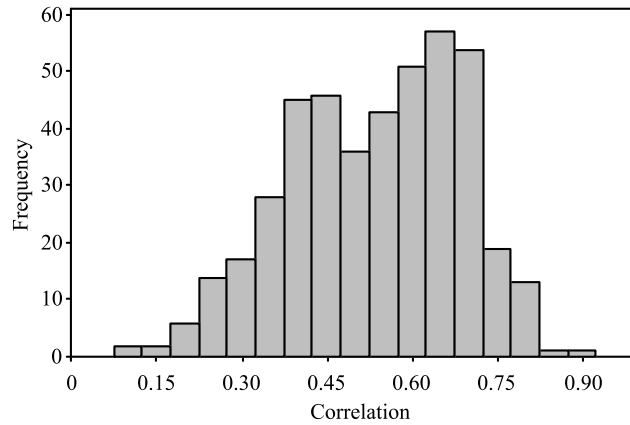**Fig. 8.** Correlation of true positive indications

Fig. 9. Correlation of false positive indications

Fig. 10 presents the relationship between the number of ANNs that inferred the steady-state and the difference between the refrigerating capacity at such instant and the final value for the performance test. The solid line is the average absolute for all

the performance tests in the training data set which had inferences; dotted lines are two standard deviations limits.

The ANNs-outputs integration was based on the analysis of the relationship between the number of ANNs that should indicate the steady-state condition and the time that should be waited for the confirmation.

As can be seen in Fig. 10, a minimum of 27 ANNs positive answers should be waited to achieve inferences into the ±2% steady-state region. Nonetheless, intermediate rules can be created to achieve a more effective reduction on test times. The rules combine number of ANN positive answers with confirmation waiting times. So, even small numbers of ANN positive answers may produce proper steady-state inferences.
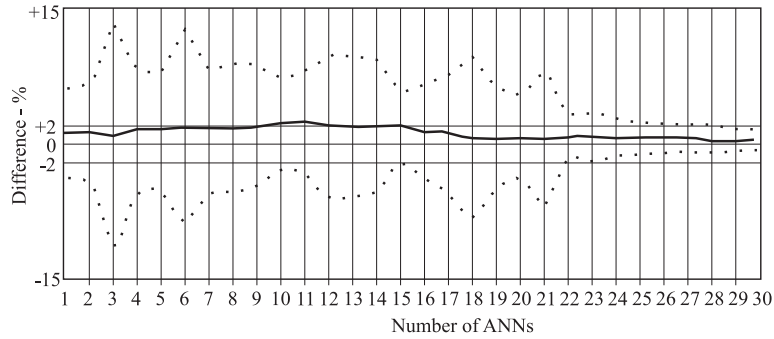

Fig. 10. Difference of refrigerating capacity versus number of ANNs

Again, the 50-example training data set was evaluated using all the 30 neural networks to create graphs as presented in Figs. 11 and 12. Each one represents the time that should be waited after a specified number of ANNs indicates the steady-state and the related average absolute relative difference for all the training data set (solid line). This difference is between the refrigerating capacity at the instant which a certain number of ANNs indicated steady-state and final value of the performance test. The dotted lines represent two standard deviations added and subtracted from the average. This analysis was run considering numbers of ANNs varying from 1 to 30, but for the sake of space, only the results for 10 and 25 networks are presented in this article.

The whole set of graphs was used to define the rules to integrate the ANNs outputs, as presented in Table 1, where numbers of ANNs and confirmation times were related by the analysis of Fig. 10 and all the 30 graphs (as shown in Figs. 11 and 12). Then, all the 210 test data set examples were evaluated using the 30-ANN ensemble and the four rules.
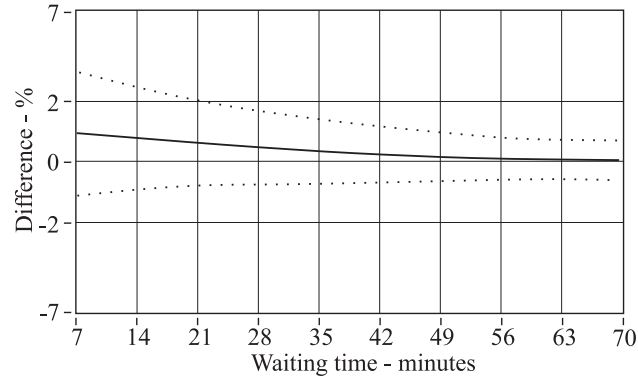
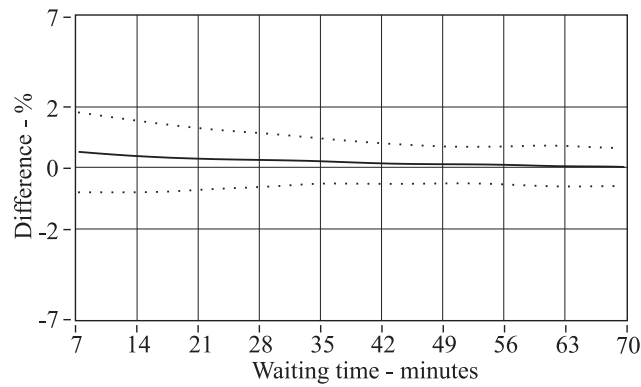**Fig. 11.** Confirmation waiting time graph for 10 ANNs

Fig. 12. Confirmation waiting time graph for 25 ANNs

Table 1. ANN Integration Rules

Rule	Number of ANNs	Confirmation waiting time
I	< 15	40 minutes
II	≥ 15 and < 20	30 minutes
III	≥ 20 and ≤ 27	10 minutes
IV	> 27	0

Fig. 13 presents a histogram of the relative differences between the refrigerating capacities at instant the steady-state was inferred by ensemble and the expected values for all the 210 performance tests. It can be noted that most of the tests had the steady-state inference within the range of ±2% of the final value, actually, more than 91%, what is a great improvement if compared to the true positive probabilities presented by the neural networks individually.
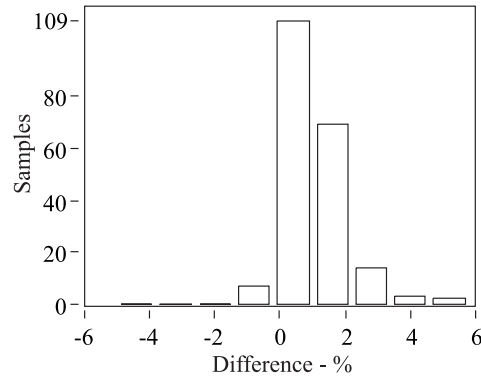
Fig. 13. ANN ensemble results

4   Conclusions

This paper presented the application of an ANN ensemble for steady-state inference that is part of a more complete tool to predict the refrigerating capacity in performance tests of refrigeration compressors. It was shown that ANN individual errors can be minimized when the networks are combined. Correlation analysis can be a good indication of diversity for an ANN ensemble. The inference allows a significant improvement in quality control and R&D activities. The average test time for the 210 performance tests was 2½ hours; when these tests are evaluated using the ensemble it is possible to reduce to half of that time. The application of steady-state inference could be extended to other performance variables like the power consumption, isentropic efficiency and coefficient of performance (COP).

References

1. Krogh, A., Vedelsby, J.: Neural networks ensembles, cross validation and active learning. Adv. in Neural Inf. Proc. Sys. 7, (1995)
2. Efron, B., Tibshirani, R.: An Introduction to the Bootstrap. Chapman & Hall (1993)
3. Sharkey, A.J.C.: Combining Artificial Neural Networks: Ensemble and Modular Multi-Net Systems. Springer-Verlag (1999)
4. Winkler, R. L., Clemen, R.T.: Multiple experts vs. multiple methods: combining correlation assessments. Decis. Anal., v.1, n.3 (2004)
5. Ahmad, Z., Zhang, J.: Bayesian selective combination of multiple neural networks for improving long-range predictions in nonlinear process modeling. Neural Comp. & Applicat., v.14 (2005)

6. Zio, E.: A study of the bootstrap method for estimating the accuracy of ANN in predicting nuclear transient processes. IEEE Trans. on Nucl. Sci., v.53, n.3 (2006)
7. Peretti, C. de, Siani, C., Cerrato, M.: A bootstrap neural network based heterogeneous panel unit root test: application to exchange rates. Working Papers in Economics, Univ. of Glasgow (2010)
8. Geman, S., Bienenstock, E., Doursat, R.: Neural networks and the bias/variance dilemma. Neural Computation, v.4 (1992)
9. Ahmad, Z., Zhang, J.: A comparison of different methods for combining multiple neural networks models. Proc. of the Int. Joint Conf. on Neural Netw., v.1 (2002)
10. Edwards, P. J., Peacock, A. M., Renshaw, D., Hannah, J. M., Murray, A. F.: Minimizing risk using prediction uncertainty in neural network estimation fusion and its application to papermaking. IEEE Trans. on Neural Netw., v.13, i.3 (2002)
11. Hu, Y. H., Hwang, J.: Handbook of Neural Network Signal Processing. CRC Press (2002)
12. Granitto, P. M., Verdes, P. F., Ceccatto, H. A.: Neural net-works ensembles: evaluation of aggregation algorithms. Artificial Intell., v.163, i.2 (2005)
13. Fortuna, L., Giannone, P., Graziani, S., Xibilia, M. G.: Virtual instruments based on stacked neural networks to improve product quality monitoring in a refinery. IEEE Trans. on Instr. and Meas., v.56, i.1 (2007)
14. Yu, J. B., Xi, L. F.: A neural network ensemble-based model for on-line monitoring and diagnosis of out-of-control signals in multivariate manufacturing processes. Exp. Sys. with Applicat., v.36, i.1 (2009)
15. Wu, B., Yu, J.: A neural network ensemble model for on-line monitoring of process mean and variance shifts in correlated processes. Exp. Sys. with Applicat., v.37, i.6 (2010)
16. Trichakis, I., Nikolos, I., Karatzas, G. P.: Comparison of bootstrap confidence intervals for an ANN model of a karstic aquifer response. Hydr. Proc. (2011)
17. Sharkey, A. J. C., Sharkey, N. E.: How to improve the reliability of artificial neural networks. Dept. of Computer Sci., Univ. of Sheffield, Report CS-95-11 (1995)
18. Zhang, J.: Developing robust non-linear models through bootstrap aggregated neural networks. Neurocomputing, v. 25 (1999)
19. Sharkey, A. J. C., Sharkey, N. E., Chandroth, G. O.: Neural nets and diversity. Neural Comp. and Applicat., v.4 (1996)
20. Brown, G., Wyatt, J., Harris, R., Yao, X.: Diversity creation methods: a survey and categorization. J. of Inf. Fusion, v.6 (2005)
21. Rokach, L.: Taxonomy for characterizing ensemble methods in classification tasks: a review and annotated bibliography. Computational Stat. & Data Anal. (2009)
22. ISO Testing of refrigerant compressors, ISO Standard 917 (1989)