Ekonomika Economics
Management Management
Inovace Innovation

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

COMPLEX SYSTEMS MODELLING USING EXTENDED FUZZY NON-LINEAR REGRESSION ANALYSIS

MODELOVÁNÍ KOMPLEXNÍCH SOUSTAV METODOU ROZŠÍŘENÉ FUZZY NELINEÁRNÍ REGRESNÍ ANALÝZY

Miroslav Pokorný
Moravská vysoká škola Olomouc, Ústav informatiky a aplikované matematiky
miroslav.pokorny@mvso.cz

Abstract:
The parameters and sometimes even the structure of the real complex systems are not exactly known in advance, and which moreover often change in unpredictable way. Adequacy of models of such complex systems must be then ensured by identification of their structure and parameters. One of suitable method of modelling of such ill-known and difficult measure systems appears a fuzzy non-linear regression analysis represets by Takagi-Sugeno fuzzy model. In the paper an extended TS model is presented with the regression coefficients in the shape of fuzzy numbers. The difference between the actual and computed values of the dependent variable in the new fuzzy regression models - with the fuzzy regression coefficients - are due to the "indefiniteness" of the system structure and parameters and - in the end - the level of system fuzziness is expressed through to fuzziness of model output variable.

Abstrakt:
Parametry a někdy i struktury reálných komplexních systémů nejsou přesně známy a navíc se často mění nepředvídatelným způsobem. Adekvátnost modelů komplexních systémů je nutno zajistit odpovídající identifikaci jejich struktury a parametrů. Jednou z vhodných metod modelování těchto špatně definovaných a obtížné měřitelných systémů je fuzzy nelineární regresní modelování typu Takagi-Sugeno. V práci je prezentován rozšířený model TS model, jehož regresní koeficienty jsou definovány ve tvaru fuzzy čísel. Rozdíl mezi skutečnými a vypočtenými hodnotami závislé proměnné jsou pak dány neurčitostí struktury systému a parametrů. Úroveň neurčitosti systému je vyjádřena prostřednictvím neurčitosti fuzzy čísla jako hodnoty výstupní proměnné modelu.

Key words:
Complex system, vagueness, fuzzy set, fuzzy logic, fuzzy regression, extended T-S model.

Klíčová slova:
Složitý systém, vágnost, fuzzy množina, fuzzy logika, fuzzy regrese, rozšířený TS model.

JEL Classification: C51, C63

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

1 Introduction

Many studied processes are represented by systems, the parameters and sometimes even the structure of which are not exactly known in advance, and which moreover often change in unpredictable way. Adequacy of models of such complex systems must be then ensured by identification of their structure and parameters. The suitable modelling methods are usually products of scientific field of Artificial intelligence [1], [2].

One of suitable method of modelling of complex, ill-known and difficult measure systems appears a fuzzy linear and fuzzy non-linear regression analysis [7].

The difference between the actual and computed values of the dependent variable in the fuzzy linear regression models - with the fuzzy regression coefficients - are due to the “indefiniteness“ of the system structure and parameters and - in the end - the level of system fuzziness is expressed through to fuzziness of model output variable.

The non-linear regression fuzzy models type Takagi-Sugeno (TS-models) include the linear regression equations with the crisp regression coefficients [5], [8]. Therefore, the global output of classical TS-model models is in the crisp form and, in this case, there is not possible to judge an output variable fuzziness. There is disadvantage in the case when the TS-model is used in tasks of system state prediction or fouls estimation. On the other hand, the crisp form of output variable is advantage in the case when the TS-model is used in tasks of system control (Sugeno controller).

The new fuzzy non-linear regression model which is presented in my paper reflects the level of system fuzziness better - through a fuzzy global output variable [12].

2 Fuzzy linear regression model (flrm)

The fuzzy parameters of the fuzzy linear regression model (FLRM) represent the described system fuzziness. The fuzzy parameters of the FLRM denoted $\underline{k}$ are defined through the normal convex fuzzy sets (fuzzy numbers). The fuzzy linear regression model is given in the form [9]

$$\underline{y}^* = \underline{k}_0 + \underline{k}_1 . x_1 + ... + \underline{k}_n . x_n$$

where $\underline{y}^*$ is the computed output variable in the fuzzy form, $\underline{k}_j$ are the fuzzy regression parameters in the form of triangular fuzzy numbers [3], [5].

In case of fuzzy linear regression model FLRM, the one regression relation validity exists in full scale of the input variables. It means that the model validity exists in full n- dimension input space defined for n- input variables.

3 Fuzzy non-linear regression model (fnlrm)

If the hypothesis exists that in the various intervals of input variables x exist the various regression relations

$$y = f(x)$$

the division of input space is necessary together with the definitions of the appropriate partial regression models. Thus it is possible to design the production rules set which represents the rule based model of system under study with higher degree of fitting than the simple fuzzy linear regression model.

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

The fuzzy rules therefore define the division of fuzzy input space, and at the same time in each sub-spaces the appropriate input/output regression relation is defined. Thus, *r*-th rule is then constructed in the form [4], [5], [6], [7]

$$\mathrm{R_r}: \mathrm{IF}\left(x_1 \text{ is } \underline{A}_{1r}\right) \text{and} \ldots \text{and} \left(x_n \text{ is } \underline{A}_{nr}\right)$$

$$\text{THEN } \; y_r^* = k_{0r} + k_{1r}.x_1 + \ldots + k_{nr}.x_n$$

where r = 1, 2, ..., R is the number of rules. The widely known fuzzy non-linear regression model FNLRM which is defined in this way is named Takagi-Sugeno (T-S model) [10], [11]. The partial result of rules y_r^* are calculated using the regression formulas and the global resulting value of whole FNRLM model y^* is given using expression

$$y^* = \frac{\sum_{r=1}^{R} w_r . y_r^*}{\sum_{r=1}^{R} w_r}$$

The weight value w_r is given as the minimum relation

$$w_r = \min_j \mu_{\underline{A}_{jr}} . \left(x_j^0\right)$$

where $\mu_{\underline{A}_{jr}}(x_j^0)$ is grade of membership of sampled values x_{ij}^0 to the fuzzy set $\underline{A}_{jr}$.

The classical T-S model is considered using the rules in which the consequent regression coefficients are identified as the non-fuzzy (crisp) numbers k. Therefore, the global output of classical T-S model y^* is in the crisp form and, in this case, there is not possible to judge the fuzziness of output variable. The crisp output variable in the fuzzy form is advantage in the case when the T-S model is used in tasks of system control (Sugeno controller). But, there is disadvantage in the case when the T-S model is used in tasks of system state prediction or foults estimation.

4 Fuzzy extended non-linear regression model (ferm)

To rich the fuzzy form of global output variable of fuzzy non-linear regression model the more general regression method was proposed. This method is one in which the consequent regression coefficients of the fuzzy rules are identified as the fuzzy numbers $\underline{k}$ and now the global output of such FERM is in the fuzzy number form $\underline{y}^*$ as well. The r- th rule of such FERM model is then constructed in the form

$$R_r: \mathrm{IF}\left(x_1 \text{ is } \underline{A}_{1r}\right) \text{and} \ldots \text{and} \left(x_n \text{ is } \underline{A}_{nr}\right)$$

$$\text{THEN } \; \underline{y}_r^* = \underline{k}_{0r} + \underline{k}_{1r}.x_1 + \ldots + \underline{k}_{nr}.x_n$$

The construction of model FERM involves the identification procedures of the regression coefficients (consequent parameters) in triangular fuzzy numbers form.

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

5 IDENTIFICATION OF THE MODEL FERM

5.1 Premise Structure Identification

Identification of the premise structure consists of two tasks: determination of optimal structure of independent variables and determination of optimal partition of fuzzy space of input variables as a proper problem of fuzzy modelling.

The algorithm of structure optimization is starting its identification process with one rule (linear model) and then is increasing the number of rules with simultaneous observation of the trend of the value of *FCR* criterion as a lost function. Assuming the *m*-th step of identification process the model has *m*- rules.

To determine the fitting criterion the simple Kondo criterion was chosen in form

$$FCR = \frac{100}{N} \cdot \sum_{i=1}^{N} \frac{\left| y_i^0 - y_i^* \right|}{y_i^0} \quad [\%]$$

The search is stopped and identification process is finished, if the minimal values of *FCR* criterion in (*m*-1)-th and *m*-th steps are in relation

$$FCR(m) > FCR(m-1)$$

and then model in (*m*-1)-th step is described as optimal one.

5.2 Premise Parameters Identification

The estimation of parameters of the premise is performed in such a way, so that the error of the estimate of dependent variable decreases in the process of identification. For observed values (x_1^0, x_2^0, ... , x_n^0, y^0) the error E of estimate y^* is formulated using formula

$$E = y^0 - y^* = y^0 - \frac{\sum_{r=1}^{R} \left[\min_j \ \mu_{\underline{A}_{rj}} \left(x_j^0 \right) \right] . y_r^*}{\sum_{r=1}^{R} \min_j \ \mu_{\underline{A}_{rj}} \left(x_j^0 \right)}$$

The identification algorithms determines new parameters of fuzzy sets $\underline{A}_{rj}$ under condition

$E \to$ min.

Identification of the consequent is realized for variables of the premise given in previous phase of identification of the premise.

5.3 Consequent Parameters Identification

Given vector of inputs $x_{ij}^0 = (1, x_{i1}^0, x_{i2}^0, \ldots, x_{in}^0)$ we can calculate partial values of outputs

$$y_r^* = k_0(w_r) + k_1(w_r).x_1 + \ldots + k_n(w_r).x_n = \sum_{j=0}^{n} k_j(w_r).x_j$$

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

The regression coefficients $k_j(w_r)$ are now the non-linear coefficients, dependent on value of the weight coefficient w_r . The regression task is solved using linear programming method for the vector of transformed input variables $\mathbf{z}_i$ with elements z_{ij} where

$$z_{ij} = g_{ir} . x_{ij}^0$$

and

$$g_{ir} = \frac{w_{ir}}{\sum_{r=1}^{R} w_{ir}}$$

Then, we solve the linear programming task with objective function min$J$ and constraints

$$\alpha^T . z_i + (1 - \mathrm{H}) . \sum_{j=0}^{n} c_j . \left| z_{ij} \right| \geq y_i^0 + (1 - \mathrm{H}) . e_i$$

$$- \alpha^T . z_i + (1 - \mathrm{H}) . \sum_{j=0}^{n} c_j . \left| z_{ij} \right| \geq - y_i^0 + (1 - \mathrm{H}) . e_i$$

5.4 Consequent Structure Identification

The procedure of structure identification of consequents is very simple. If the value of regression coefficient k_j is lower than determined limit K_{min} (where K_{min} is very small number approaching zero), i.e.

$$k_j \leq \mathrm{K}_{\min}$$

then the variable x_j is implicite eliminated from the consequent.

6 Computer program modul enfis

To realize the method FERM the special computer program modul called ENFIS (Extended Non-Linear Fuzzy Identifical System) was developed and created.

The block CFG determines all the constants and parameters of future computing. The input blocks PD1 - PD3 perform the data acquisition and data preliminary processing. The blocks PP1 - PP4 are used for heuristic searching of sub-models premise structure and for determination of the sub-models rules. The blocks IM1 - IM6 realize the proper procedures of structure and parameters identification and evaluate the fitting and stop criterions. The output block OUT serves for presentation of the results in suitable output form.

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

7 Numerical examples

Langari´s Synthetical Model and ENFIS Efficiency Comparison

In this example the fuzzy regression modelling using expanded system ENFIS is applied to the empirical data to illustrate its power in comparison with other systems. The set of 24 data points was used to identify fuzzy model.

Using the input/output data points the following model and results were obtained using the procedure ENFIS [the value of fuzziness (> *0*) of partial consequent coefficients are written in the brackets]:

R1: IF (*x1* is *SMALL*) & (*x2* is *SMALL*) THEN *y1 = 0.24x1 + 0.51x2 + 0.01(0.01)*
R2: IF (*x1* is *LARGE*) & (*x2* is *SMALL*) THEN *y2 = 0.35x1 + 0.53x2 + 0.00(0,18)*
R3: IF (*x1* is *SMALL*) & (*x2* is *LARGE*) THEN *y3 = 0.62x1 + 0.40x2 + 0.00(0.12)*
R4: IF (*x1* is *LARGE*) & (*x2* is *LARGE*) THEN *y4 = 0.47x1 + 0.51x2 + 0.00(0.04)*

J = 0.36, FCR = 10.35%

The premise parameters of two linguistic terms (SMALL, LARGE - both of trapezoidal shapes) of two linguistic premise variables (x_1, x_2) were identified as follows:

SMALL {*0.00, 0.00, 0.30, 0.70* }
LARGE {*0.30, 0.70, 1.00, 1.00* }

The fitness of desirable values y^0 and predicted values y^* of input variable *y* is descripted using the graph in Fig. 2.

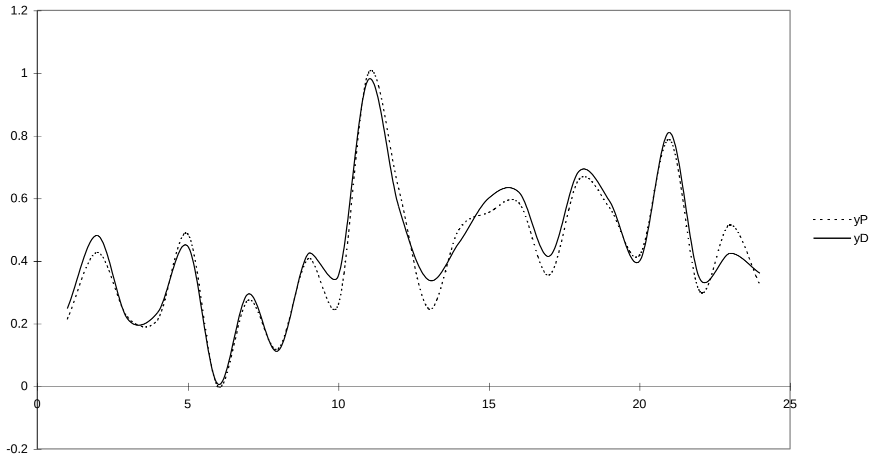

Fig. 1 Measured values y^0 and predicted values y^* of input variable *y*

Now the fitting criterion was composed as sum of square errors.

$$Q = \sum_{k=1}^{24} (y_k^0 - y_k^*)^2$$

To judge the obtained results the comparison with some other methods is presented and the prediction power of fuzzy model ENFIS is proved [13].

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

METHOD	FITTING CRITERION
Zimermann (1980)	0.080
Dyckhoff (1984)	0.066
Krishnapuram (1992)	0.064
FERM – ENFIS	**0.049**
Langari (1995) - Hybrid Learning Algorithm	0.044
Langari (1995) - Supervised Learning Algorithm	0.043

Tab. 1 Quality of selected modelling methods comparison [13]

It can be seen that system ENFIS compare favourably with these existing algorithms.

Fuzzy Regression Coeficients Identification

The final part of extended TS model identification procedure we can see in Tab.1. The search of the best structure is presented through last four submodels. The best of them is this one with 8 rules and lost function value equal 1,05%. The vagueness of regression coefficients is expressed trough the half basics of their triangular shape membership functions (in brackets).

```
Ucelova funkce  4.92
x1V         y1= 1.23( 0.68)*x1+ 1.52( 0.45)*x2+ 2.91( 0.00)*x3+ 0.00( 0.00)
x1M x2M x3V y2= 1.46( 0.00)*x1+ 1.19( 0.00)*x2+ 0.73( 0.16)*x3+ 0.00( 0.00)
x1M x2V x3V y3= 2.77( 0.00)*x1+ 1.14( 0.39)*x2+ 0.00( 0.00)*x3+11.07( 0.00)
x1M     x3M y4= 1.22( 0.10)*x1+ 2.31( 3.09)*x2+ 1.51( 0.00)*x3+ 1.00( 0.05)
Ztratova funkce:   10.27[procent]
Zmena presnosti ulohy LP 0.001000
Ucelova funkce  2.34
x1V         y1= 1.23( 0.68)*x1+ 1.52( 0.45)*x2+ 2.91( 0.00)*x3+ 0.00( 0.00)
x1M x2M x3M y2= 1.19( 0.07)*x1+ 0.15( 0.00)*x2+ 0.66( 0.00)*x3+ 1.00( 0.05)
x1M x2V x3M y3= 2.56( 0.00)*x1+ 2.15( 0.28)*x2+ 2.93( 0.00)*x3+ 5.28( 0.00)
x1M     x3V y4= 2.16( 0.00)*x1+ 1.65( 0.00)*x2+ 0.00( 0.81)*x3+ 6.06( 0.00)
Ztratova funkce:    5.53[procent]
Zmena presnosti ulohy LP 0.001000
Ucelova funkce  6.06
x1V x2M     y1= 1.11( 0.52)*x1+ 2.23( 0.00)*x2+ 3.00( 0.00)*x3+ 0.00( 0.00)
x1V x2V     y2= 1.30( 0.00)*x1+ 1.52( 0.85)*x2+ 2.78( 0.00)*x3+ 0.00( 0.00)
x1M     x3M y3= 1.20( 0.06)*x1+ 2.54( 3.55)*x2+ 1.52( 0.00)*x3+ 1.00( 0.05)
x1M     x3V y4= 2.35( 0.00)*x1+ 1.95( 0.00)*x2+ 0.26( 1.02)*x3+ 0.00( 0.00)
Ztratova funkce:   13.16[procent]
Zmena presnosti ulohy LP 0.001000
Ucelova funkce  1.79
x1M x2M x3M y1= 1.22( 0.09)*x1+ 0.13( 0.00)*x2+ 1.36( 0.18)*x3+ 1.00( 0.05)
x1M x2M x3V y2= 1.46( 0.00)*x1+ 1.17( 0.00)*x2+ 0.55( 0.12)*x3+ 1.95( 0.00)
x1M x2V x3M y3= 2.44( 0.00)*x1+ 2.17( 0.26)*x2+ 2.93( 0.00)*x3+ 5.56( 0.00)
x1M x2V x3V y4= 2.63( 0.00)*x1+ 1.37( 0.20)*x2+ 0.39( 0.00)*x3+ 5.85( 0.00)
x1V x2M x3M y5= 0.96( 0.12)*x1+ 2.93( 0.00)*x2+ 2.43( 0.00)*x3+ 0.00( 0.00)
x1V x2M x3V y6= 1.32( 0.23)*x1+ 1.96( 0.00)*x2+ 2.95( 0.07)*x3+ 0.00( 0.00)
x1V x2V x3M y7= 1.77( 0.09)*x1+ 0.90( 0.05)*x2+ 1.84( 0.00)*x3+ 0.00( 0.00)
x1V x2V x3V y8= 2.23( 0.20)*x1+ 1.90( 0.10)*x2+ 1.73( 0.02)*x3+ 0.00( 0.00)
Ztratova funkce:    1.05[procent]
```

Tab.1 Final results of extended TS model identification using procedure ENFIS

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

8 CONCLUSION

The proposed fuzzy non-linear regression model FERM involving the fuzzy linear regression equations is interesting through the vague phenomenon of the system structure/parameters which is better reflected. The realization of the computer program ENFIS which solves the tasks of construction FERM and its parameters identification enables application of the proposed method in modelling of the real complex systems.

Acknowledgements:
This work has been supported by Project GAČR P403-12-1811: Unconventional Managerial Decision Making Methods Development in Enterprise Economics and Public Economy.

References

[1] RUSSEL,S., NORVIG,P.: Artificial Intelligence, Prentice-Hall, Inc., 1995, ISBN 0-13-360124-2, second edition 2003, ISBN 0-13-080302-2, third edition 2010, ISBN 0-13-604259-7
[2] MUNAKATA, T.: Fundamentals of the New Artificial Intelligence, Springer-Verlag New York, Inc., 2008. ISBN 978-1-84628-838-8
[3] NOVÁK,V.,PERFILIEVA,I.,MOČKOŘ,J. *Mathematical Principles of Fuzzy Logic.* Kluwer, Boston. 1999. ISBN 0-7923-8595-0.
[4] CELIKZILMAZ,Asli, TURKSEN,Burkhan. *Modelling Uncertainty with Fuzzy Logic. S*pringer-Verlag.2009. ISBN 978-3-540-89923-5.
[5] NOVÁK,Vilém. *Základy fuzzy modelování.* BEN Praha, 2000, ISBN 80-7300-009-1
[6] ALIEV,R.A, ALIEV,R.R.: Soft Computing and its Application, World Scientific Publishing Co. Pte. Ltd., 2001, ISBN 981-02-4700-1
[7] KACPRZYK,J. FEDRIZZI,M. (Ed.). *Fuzzy Regression Analysis. Studies in Fuzziness and Soft Computing,* ISBN-13: 978-3790805918. Publisher: Physica-Verlag HD, 1992.
[8] BARDOSSY,A. Note on fuzzy regression. *Fuzzy Sets and Systems*, volume 37, pages 65–75, 1990.
[9] BUCKLEY, J.J. JOWERS, L.J. Fuzzy Linear Regression I. *Studies in Fuzziness and Soft Computing*, volume 22, ISBN 978-3-540-76289-8. Springer, 2008.
[10] SHAPIRO,A.F. Fuzzy regression models. [online]. [cit. 2013-0$-10]. Dostupné z: http://www.soa.org/library/research/actuarial-research-clearing-house/2006/january/arch06v40n1-ii.pdf
[11] CETINTAV,B. ZDEMIR,F. *LP Methods for Fuzzy Regression and a New Approach* E.Krause, Ed. In Synergies of Soft Computing and Statistics for Intelligent Data Analysis, volume 22, ISBN 978-3-642-33041-4. Springer, 2013.
[12] POKORNÝ,M. *Fuzzy nelinearní regresní analyza* (*Fuzzy Non-Linear Regression Analysis*). Doctoral Thesis (in Czech). VUT Brno, FEL, Brno, 1993.[13] LANGARI,R. WANG,L. Fuzzy Models, Modular Networks and Hybrid Learning. *Fuzzy Sets and Systems,* 79, 1996, pp.141/150

EMI, Vol. 7, Issue 1, 2015
ISSN: 1804-1299 (Print), 1805-353X (Online)

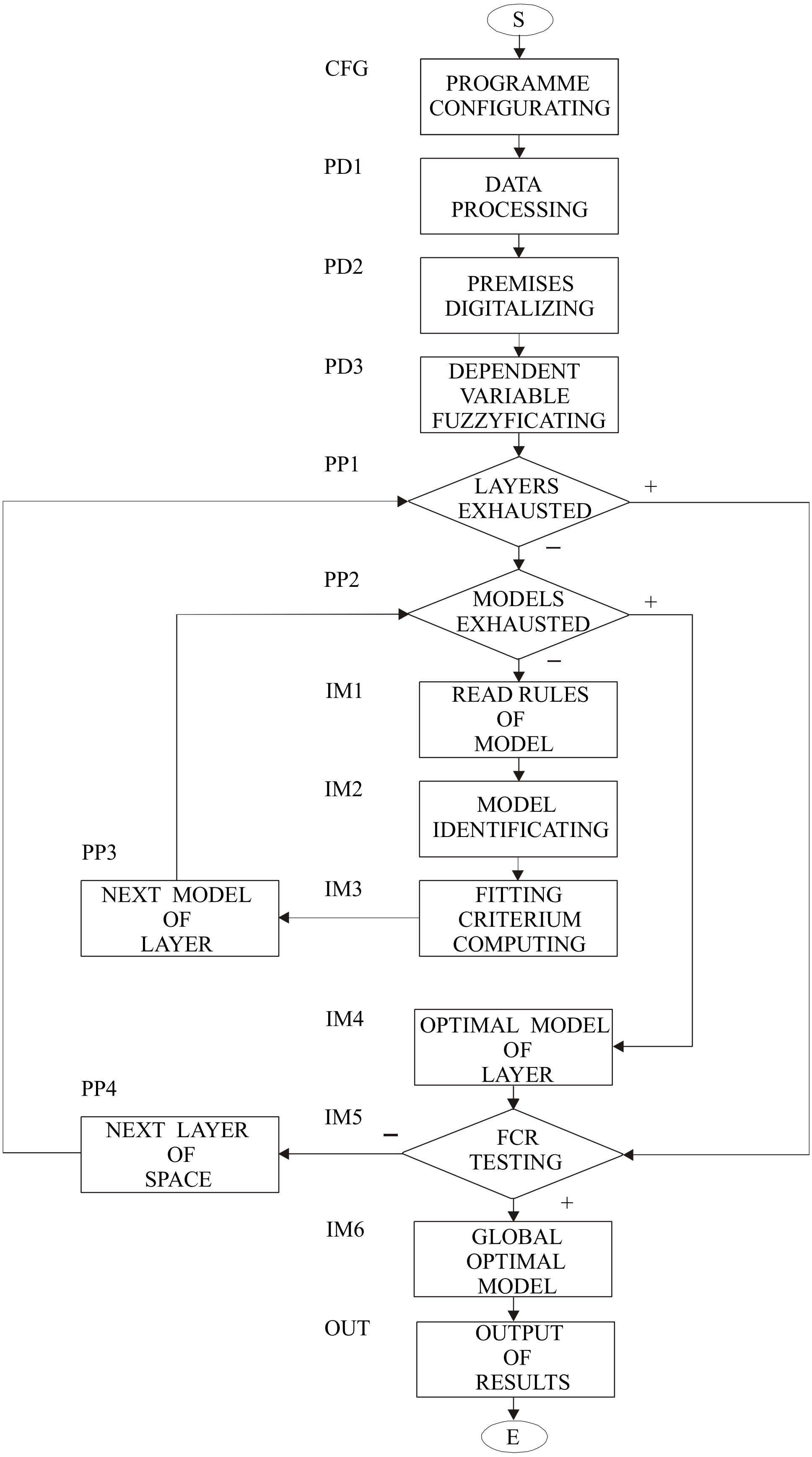

Fig.2 Flow diagram of programme system ENF