



HAL
open science

A new method for dynamic modelling of bread dough kneading based on artificial neural network

Bouchra B. Lamrini, Guy G. Della Valle, Cristian C. Trelea, Nathalie N. Perrot, Gilles Trystram

► **To cite this version:**

Bouchra B. Lamrini, Guy G. Della Valle, Cristian C. Trelea, Nathalie N. Perrot, Gilles Trystram. A new method for dynamic modelling of bread dough kneading based on artificial neural network. *Food Control*, 2012, 26, pp.512-524. 10.1016/j.foodcont.2012.01.011 . hal-01000846

HAL Id: hal-01000846

<https://hal.science/hal-01000846>

Submitted on 6 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **A New Method for Dynamic Modelling of Bread Dough Kneading Based on**
2 **Artificial Neural Network**

3

4 Bouchra LAMRINI ^{a*}, Guy DELLA VALLE ^b, Ioan Cristian TRELEA ^c, Nathalie PERROT ^c, Gilles
5 TRYSTRAM ^a

6

7 ^a UMR GénIAL, AgroParisTech, INRA, F 91300 MASSY, France

8 ^b UR BIA, INRA, BP 71627, 44316 NANTES Cedex 3, France

9 ^c UMR GMPA, AgroParisTech, INRA, BP 1, F 78850 THIVERVAL-GRIGNON, France

10

11 All correspondence should be addressed to:

12 (blamrini@yahoo.fr) and (gilles.trystram@agroparistech.fr) / Tél: +33-1-6993-5065 / Fax: +33-1-6993-5185)

13

14

15 **Abstract.** This paper presents a dynamic model of the kneading process based on artificial neural networks.
16 This dynamic neuronal model allows predicting the bread dough temperature and the delivered power necessary
17 to carry out mechanical work. This neuronal technique offers the advantage of very short computational times
18 and the ability to describe nonlinear relationships, sometimes causal, explicit or implicit, between the input and
19 output of a system. We used the recurrent neural networks to capture the dynamic of the process. The type and
20 the number of inputs to the neural networks, as well as the nature of the learning set, the architecture and the
21 parameter learning technique have been studied. The comparison of the results with experimental data shows the
22 possibility to predict the temperature and the power delivered to the dough for various operating conditions.

23

24 **Keywords:** Bread-making; Kneading; Recurrent Neural Networks; Artificial learning; Performance modelling.

25

26

* Present address: IRCAM-Centre Georges Pompidou, 1, Place Igor stravinsky, 75004 Paris, France.
Tél: +33-1-4478-4861 / Fax: +33-1-4478-1540

27 1 Introduction

28

29 The increasing complexity of systems used in food industry, in terms of functions and methods implemented,
30 promotes the use of innovative tools for modelling processes (Norton & Sun, 2006; Bimbenet et al., 2007;
31 Broyart & Trystram, 2002; Perrot et al., 1996; Ndiaye et al., 2005). This increasing complexity of the operation
32 requires an increased accuracy of the description of the technology involved, in particular for the complex
33 systems where instrumentation or analytical approach are difficult. This difficulty, which arises from the
34 processes nonlinearities and of strong interactions between operating variables, is therefore added to the
35 complexity of the analysis of physical phenomena.

36

37 These difficulties are well illustrated in wheat flour dough kneading, the fundamental process in bread making,
38 responsible in large part of the quality of the finished product. From two main components, water and flour, this
39 process allows the preparation of homogeneous and viscoelastic dough (Bloksma, 1990; Roussel & Chiron,
40 2003). The dough is subject to intense forces of extension and shear, depending on the geometry of the mixer
41 bowl, stirrer and on the rotor speed as well as its rheological properties. After the phase of mixing for
42 moisturizing the components, the protein network is obtained; the dough undergoes an important evolution
43 which causes a modification of its elasticity and its viscosity. These changes reflect the modifications in the
44 nature of hydrogen and disulfides bonds, and of the hydrophobic interactions which associate the various
45 components. The dough properties can be profoundly changing because of the flours quality, the nature and the
46 quantity of the technological auxiliaries added, and the water content in the recipe. The process causes an
47 intensive contribution of mechanical work, essential for the dough quality (Tanaka & Bushuk, 1973). The
48 mechanical energy provided by the stirrer, is partially converted by viscous dissipation into a thermal energy
49 contribution to the dough. This energy, absorbed by the dough, is practically measured by temperature raise
50 during kneading process (Contamine et al., 1995; Charun et al., 2000). Another part of the energy transmitted by
51 the moving arm is dissipated towards the outside. The energy balance is difficult to set, mainly because of the
52 discontinuity of the limit surface of the dough with environment. Besides the contribution of mechanical energy,
53 the value of instantaneous specific power will also affect the dough characteristics. According to Hlynka (1962),
54 consistency increases with the dough temperature and for equal consistency, the speed of dough development
55 increases with the dough temperature. Bloksma and Nieman (1975) indicate that the dough stickiness decreases
56 according to the temperature until 40 and 45 °C. These results illustrate the complex dependency of the

57 rheological behaviour of dough during kneading, between the process variables and, finally, the apparent
58 contradiction the operator has to face, either engineer or baker.

59

60 At present, the behaviour of the kneading operation depends on operators “know-how”, developed empirically,
61 and for whom there is no pertinent sensor, but whose knowledge is necessary to control the kneading behaviour
62 and the bread final quality. This lack of knowledge is basically due to ignorance of the underlying physical and
63 chemical mechanisms. As an example, to make fermentation progress satisfactory, it is recognised necessary to
64 have a final temperature of dough between 20 and 26°C. This basic temperature allows to determine the
65 temperature of the water added to the flour, and so to acquire at the end of kneading dough with the expected
66 temperature. It is an empirical method that allows obtaining constant results by means of a simple calculation
67 (Roussel & Chiron, 2003). To fill up these gaps, different studies were led on the process kneading (Chin &
68 Campbell, 2005; Levavasseur et al., 2006; Mann et al., 2008) and its modelling (Binding et al., 2003) to describe
69 the evolution of dough characteristics according to operating conditions. They illustrate the specificity of the
70 technology of the kneading equipment and the used recipe. The extension of such works to all operations of
71 bread-making process would allow defining tools for decision support and innovation, and to predict their impact
72 on the properties sensory of the product (Rousu et al., 2003). Previous approach based on the incorporation of
73 “know-how” (Ndiaye et al., 2009) permits a qualitative model which considers bread-making process as a chain
74 of various discrete operations without taking into account the dynamics of each operation. Currently no
75 mathematical model representing the power and temperature dynamics of the dough during kneading is
76 available. Therefore, the development of a method for prediction and control of these two variables for a given
77 formulation present a real interest for the operators. This work describes the methodology, based on automatic
78 learning, to model one of the most complex operations in bread-making: the kneading process.

79

80 In this context, we suggested a dynamic model based on artificial neural networks called "Model-Operation".
81 Developments in recent years with nonlinear modelling approaches based on automatic learning, in particular
82 artificial neural networks, justify our interest carried in these approaches (Dreyfus et al., 2002; Acunà, 1995;
83 Fravolini, 2003; Hernandez-Perez, 2002; Olmos-Perez, 2003). Artificial Neural Networks (ANN) are a
84 promising technique because of their parsimony and their flexibility towards the environment variations to be
85 analysed, in other words their ability to adapt to novelty. Neural approach is well adapted for a complex and
86 nonlinear process such as dough kneading in order to improve knowledge of the effect of kneading conditions on

87 temperature and instantaneous power evolutions, by assuming that these variables allow to describe the rheology
88 and the dough quality at the end of the kneading process (Lamrini et al., 2009). The neural model introduces
89 same advantages in term of cost, of time and of optimization, for the development of prediction and decision
90 support tool computing the temperature by using the power dissipated during the kneading process. Considering
91 control purposes, the ability to describe dough temperature evolution and power only from initial operating
92 conditions is an important objective. In this study, based on the general architecture of system proposed (Lamrini
93 et al., 2009), is the first one to apply the connexionist techniques for dynamic process modelling such as the
94 kneading. In this purpose, a set of experimental data involving formulations variables and operating conditions
95 was exploited to define the model.

96
97 In a general way, the modelling techniques based on automatic learning allow to approach the modelling of
98 physical phenomena which description is arduous. So, this widens the field of possible as far as modelling, but
99 also to separate from a detailed physical description for known processes and to reduce the development time of
100 a particular model. The scope of the paper covers an introduction; a description of database is given before
101 introducing the methodology adopted to model the power dynamics and that of temperature of dough in the
102 course of kneading. Finally, experimental results and model performance are presented and discussed.

103

104

105 **2 Materials and methods**

106

107 **2.1 Data base composition**

108

109 We consider the kneading process as a system evolving in its environment and that we can interact with it. It is
110 therefore important to distinguish the input variables allowing the system to evolve from the ones affecting the
111 state of its output variables we are interested to predict. Then, it is important to determine the relationships
112 between inputs and outputs and to know the nature and the modes of interaction with the environment. The
113 outputs can be modified by the inputs action or due to the effect of disturbances from the environment.
114 According to a systemic study, carried out on different stages of bread-making, a first presentation of the
115 kneading process and its environment can be illustrated by Fig. 1. Through this structure and with various

116 operating conditions (amounts of ingredients, their temperature, air temperature, rotor speed, kneading duration,
117 etc.), we proposed three models as follows:

- 118 1. Operation-Model: used to predict the temperature dynamics 'T (t)' and that of power 'P (t)' delivered
119 during the kneading process.
- 120 2. Rheology-Model: used to predict the rheological and viscoelastic behaviour of the dough during the
121 leavening process. The considered variables in this model are: dough porosity /or volume fraction 'Fv
122 (t)', dough stability 'H / L (t)' defined by [Shehzad et al. \(2010\)](#) and dough elasticity 'E '(T_{lev})' (T_{lev} is
123 the dough temperature during leavening process).
- 124 3. Quality-Model: used to predict automatically, at the end of the process, the sensory variables of dough
125 such as, smoothness, stickiness, elasticity, extensibility, stability, and identify the three dough states
126 (normal, excessive and insufficient) by means of fuzzy classification. The validation results of the two
127 models "Rheology-Model" and "Quality-Model", introduced by ([Lamrini et al., 2009](#)), will not be
128 discussed in this paper.

129

130 The work in this paper is only dedicated to the operation model. The structure of this "Operation-Model" has
131 been proposed starting from the collected data on the bread dough kneading through the literature,
132 experimentations and technological expertise ([Fig. 2](#)). The Operation model is designed to predict the two
133 'Variables-Operation': the power P(t) and the dough temperature T(t) at time 't +1'. The following operating
134 variables are used as inputs of model: the ingredients mass 'M', the rotation speed 'Vr', the water temperature 'Te'
135 and the flour temperature 'Tf'. According to the expert knowledge, the environment temperature has little
136 influence on the temperature evolution of the dough, especially in a classical range of operation (15-25 °C),
137 which justifies that it is not present as an input of neural Operation-Model. To take into account the dynamics of
138 the two outputs, we used the recurrent neural networks.

139

140

141 [Fig. 1](#). Representative Diagram of the kneading process and its environment

142

143

144 [Fig. 2](#). Operation model structure adopted to predict the temperature dynamics of the bread dough and those of
145 power delivered during kneading process.

146 For the modelling stage, we established, during the dough kneading, a database of 36 power and temperature
147 kinetics (Fig. 3, Fig. 4 and Fig. 5). A spiral mixer (DIOSNA GmbH, Osnabrück, Germany) was used according
148 to a procedure detailed by Chiron et al. (2008), the data are collected for different conditions of rotation speed
149 [81-329 rpm], ingredients temperature [5-25 °C] and mixture weight [3-10.5 kg]. The flour used is a type 55,
150 containing neither enzyme nor additive. The flour mass used varied between 1.8 and 6.3 kg, and moisture was
151 kept constant (water added = 62% of flour mass). Kneading is carried in two stages: a mixing stage, for 4 min
152 and with a speed of 100 rpm, to moisturize the components and obtain a smooth dough, followed by a kneading
153 stage which varies according to an experimental design fixed by the expert in order to explore the operating
154 possibilities of these mixers and, possibly, determine the combination of these variables allowing to separate the
155 temperature and energy variables. Fig. 3 shows the domain of experimentation and so the model definition. The
156 raw data of signals measured every 3 s (temperature and power) have been smoothed by mean of moving
157 average on 10 measurements. Data collection was carried in such a way that the data sets intended for learning
158 are representative of those intended for validation phase and of the entire operating range of the mixer.

159

160

161 Fig. 3. Experimental conditions variations according to definition domain.

162

163

164 Fig. 4. A streaming example from database of powers and temperatures measurements.

165 Measured (up) and smoothed (down) for six different operating conditions.

166

167

168 Fig. 5. Temperature evolution during kneading process (mixing + texturing) according to an experiment given.

169

170

171 The observation of experimental data showed maximum values of temperature ($T > 25$ ° C) and power until 1200
172 watts. The power evolution suggests that the dough consistency increases, reaches a maximum before decreasing
173 gradually (Fig. 4, curves ▽ and +). This phenomenon corresponds to over-kneading, and is frequently associated
174 with the appearance of a strong stickiness that can be disadvantageous for the dough handling. Moreover, for
175 identical conditions of kneading process, increasing the speed of kneading is accompanied by an increase of

176 shear forces within the dough. For example, [Kilborn and Tipples \(1972\)](#) observed the existence of a minimum
177 threshold of power below which the bread volume decreases strongly in the CBP process (Chorleywood Bread
178 Process), and obtained an optimal development of dough for a minimum energy contribution equal to 40 kJ.kg^{-1}
179 of dough. Moreover, according to these authors, a kneading with constant power implying a varying speed
180 would significantly reduce the time of dough development. In this work, the analysis of various data used for
181 modelling, shows that the power variable has a significant influence since the higher powers promote the
182 emergence of higher values of temperature. This justifies our decision to introduce the power as an input variable
183 to describe correctly the evolution of the dough temperature during kneading process.

184
185 The second part of this work will consider the identification of model parameters. For this, experimental data
186 were separated into two sets: a data set constituting the learning basis upon which the weights assigned to neural
187 network connections are determined, and a data set to validate the model but not used during the learning phase.
188 Twenty kinetics (about 56% of initial database) were used to define the Operation model. The remainder (sixteen
189 kinetic), is used as independent validation set.

190

191

192 **3 Artificial Neural Networks and expert knowledge insertion**

193

194 In our modelling problem, according to the kneading variables, we used the Recurrent Neural Networks (RNR)
195 as dynamical system. This type of networks forms a class of networks where the output of same neurons is used
196 as input to other neurons located upstream in comparison to propagation direction of the information. So that a
197 system is causal, it must be associated any connection with a delay: a recurrent neural network is a dynamics
198 ([Dreyfus et al., 2002](#), [Dreyfus & Ploix, 1998](#)), governed by differential equations. As most applications are
199 carried by computer programs, we stand in the context of discrete-time systems, where differential equations are
200 replaced by difference equations. A feedback neural network of discrete-time is therefore governed by one or
201 more non-linear difference equations, resulting from the composition of functions carried by each neuron and
202 delays associated with each connection. The canonical form of equations governing a recurrent neural network is
203 following:

204

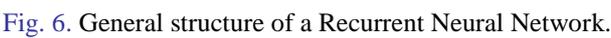
$$x(k+1) = \phi[x(k), u(k)] \quad \text{and} \quad y(k) = \psi[x(k), u(k)] \quad (1)$$

205

206 Where ϕ and ψ are nonlinear functions implemented by a neural network forward, and k designates the
207 discrete time. Fig. 6 (Personnaz & Rivals., 2003) shows the canonical form of a RNN. Any neural network can
208 be put in this canonical form, so completely automatic.

209

210

211  Fig. 6. General structure of a Recurrent Neural Network.

212

213

214 **3.1 Taking into account knowledge in neural modelling**

215

216 The works conducted to model the kneading process (that is: the prediction of the temperature evolution during
217 kneading operation) follow a systematic study based on expert knowledge and physical analysis of phenomena
218 occurring during kneading process (Roussel & Chiron, 2002, Chiron et al., 2008). In fact, we utilized this study
219 to constitute a data set for the learning phase, as well as information on the choice of inputs and outputs model. It
220 was established that the dough temperature and power are influenced by different variables (Fig. 1), and can be
221 described by a dynamic model, with four operating variables. The evolution of the process allows to introduce
222 the power into the neural network as a second state variable representing the thermal source which affects much
223 on the one hand the dough overheating and on the other hand, its texture at end kneading. The idea is that, since
224 a neural network will seek to minimize the cost function, we can build a network for minimizing this function by
225 associating its variables to variables of the modelling problem. For this, the systematic approach by neural
226 networks consists to solve successively the three sub-problems: 1) find a network topology that connections are
227 interpreted as functions carrying the knowledge sought for problem resolution in output system, 2) build a cost
228 function of the problem whose minimum is the best solution, i.e. the convergence conditions must be verified, 3)
229 and deduce the weights of the connections between neurons and input applied to each neuron from the cost
230 function found.

231

232 **3.2 Optimization tools: Learning phase**

233

234 Most learning algorithms of artificial neural networks are optimization algorithms that seek to minimize, by
235 means of nonlinear optimization methods, a cost function which is a measure of the difference between the real
236 responses of the network and its estimated responses. This optimization is carried iteratively, by changing the
237 weights according to the gradient of the cost function. The gradient is estimated by back-propagation method.
238 The back-propagation method has been successfully applied to a wide variety of problems such as, for example,
239 leavening optimization in bread-making (Fravolini, 2003), modelling of thermophilic lactic acid starters
240 production (Acunà, 1995), and determination of the roasting degree of coffee in real-time (Hernandez-Perez,
241 2002). The weights are randomly initialized before training, and then iteratively modified until a satisfactory
242 compromise between approximation accuracy on the training set and the approximation accuracy on a validation
243 set. The learning of neural networks is not particularly slow: there are algorithms for nonlinear optimization
244 extremely fast (Press et al., 1992), which can carry industrial developments with simple computers. Learning of
245 neural networks with feedback connections (to build dynamic models) is very similar to the one of feed-forward
246 networks.

247

248 In this work, the choice of the architecture (number of hidden layers, hidden neurons, and activation function) of
249 the Operation-Model neural network is justified by the universal approximation and parsimonious ANN
250 properties. The number of neurons in the hidden layer was optimized with the Bayesian regularization method
251 (Mackay, 1992) in combination with the Levenberg-Marquardt algorithm allowing weak weights to be penalized
252 (the connections with weak weight are eliminated). In this framework, the weights and biases of the network are
253 assumed to be random variables with specified distributions. The advantage of this algorithm is that it provides a
254 measure of the number of parameters (weights and biases) effectively used by the network. The regularization
255 parameters are related to the unknown variances associated with these distributions. We can then estimate these
256 parameters using statistical techniques. This approach consists of minimizing the cost function $C(w)$, defined in
257 (2), using training and modifying this objective function to:

258

$$C'(w) = C(w) + \alpha \Omega(w) \quad (1)$$

259

260 Where:

261

$$C(w) = \frac{1}{2} \sum_{i=1}^N (y_i^{pred}(w) - y_i^{mes})^2 \quad \text{and} \quad \Omega(w) = \frac{1}{n} \sum_{i=1}^n w_i^2 \quad (2)$$

262 n is the number of network weights, $C(w)$ is the sum of squared errors between y_i^{pred} neural outputs and
 263 y_i^{mes} real outputs. α is a parameter that determines the importance of the two terms in the new performance
 264 function $C'(w)$. Using this, performance functions will cause the network to have smaller weights and biases,
 265 and this will force the network response to be smoother and less likely to over-fit. This method has the advantage
 266 of being simple to implement, since the gradient of C' can be easily calculated from the gradient of C and the
 267 network weights. It is sufficient to add the quantity αw to the gradient vector ∇C calculated by the
 268 backpropagation algorithm:

$$\nabla C' = \nabla C + \alpha w \quad (3)$$

270
 271 The Levenberg-Marquardt algorithm was designed to approach the second order speed of training without
 272 computing the Hessian matrix. When the cost function is the sum of squared errors, the Hessian matrix can be
 273 approximated by:

$$H_k = \nabla^2 C(w_k) \quad (4)$$

275
 276 The update is performed using the following equation:

$$w_{k+1} = w_k - [\eta_{(t)} R_k]^{-1} \nabla C(w_k) \quad \text{with} \quad R_k = H_k + I \quad (5)$$

278
 279 When the learning parameter $\eta_{(t)} = 0$, this is the Newton Method:

$$\Rightarrow w_{k+1} = w_k - H_k^{-1} \nabla C(w_k) \quad (6)$$

280
 281 When $\eta_{(t)}$ is large, this is the gradient method:

(7)

$$\Rightarrow w_{k+1} = w_k - [\eta_k I]^{-1} \nabla C(w_k)$$

282

283 The learning parameter $\eta_{(t)}$ allows to adapt the algorithm to cost function form and to carry a good compromise
 284 between the Newton method ($\eta_{(t)} = 0$), which converges very rapidly towards a minimum, and the simple
 285 gradient method ($\eta_{(t)} \gg 1$), effective far from minima.

286 The learning neural model is normally carried as follows:

- 287 1) Initialization of network parameters;
- 288 2) Propagation through the network temporally unfolded;
- 289 3) Computing the gradient of the cost function (taking into account the weight shared);
- 290 4) Update the parameters network, then return to step 2 for another iteration.

291

292 3.3 Determination of the model performances

293

294 The performances of the optimal model chosen are presented in Section IV. They will be estimated using a
 295 number of indicators presented in this section. In the literature, there are usually two groups of performance
 296 indicators that can be used to assess the quality of models (Willmott et al., 1985). A first group represents the
 297 overall agreement between observed and predicted values. The second group represents the quality of modelling,
 298 in terms of excess thresholds prediction (Dutot et al., 2007). In this work, we used indicators of the first group:
 299 the absolute mean bias (*AMB*) and square root of mean squared error (*SMSE*), as defined below:

300

$$AMB = \frac{1}{N} \sum_{i=1}^N |y_i^{pred} - y_i^{mes}| \quad \text{and} \quad SMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i^{pred} - y_i^{mes})^2} \quad (8)$$

301

302 With y_i^{mes} and y_i^{pred} are measured and predicted values by the model, respectively. Note that a positive value
 303 of mean bias indicates a trend of model to over-estimate, and a negative value indicates the trend to under-
 304 estimate. An additional indicator called *d* (Willmott et al., 1985), reflects the degree to which the observed
 305 variable is correctly estimated by the simulated variable.

$$d = 1 - \frac{\sum_{i=1}^N (y_i^{pred} - y_i^{mes})^2}{\sum_{i=1}^N (|y_i^{pred} - \bar{y}^{mes}| + |y_i^{mes} - \bar{y}^{mes}|)^2} \quad (9)$$

306

307 This indicator is not a measure of correlation in the formal sense, but rather a measure of the degree (based on a
 308 set averaged) predictions without error. It varies between 0 and 1. A value of 1 indicates perfect agreement
 309 between measurement and model, while the value 0 indicates total disagreement.

310

311

312 **4 Results and discussion**

313

314 **4.1 Results of prediction**

315

316 The use of neural prediction involves the choice of input parameters but also the optimization of the architecture
 317 of the neural network itself. The optimization of both the type of nonlinear function and the neural network
 318 architecture (number of neurons in the hidden layer) was made after numerous empirical testing. According to
 319 [Lepage et Solaiman \(2003\)](#), for a neural network with one hidden layer, the recommended number of neurons
 320 approximately equals $\sqrt{N(M+2)}+1$, where N is the number of input parameters and M is the number of
 321 neurons in the output layer. In our case, we initially used a multilayer perceptron with 6 inputs (Tr, M, Te, Tf, P
 322 (t), T (t)), 6 sigmoid (to generate a degree of non-linearity between the neurone's input and output) in the hidden
 323 layer and, in output, the two state variables P(t+1) and T (t+1). So there are 56 connections/or weight at the
 324 beginning of learning. The final qualified network employs approximately 37% of the effective parameters, with
 325 a percentage error of 4.5% on the training set. All operations carried through different network links, summarizes
 326 the energy exchange between the inputs and outputs of the system. [Fig. 7](#) presents the optimal architecture
 327 adopted for power and temperature prediction. [Fig. 8 to Fig. 10](#) provide a visual overview of model performance
 328 for each validated kinetic. For all curves involved that of validation phase, we adopted in this report the
 329 following convention of symbols: the curve with marks (shown on each point) represents the power and the
 330 temperature measured belonging to the validation database, the continuous line curve representing the variables

331 predicted by the model, and in dashed lines the confidence interval at 95% of data predicted. Fig. 10 presents a
332 simulation result for 4 kinetic of temperature and power variables on various conditions.

333

334

335 [Fig. 7.](#) Operation-Model architecture.

336

337

338 [Fig. 8](#) presents the results for the power prediction. On [Fig. 8a](#) it is clear that the prediction is good and the
339 model reaches to reproduce correctly the dynamic variability. The different curves show a good dynamic
340 behaviour of the network; the envelopes are properly fulfilled. Looking at the correlation curves, this overall
341 good behaviour is confirmed, but it appears in some situation that the input variables which are not taken into
342 account, and probably other variables, have a significant effect. It is the same with temperature prediction ([Fig.](#)
343 [9a-b](#)). Nevertheless, it is evident the good correspondence for temperature and power between measure and
344 model. For power, the network response is a bit out of step with the real output. Consequently, the network will
345 generalize well to new data. The neural Operation-Model has a global prediction rate (on the two variables) close
346 to 91%. The various performance indicators and correlation coefficients between the outputs predicted and real
347 outputs are presented in Table 1.

348

349

350 [Fig. 8a.](#) Powers measured and predicted by means of Operation-Model on validation phase.

351

352

353 [Fig. 8b.](#) Correlations between power measured and that predicted by means of Operation-Model on validation
354 phase.

355

356

357 [Fig. 9a.](#) Temperatures measured and predicted by means of Operation-Model on validation phase.

358

359

360 Fig. 9b. Correlations between temperature measured and that predicted by means of Operation-Model on
361 validation phase.

362

363

364 Fig. 10. A streaming example from simulation result of temperature and power variables (on 4 data series) in
365 various conditions.

366

367

368 Table 1. Indicators of performance on validation data set for power and temperature variables.

369

370

371 Fig. 10 illustrates how the model is working and presents some pure simulation from input variable shown on the
372 figure. Although performances, in terms of *AMB* and *SMSE*, are best for temperature, the index *d* indicates a
373 better attractiveness for two variables. This coefficient is between 0.28 and 0.8 for power and 0.99 for
374 temperature. Apart from the kinetic 14 ($d = 0.28$, $R^2 = 0.37$), which is far from the desired performances, the
375 model well reproduces the dynamics variability of the two recorded outputs. The nonlinearity between the two
376 state variables is still well taken into account by the model, since the general trend of power profiles is respected
377 on the whole, if we discard the data stemming from the kinetics 14, the two coefficients *d* and R^2 increase
378 respectively from 0.27 to 0.82 and 0.37 to 0.76.

379

380 A residues analysis measure-model showed in Fig. 11 highlights the fact that the Operation-Model presents the
381 best prediction performance for temperature and power. The information part, provided by each connection of
382 the neural network on real observations, is also higher. For temperature variable, the model explains 99% of
383 variance and has near-nil dispersion around the ideal line. For the power variable, the model has a visible
384 dispersion, but low compared to all data set. The *SMSE* of the model at this level is 9.85 to 52.7 but the
385 variation of almost all residues is systematically lower at 1.97 standard deviation, in other words, the deviation is
386 very small and is bounded by the interval [-2, 2].

387

388

389 Fig. 11. Standardised Residual according to predicted values of power and temperature variables.

390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419

4.2 Analysis of relative importance of input variables

The importance of each variable must be determined so as to confirm the influence of parameters selected in neural network input on the temperature and power evolutions. However, the function implemented by the ANN is complex. The analysis of the contribution of each variable is extremely difficult. Indeed, from the input variables, the network is able to predict the output parameter, but the inner mechanisms of network are completely ignored. The interpretation of an ANN is not so evident as the one of simple linear regression models. Thus, various methods have been implemented to characterize the ANN "black-box". These methods, of quantitative nature, which allow analyzing the contribution of input variables on the ANN prediction. Olden et al. (2004) assessed the pertinence of each method to estimate the contribution of variables as far as ANN and 9 methods described in the literature were tested. The results show that one of them, called "connections weights approach", is often the most accurate and reliable. We have therefore decided to use this approach in order to examine the contribution of predictor variables on the two model outputs. The connections weights approach uses the weights of input-hidden connections and hidden-output connections of an ANN. According to (Olden et al., 2004), it provides the best methodology for precise identification of the importance parameters of model. This approach allows therefore to successfully identify the real importance of all variables of neural network. It computes the result of the 'input-hidden' and 'hidden-output' connexions for each input neuron and output neuron, and recaps the sum of results through all hidden neurons. The relative contribution of independent variables on the network prediction depends on the magnitude of the numerical value obtained for each of the variables considered. An input variable with strong weights connections represents high-intensity on signal transfer. In other words, it has a strong contribution on the network prediction in comparison to other variables characterized by smaller weight connections. The computation procedure consists in multiplying the value of the weights connection of 'hidden-output' neurons, for each hidden neuron, by value of the weights connection of 'input-hidden' layer. The sum of previous products, for each input neuron, is then performed. In this work, the same architecture adopted has been tested with only two inputs M and Vr, and then with the flour temperature as a third input variable. By performing the process described above for each input neuron, we identify its contribution to the output. In Fig. 12 we report the product carried and the sum gives. The various indices and indicators performance are illustrated in Table. 2 and Table. 3.

420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449

Fig. 12. Variables effect on the two outputs T(t+1) and P(t+1) according to connections weights approach.

For two architectures (Fig. 12a and Fig. 12c), the two operating variables, mixture mass and rotor speed, have a greater influence than the other two variables Tf and Te. We noted also the same for the two state variables, temperature T(t) and power P(t). The power and rotor speed inform the model on the stress intensity carried during kneading process. Since the temperature of the flour and that of water affect significantly the dough behaviour in terms of its overheating at the end kneading, we can verify by this method that the Operation Model with four independent inputs (Fig. 12c) is best adapted to model the power and temperature dynamics. The results of various performance indicators presented in the tables below (here without the kinetic 14) are almost identical for the three configurations. However, the coefficients margin d and R^2 , on validation for the power variable, change if we keep kinetic 14.

Although at the systemic study performed about kneading process, the 4 variables were selected as key-information, proof that their influence on the two outputs is not useless in the absolute, with the data set considered for this modelling problem. Therefore, the results of this analysis indicate that, in the context of network optimization, it is possible to keep these variables (including the least influents, in our case 'Ta') without disrupting significantly the results of neural modelling.

Table. 2. Validation results for tree architectures tested.

Table.3. Indicators performance on validation for tree architectures tested.

5 Conclusion

450 In this paper, we clearly proved through the validation results of Operation-Model that the kneading process is
451 relatively nonlinear with respect to temperature and power variations. We were able to predict the dynamics of
452 these two variables from four operating variables: the mixture mass, the rotor speed, the flour temperature, and
453 water temperature, and the return of motional. For this purpose, we used a data set covering the variations space
454 relative to ingredients mixing and operating conditions of kneading. The experimental results showed the
455 effectiveness and reliability of the neural approach for this type of complex operation. Model performance
456 depends on the quality and quantity of data available for learning. The neural model has a prediction rate close to
457 91%. Like any dynamic modelling approach, the use of artificial neural networks requires that the available data
458 are sufficiently numerous and representative. From the perspective of development time, it is clear that the
459 neural approach is advantageous compared to other techniques, where the deployment of other data requires the
460 built of a new model that will be much easier to carry by learning. The model variables remain the same (in a
461 first approximation and unless major evolution of technique), so that learning of a new neural model boils down
462 to elaborate a representative database, and to implement an usual learning algorithms, by varying the model
463 complexity by the number of hidden neurons. It is always desirable and often possible to use for network design,
464 the mathematical knowledge or physical laws, which are available for the phenomenon to model: neural
465 networks are not necessarily "black boxes". As a first perspective of this work, we plan to expand this type of
466 dynamic modelling approach to predict the rheology and sensory proprieties of bread dough at the end kneading
467 process. The final aim is to aggregate and deal all information streaming from these different operations
468 (fermentation or proofing..., and baking) and to detect defects and faults as soon as possible, so as to understand
469 the mechanisms that influence the quality of the final product.

470

471

472 **Acknowledgements**

473

474 The authors gratefully thank all members of the ANR-INCALIN project, for their constructive comments and
475 their assistance during this study. We would like also to thank Hubert Chiron and the MC2 group from the BIA
476 Research Unit (INRA Nantes) for their cooperation and assistance in this project.

477

478 **References**

479

480 [Acunà, G.](#) (1995). Contribution à la modélisation et l'estimation de la production de levains lactiques
481 thermophiles. UNIV Thesis, Institut National Polytechnique de Grenoble, France.
482

483 [Bimbenet, J.J.](#), Schubert, H., & Trystram, G. (2007). Advances in research in food process engineering as
484 presented at ICEF9. *Journal of Food Engineering*, 78, 390-404.
485

486 [Binding, D.M.](#), Couch, M.A., Sujata, K.S., & Webster, M.F. (2003). Experimental and numerical simulation of
487 dough mixing in filled geometry. *Journal of Food Engineering*, 58, 111-123.
488

489 [Bloksma, A.H.](#) (1990). Rheology of breadmaking process. *The American association of cereal chemists*, 35(2),
490 228-236.
491

492 [Bloksma, A.H.](#), & Nieman, W. (1975). The effect of temperature on some rheological properties of wheat flour
493 doughs. *Journal of Texture Studies*, 6, 343-361.
494

495 [Broyart, B.](#), & Trystram, G. (2002). Modelling heat and mass transfer during the continuous baking of biscuit.
496 *Journal of Food Engineering*, 51, 47-57.
497

498 [Charun, E.](#), Abecassis, J., Contamine, A.S., Roulland, T. M, Vergnes, B., & Morel, M.H. (2000). Effects of
499 Temperature and Mechanical Input on Semisweet Biscuit (Cookie) Quality and Dough Characteristics (1).
500 *Cereal Chemistry*, 77 (3), 265-271.
501

502 [Chin, N.L.](#), & Campbell, G.M. (2005). Dough aeration and rheology: Part 2. Effects of flour type, mixing speed
503 and total work input on aeration and rheology of bread dough. *Journal of the Science of Food and Agriculture*,
504 85(13), 2194-2202.
505

506 [Chiron, H.](#), Chaunier, L. Della Valle, G., & Morel-Fatio, L. (2008). Mixing behaviour: Energy balance for
507 oblique axis to spiral mixer. *Industries des Céréales*, 159, 32-33
508

509 [Contamine, A.S.](#), Abecassis, J., Morel, M.H., Vergnes, B., & Verel, A. (1995). Effect of Mixing Conditions on
510 the Quality of Dough and Biscuits. *Cereal Chemistry*, 72 (6), 516-522.
511

512 [Dreyfus, G.](#), Martinez, J. M., Samuelides, M., Gordon, M. B., Badron, F., Thiria, S., & Hérault, L. (2002).
513 Réseaux de neurones : Méthodologie et applications. (ed). France: Eyrolles, (chapter 1 and 5).
514

515 [Dutot, A.L.](#), Rynkiewicz, J., Steiner, F., & Rude, J. (2007). A 24-h forecast of Ozone peaks and exceedance levels
516 using neural classifiers and weather predictions. *Environmental Modelling & Software*, 22, 1261-1269.
517

518 [Fravolini, M.L.](#), Ficola, A., & La Cava, M. (2003). Optimal operation of the leavening process for a bread-
519 making industrial plant. *Journal of Food Engineering*, 60, 289-299.
520

521 [Hernandez-Perez, J. A.](#) (2002). Etude de la torréfaction : modélisation et détermination du degré de torréfaction
522 du café en temps réel. UNIV Thesis, Ecole Nationale supérieure des industries agricoles et alimentaires, 81-89,
523 France.
524

525 [Hlynka, I.](#) (1962). Influence of temperature, speed of mixing and salt on some rheological properties for dough
526 farinograph. *Cereal Chemistry*, 39, 286-303.
527

528 [Kilborn, R.H.](#), Tipples, K.H. (1972). Factors affecting mechanical dough development. I. Effect of mixing
529 intensity and work input. *Cereal Chemistry*, 49, 34-47.
530

531 [Lamrini, B.](#), Trystram, G., Della Valle, G., Chiron, H., Trelea, I.C., & Perrot, N. (2009). A dynamic artificial
532 neural network based model to describe bread dough proofing. *EFFoST conference*, 11-13 November, Budapest,
533 Hungry.
534

535 [Lepage, R.](#), Solaiman, B. (2003). Les réseaux de neurones artificiels et leurs applications en imagerie et en vision
536 par ordinateur. (ed). Montreal : Presses de l'Ecole de technologie supérieure, ISBN 2-921145-40-5.

537 [Levavasseur, L.](#), Rakotozafy, L., Manceau, E., Louarme, L., Robert, H., Baret, J.L., Potus, J., & Nicolas, J.
538 (2006). Discrimination of wheat varieties by simultaneous measurements of oxygen consumption and
539 consistency of flour dough during mixing. *Journal of the Science of Food and Agriculture*, 86 (11), 1688-1698.
540

541 [MacKay, D.J.C.](#) (1992). A Practical Bayesian Framework for Backpropagation Networks. *Neural Computation*,
542 4(3), 448-472.
543

544 [Mann, G.](#), Diffey, S., Allen, H., Pumpa, J., Natha, Z., Morell, M.K., & Cullis, B. (2008). A. Smith. Comparison
545 of small-scale and large-scale mixing characteristics: Correlations between small-scale and large-scale mixing
546 and extensional characteristics of wheat flour dough. *Journal Cereal Science*, 47, 90–100.
547

548 [Ndiaye, A.](#), Chiron, H., Della Valle, G., & Roussel, P. (2005). A Qualitative Representation of French
549 Breadmaking Process in View to Develop a Computerised Decision Support System: the Mixing Unit Process.
550 *Acta Horticulturae*, 585-592.
551

552 [Ndiaye, A.](#), Della Valle, G., & Roussel, P. (2009). Qualitative modelling of a multi-step process: The case of
553 French bread making. *Expert Systems with Applications*, 36(2), Part 1, 1020-1038.
554

555 [Norton, T.](#), & Sun, D. W. (2006). Computer fluid dynamics (CFD) – an effective and efficient design and
556 analysis tool for the food industry. *Trends in Food Science & Technology*, 17(11), 600-620.
557

558 [Olden, J.D.](#), Joy, M.K., & Death, R.G. (2004). An accurate comparison of methods for quantifying variable
559 importance in artificial neural networks using simulated data. *Ecological Modelling*, 178, 389-397.
560

561 [Olmos-Perez, A.](#) (2003). Contribution à l'optimisation de la conduite des procédés alimentaires, UNIV Thesis,
562 Ecole Nationale supérieure des industries agricoles et alimentaires, 81-89, France.
563

564 [Perrot, N.](#), Trystram, G., Le Gennec, D., & Guely, F. (1996). Sensor fusion for real time quality evaluation of
565 biscuit during baking : Comparison between Bayesian and Fuzzy approaches. *Journal of Food Engineering*, 29,
566 301-315.

567

568 [Personnaz, L.](#), Rivals, I. (2003). Réseaux de neurones formels pour la modélisation, la commande et la
569 classification. (ed). CNRS, France, Chapter I.

570

571 [Press, W.H.](#), Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. (1992). Numerical Recipes in C: the Art of
572 scientific Computing. (ed). Cambridge University Press.

573

574 [Roussel, P.](#), & Chiron, H (2003). Les pains français : évolution, qualité, production. (ed). Erti, France, Chapter 3
575 and 5.

576

577 [Rousu J.](#), Flander, L., Suutarinen, M., Autio, K., Kontkanen, P., & Rantanen, A. (2003). Novel computational
578 tools in bakery process data analysis: a comparative study. *Journal of Food Engineering*, 57(1), 45-56.

579

580 [Shehzad, A.](#), Chiron, H., Della Valle, G., Kansou, K., Ndiaye, A., Réguerre, L. (2010). Porosity and stability of
581 bread dough determined by video image analysis for different compositions and mixing conditions. *Food*
582 *Research International*, 43: 1999-2005.

583

584 [Tanaka, K.](#), Bushuk, W. (1973). Changes in Flour Proteins during Dough-Mixing. I. Solubility Results.
585 *American Association of Cereal Chemists*, 50, 590-596.

586

587 [Willmott, C.J.](#), Ackleson, S.G., Davis, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., ODonnell, J., & Rowe,
588 C. (1985). Statistics for the evaluation and comparison of models. *Journal of Geophysical Research*, 90, 8995-
589 9005.

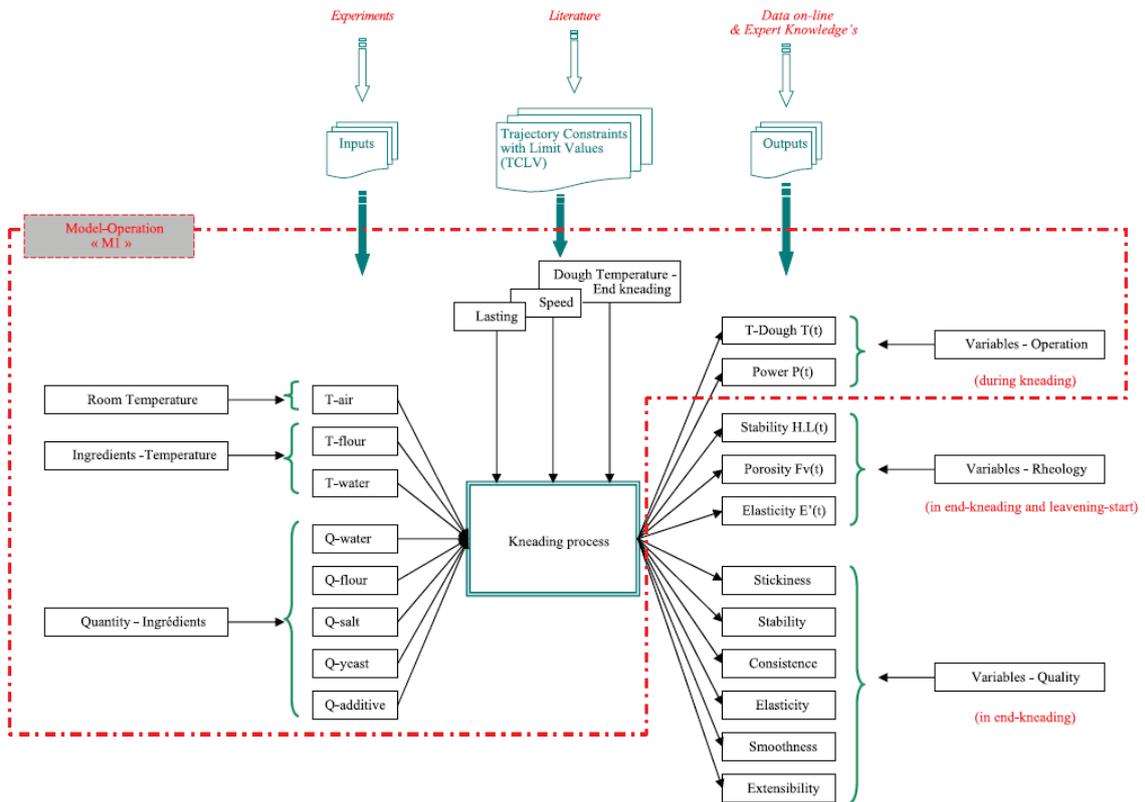


Fig. 1. Representative diagram of kneading process and its environment.

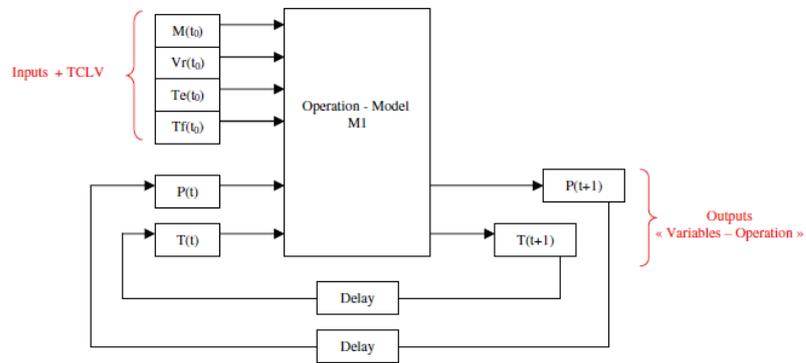


Fig. 2. Operation-Model structure adopted to predict the temperature dynamics of the bread dough and that of the power delivered during kneading process.

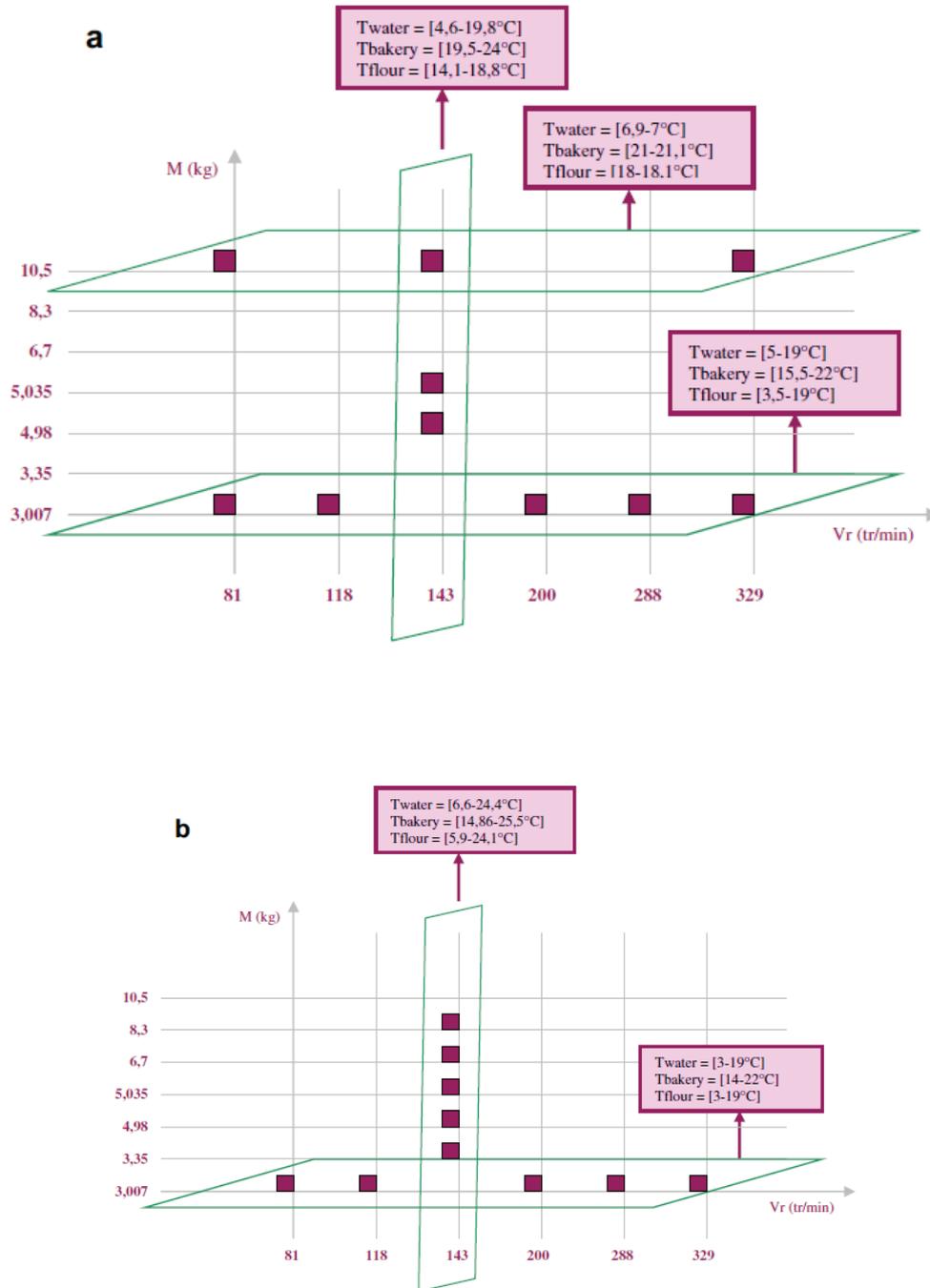


Fig. 3. Experimental conditions variations according to definition domain. a) Data learning. b) Data validation. For all information about the data experimentation, you can contact: blamrini@yahoo.fr and Guy.Della-valle@nantes.inra.fr.

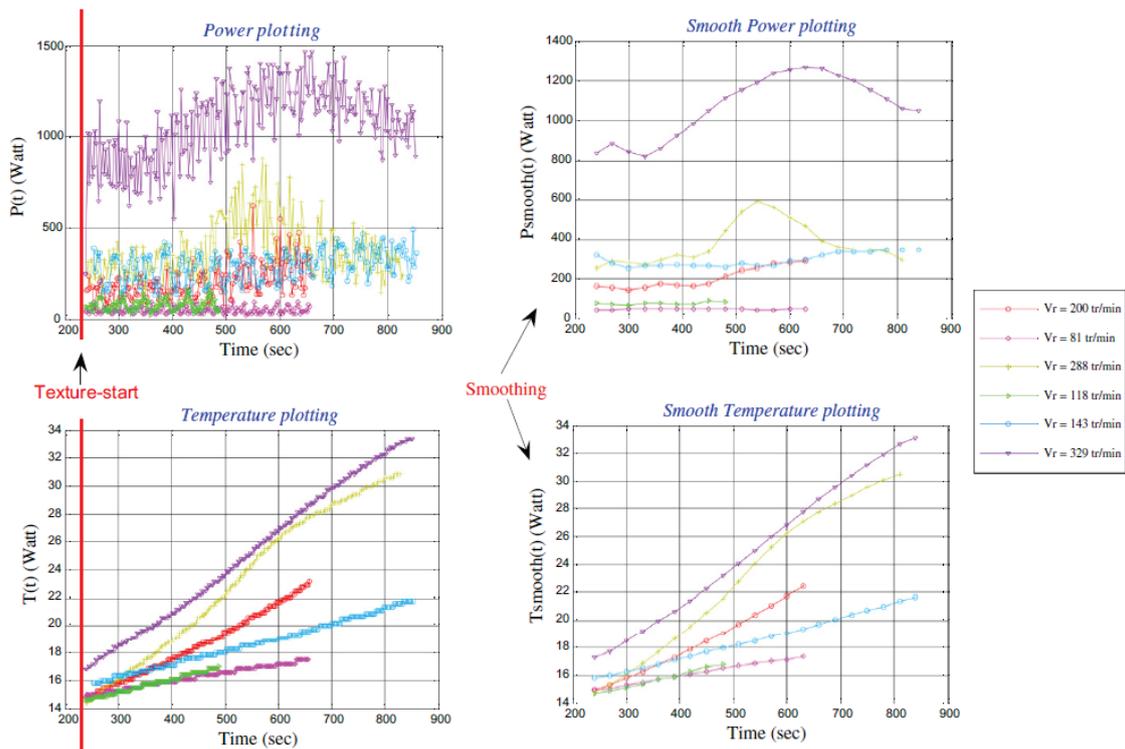


Fig. 4. An example of powers and temperatures measurements stemming from database. Measured (up) and smoothed (down) for six different operating conditions.

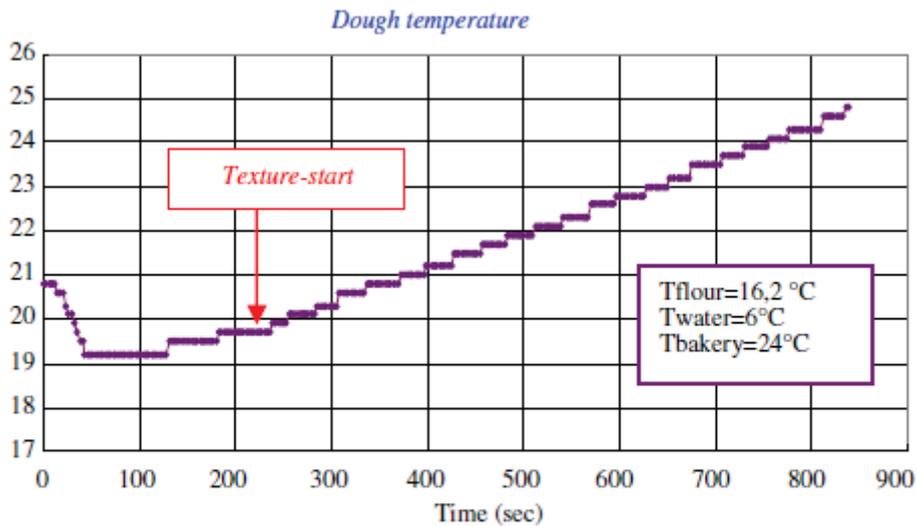


Fig. 5. Temperature evolution during kneading process (mixing + texturing) according to an experiment given.

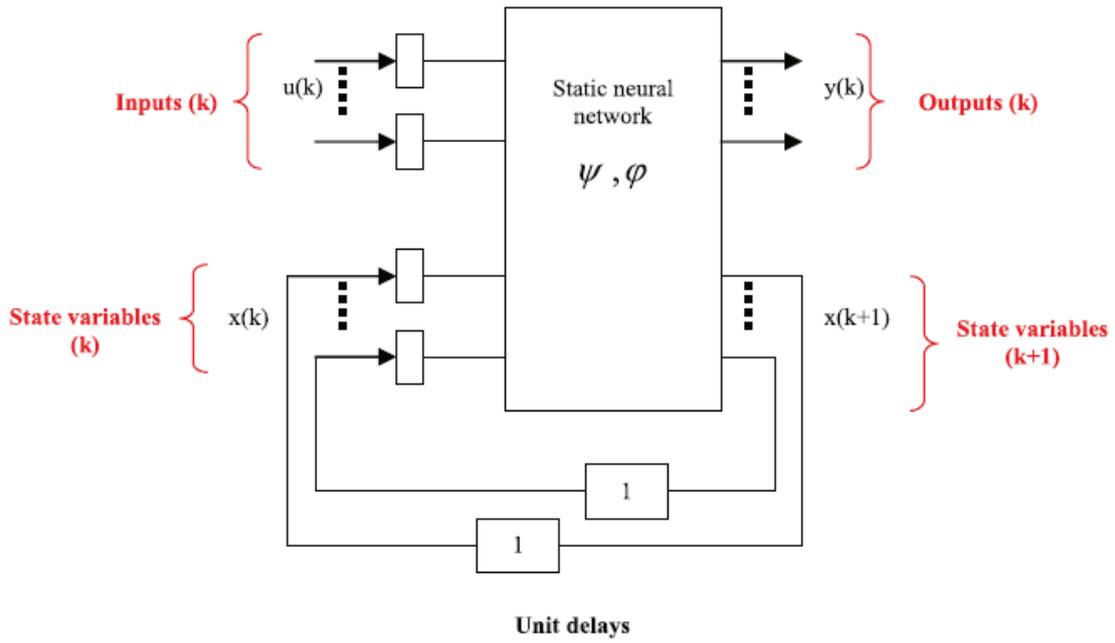


Fig. 6. General structure of a recurrent neural network.

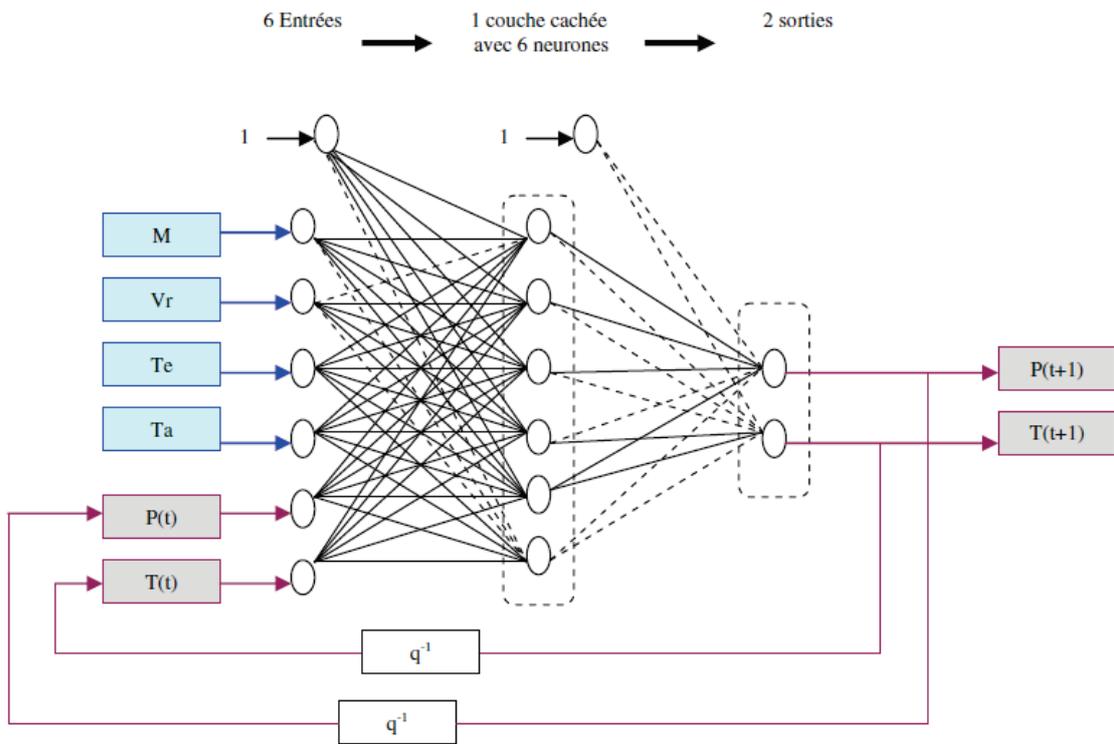


Fig. 7. Operation-Model architecture.

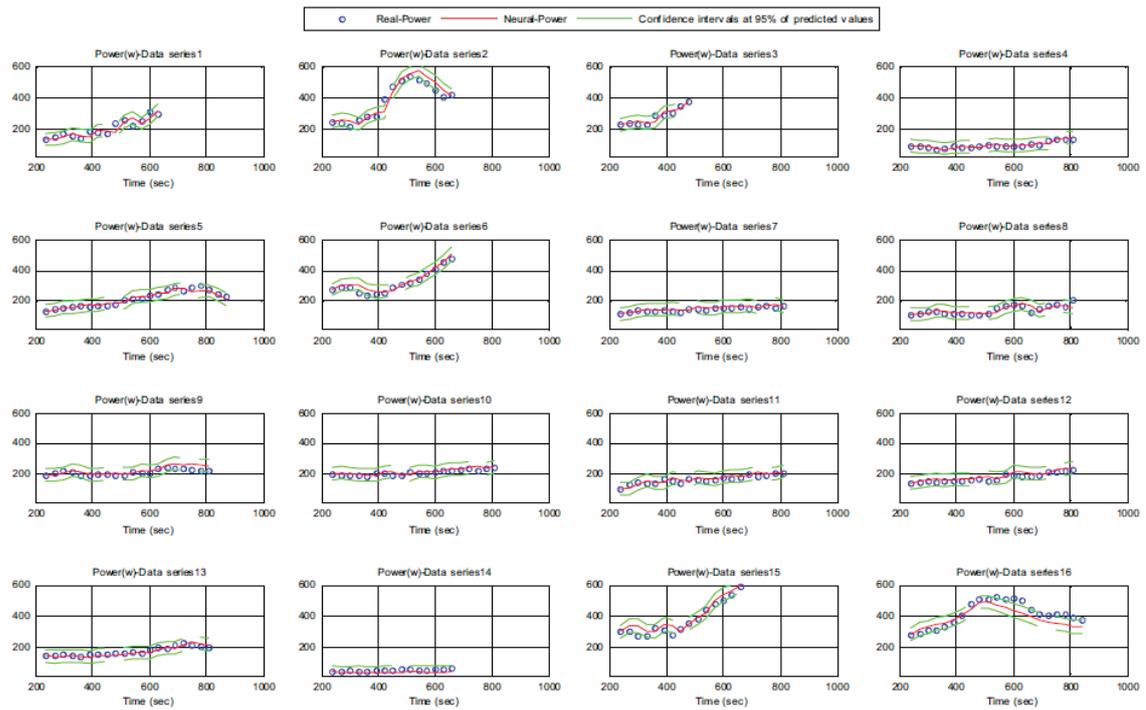
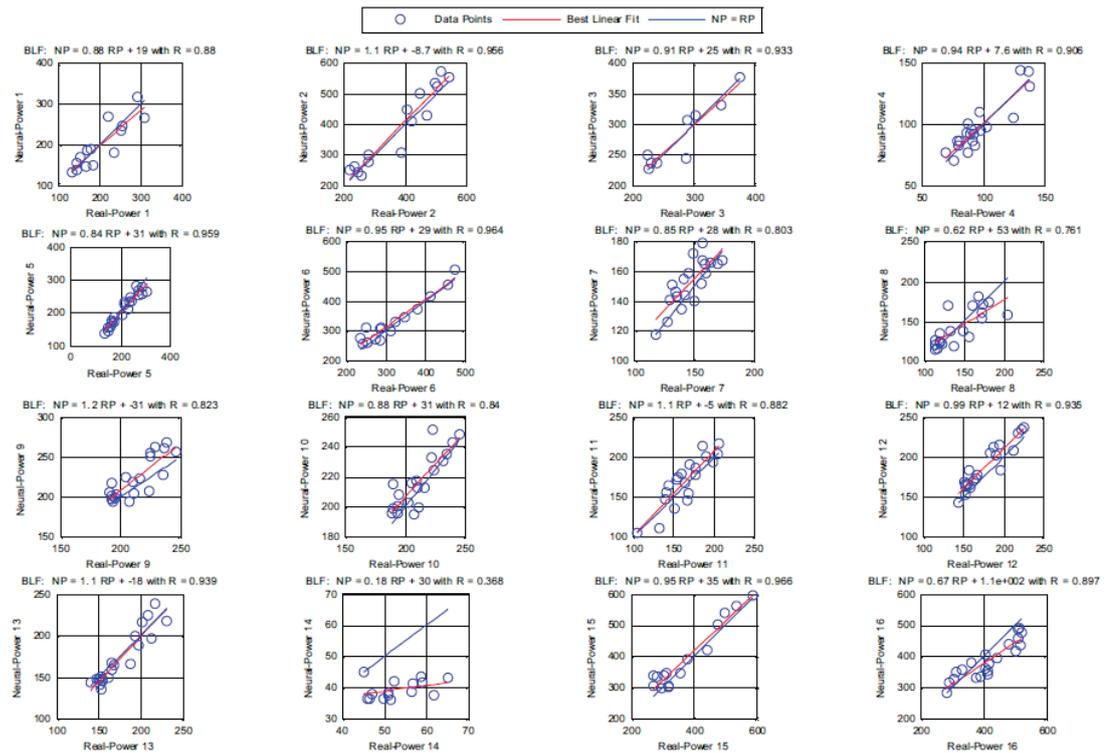
a**b**

Fig. 8. a. Powers measured and predicted by means of Operation-Model on validation phase. b. Correlations between power measured and that predicted by means of Operation Model on validation phase.

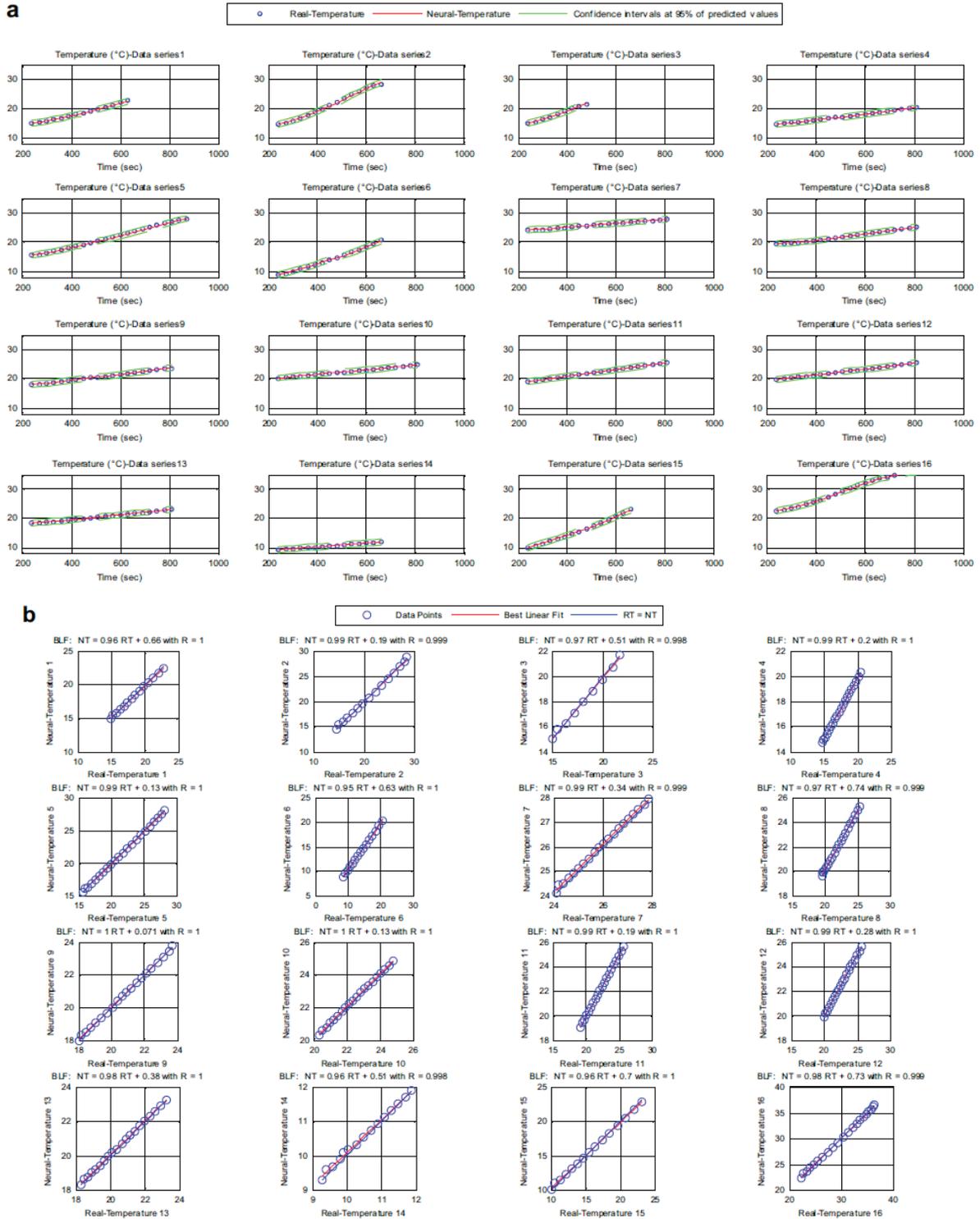


Fig. 9. a. Temperatures measured and predicted by means of Operation-Model on validation phase. b. Correlations between temperature measured and that predicted by means of Operation-Model on validation phase.

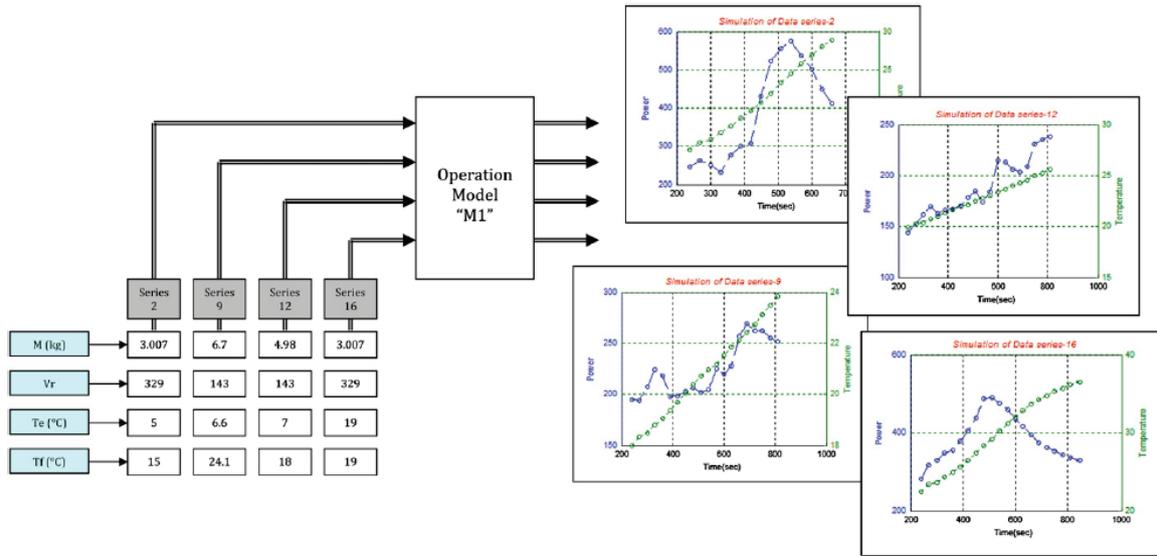


Fig. 10. A streaming example from simulation result of temperature and power variables (on 4 data series) in various conditions.

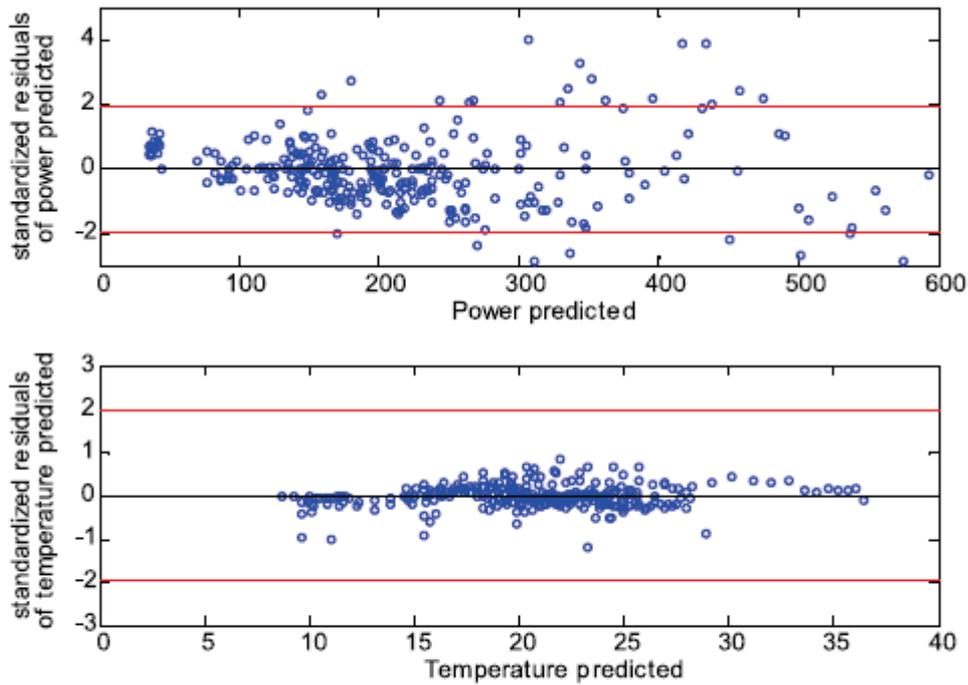
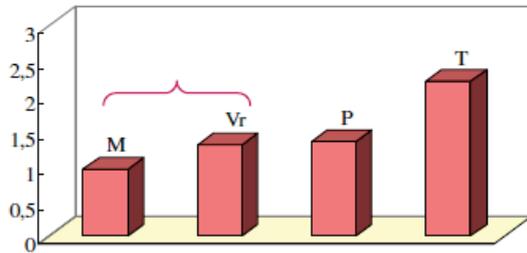
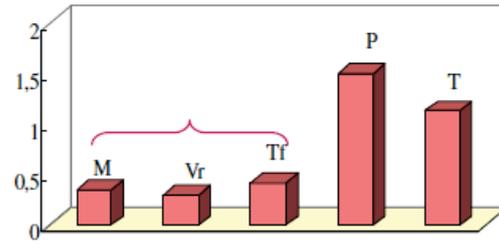


Fig. 11. Standardised Residual according to predicted values of power and temperature variables.

a Inputs importance in comparison to outputs model M1, with 2 independant inputs M and Vr



b Inputs importance in comparison to outputs model M1, with 3 independant inputs M, Vr and Tf



c Inputs importance in comparison to outputs model M1, with 4 independant inputs M, Vr, Tf and Te

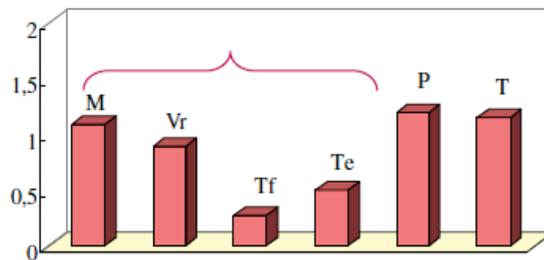


Fig. 12. Variables effect on the two outputs $T(t + 1)$ and $P(t + 1)$ according to connections weights approach.

Table.1. Indicators of performance on validation data set for power and temperature variables.

	d	R^2	AMB	$SMSE$
Temperature	0,99	0,99	[0,04-0,2]	[0,0452-0,2584]
Power	[0,28-0,8]	[0,37-0,96]	[2,55-7,65]	[9,85-52,7]

Table.2. validation results for tree architectures tested.

		% ERROR ON VALIDATION	EFFECTIVE PARAMETERS OF TOTAL	NBRE OF ITERATIONS
M1 with 4 inputs (M, Vr, Tf, Te)	Temperature	0,53%	37% of 56	800
	Power	6,97%		
M1 with 3 inputs (M, Vr, Tf)	Temperature	0,43%	35% of 50	500
	Power	6,21%		
M1 with 2 inputs (M, Vr)	Temperature	0,45%	52% of 44	450
	Power	6,8%		

Table.3. Indicators performance on validation for tree architectures tested.

		<i>d</i>	<i>R</i> ²	<i>AMB</i>	<i>SMSE</i>
<i>M1</i> <i>4 inputs</i>	Temperature	0,99	0,99	[0,04-0,2]	[0,0452-0,26]
	Power	[0,82-0,97]	[0,76-0,96]	[7,65-52,55]	[9,85-52,7]
<i>M1</i> <i>3 inputs</i>	Temperature	0,99	0,99	[0,036-0,21]	[0,04-0,245]
	Power	[0,73-0,97]	[0,73-0,97]	[7,45-26,15]	[10,50-32,58]
<i>M1</i> <i>2 inputs</i>	Temperature	0,99	0,99	[0,045-0,21]	[0,05-0,25]
	Power	[0,77-0,98]	[0,65-0,97]	[8,74-35,13]	[10,31-36]