InterCriteria analysis results
based on different number of objects

Dafina Zoteva and Olympia Roeva

Institute of Biophysics and Biomedical Engineering
Bulgarian Academy of Sciences
105 Acad. G. Bonchev Str., Sofia 1113, Bulgaria
e-mails: dafy.zoteva@gmail.com, olympia@biomed.bas.bg

Received: 22 September 2017
Accepted: 27 October 2017

Abstract: InterCriteria Analysis (ICrA) results based on different number of objects are investigated in this paper. To evaluate the influence of the number of objects, data from parameter identification procedures of an E. coli fed-batch fermentation process model are used. Model parameters are estimated applying 100 genetic algorithms with different mutation rate values. Seven different index matrices are constructed for ICrA. The results show that the number of objects in ICrA is important for the reliability of the obtained results.

Keywords: InterCriteria Analysis, Intuitionistic fuzzy sets, Genetic algorithms, Mutation rate, E. coli.

2010 Mathematics Subject Classification: 03E72.

1 Introduction

A contemporary approach for multicriteria decision making, named InterCriteria Analysis (ICrA), is proposed in [4]. This approach implements the means of the index matrices (IM) and intuitionistic fuzzy sets (IFS), aiming a comparison of predefined criteria and the objects estimated by them.

ICrA has been applied for the first time in the field of model parameter identification of fermentation processes (FP) using genetic algorithms (GAs) in [10]. Series of papers with ICrA applications in this area have been published since then, for example [1, 14, 15]. ICrA has been
proven to be an appropriate approach for establishing the correlations between model and optimization algorithm parameters, when given parameters are considered as criteria. The reported results confirm some existing dependencies that are based on the physical meaning of the FP model parameters and the stochastic nature of GAs.

There is a lack of studies in the literature about the influence of the number of objects on the ICrA results. It is important to determine the number of objects that are sufficient to obtain reliable results from the ICrA application. It can be assumed that the larger the number of objects are used in the analysis, the more reliable the results will be. But, at least how many objects are enough to be able to rely on the results?

The current research is an attempt to investigate the impact of the different number of objects on the InterCriteria analysis results in the particular test case. It is a very important issue concerning the application of ICrA in the field of model parameter identification of FP. Any additional exploring of the FP model is valuable in the case of modelling living systems, such as FP. Moreover, the relation between mathematical model and optimization algorithm will be established. In order to improve both mathematical modelling and optimization algorithm performance reliable and secure results are needed.

Data from series of parameter identification procedures of an E. coli fed-batch fermentation model are used to construct several IMs with different number of objects. ICrA is applied over the so defined IMs and the results are discussed.

The paper is organized as follows: Section 2 presents the background of ICrA. Numerical results and discussion are presented in Section 3 and conclusion remarks are given in Section 4.

2 InterCriteria analysis

Following [4] and [2], an Intuitionistic Fuzzy Pair (IFP), as the degrees of “agreement” and “disagreement” between two criteria applied on different objects, will be obtained. As a remainder, an IFP is an ordered pair of real non-negative numbers \( \langle a, b \rangle \), such that \( a + b \leq 1 \).

For clarity, let an IM [3], whose index sets consist of the names of the criteria (for rows) and objects (for columns), be given. The elements of this IM are further supposed to be real numbers, which is not required in the general case. An IM with index sets, consisting of the names of the criteria, and IFPs, corresponding to the “agreement” and “disagreement” of the respective criteria, as elements will be obtained.

Let \( O \) denotes the set of all objects being evaluated, and \( C(O) \) is the set of values assigned by a given criteria \( C \) (i.e., \( C = C_p \) for some fixed \( p \)) to the objects, i.e.,

\[
O \overset{\text{def}}{=} \{O_1, O_2, O_3, \ldots, O_n\}, \quad C(O) \overset{\text{def}}{=} \{C(O_1), C(O_2), C(O_3), \ldots, C(O_n)\}.
\]

Let \( x_i = C(O_i) \). Then the following set can be defined:

\[
C^*(O) \overset{\text{def}}{=} \{\langle x_i, x_j \rangle | i \neq j \land \langle x_i, x_j \rangle \in C(O) \times C(O)\}.
\]

Further, if \( x = C(O_i) \) and \( y = C(O_j) \), \( x \prec y \) will be written iff \( i < j \). The vectors of all internal comparisons for each criterion are constructed in order to find the agreement of different criteria. The elements of the vectors fulfil one of the three relations \( R, \overline{R} \) and \( \tilde{R} : \)

111
\[ (x, y) \in R \iff (y, x) \in \overline{R}, \quad (1) \]
\[ (x, y) \in \hat{R} \iff (x, y) \notin (R \cup \overline{R}), \quad (2) \]
\[ R \cup \overline{R} \cup \hat{R} = C^*(O). \quad (3) \]

For example, if “\( R \)” is the relation “<”, then \( \overline{R} \) is the relation “>”, and vice versa.

Hence, for the effective calculation of the vector of internal comparisons, denoted further by \( V(C) \), only the subset of \( C(O) \times C(O) \) needs to be considered, namely:

\[ C^<(O) \overset{\text{def}}{=} \{ (x, y) \mid x < y \& (x, y) \in C(O) \times C(O) \}, \]
due to Eqs. (1)–(3). For brevity, \( e^{ij} = (C(O_i), C(O_j)) \). Then, the vector with lexicographically ordered pairs as elements is constructed for a fixed criterion \( C \):

\[ V(C) = \{ e^{1,2}, e^{1,3}, \ldots, e^{1,n}, e^{2,3}, e^{2,4}, \ldots, e^{2,n}, e^{3,4}, \ldots, e^{3,n}, \ldots, e^{n-1,n} \}. \quad (4) \]

Further, the vector \( V(C) \) is replaced with \( \hat{V}(C) \), where for the \( k \)-th component, \( 1 \leq k \leq \frac{n(n-1)}{2} \), it is true:

\[ \hat{V}_k(C) = \begin{cases} 1 & \text{iff } V_k(C) \in R, \\ -1 & \text{iff } V_k(C) \in \overline{R}, \\ 0 & \text{otherwise.} \end{cases} \]

Then, the degree of “agreement” between two criteria, which are to be compared, is determined as the number of the matching components, divided by the length of the vector for the purpose of normalization. This can be done in several ways, e.g. by counting the matches or by taking the complement of the Hamming distance. The degree of “disagreement” is the number of the components of opposing signs in the two vectors, again normalized by the length. This also may be done in various ways.

If the respective degrees of “agreement” and “disagreement” are denoted by \( \mu_{C,C'} \) and \( \nu_{C,C'} \), it is obvious (from the way of computation) that \( \mu_{C,C'} = \mu_{C',C} \) and \( \nu_{C,C'} = \nu_{C',C} \). Also it is true that \( \langle \mu_{C,C'}, \nu_{C,C'} \rangle \) is an IFP.

The sum \( \mu_{C,C'} + \nu_{C,C'} \) is equal to 1 in the most of the obtained pairs \( \langle \mu_{C,C'}, \nu_{C,C'} \rangle \). However, there may be some pairs, for which this sum is less than 1. The difference

\[ \pi_{C,C'} = 1 - \mu_{C,C'} - \nu_{C,C'} \quad (5) \]
is the degree of “uncertainty”.
3 Numerical results and discussion

3.1 Parameter identification of a mathematical model of an E. coli fed-batch fermentation process

The mathematical model of the E. coli fed-batch FP has the form [6]:

\[
\frac{dX}{dt} = \mu X - \frac{F_{in}}{V} X
\]  

\[
\frac{dS}{dt} = -q_s X + \frac{F_{in}}{V} (S_{in} - S)
\] 

\[
\frac{dV}{dt} = F_{in}
\]

where

\[
\mu = \mu_{max} \frac{S}{k_S + S}, \quad q_s = \frac{1}{Y_{S/X}} \mu
\]

and \(X\) is the biomass concentration, [g/l]; \(S\) is the substrate concentration, [g/l]; \(F_{in}\) is the feeding rate, [l/h]; \(V\) is the bioreactor volume, [l]; \(S_{in}\) is the substrate concentration in the feeding solution, [g/l]; \(\mu\) and \(q_s\) are the specific rate functions, [1/h]; \(\mu_{max}\) is the maximum value of the \(\mu\), [1/h]; \(k_S\) is the saturation constant, [g/l]; \(Y_{S/X}\) is the yield coefficient, [-].

Real experimental data for biomass and glucose concentration are used in the model parameters identification. The detailed description of the process conditions and experimental data are presented in [11, 13].

3.2 Optimization criterion

The objective function is considered as a mean square deviation between the experimental data trajectories and the ones predicted by the model, defined as:

\[
J = \sum_{i=1}^{m} (X_{exp}(i) - X_{mod}(i))^2 + \sum_{i=1}^{n} (S_{exp}(i) - S_{mod}(i))^2 \rightarrow \min
\]

where \(m\) and \(n\) are the experimental data dimensions; \(X_{exp}\) and \(S_{exp}\) – available experimental data for biomass and substrate; \(X_{mod}\) and \(S_{mod}\) – model predictions for biomass and substrate with a given model parameter vector, \(p = [\mu_{max} k_S Y_{S/X}]\).

3.3 Genetic algorithm identification

Genetic algorithm, initially presented in Goldberg [8], searches a global optimal solution using three main genetic operators in a sequence selection, crossover and mutation. GA starts with a creation of a randomly generated initial population. Each solution is then evaluated and assigned a fitness value. According to the fitness function, the most suitable solutions are selected. After that, crossover proceeds to form a new offspring. Mutation is next applied with specified probability, aiming to prevent falling of all solutions into a local optimum. The execution of the GA has been
repeated until the termination criterion (i.e. reached number of populations, or when a solution with specified tolerance is found, etc.) is satisfied.

When applying GA, there are many operators, functions, parameters, and settings in the GAs that can be implemented specifically for different problems. For the parameter identification problem considered here, the GA operators and parameters are tuned as follows: crossover operators are double point; mutation operators – bit inversion; selection operators – roulette wheel selection; number of generations – \( maxgen = 100 \); crossover rate – \( p_c = 0.7 \); number of individuals – \( nind = 100 \) and generation gap – \( ggap = 0.97 \). The GA parameter mutation rate \( p_m \) is varied in the range \( p_m = [0.001 : 0.001 : 0.1] \). The \( p_m \) values are chosen based on the results in [7, 9, 12]. While the mutation rate is varied using vector \( p \), the other parameters and operators are kept constant. As a result, 100 differently tuned GAs are produced.

30 independent runs have been performed for each GA. The obtained model parameters estimates \( (\mu_{max}, k_S, Y_{S/X}) \), total computation time and objective function value are recorded. As a result, 3000 records are obtained – 30 estimates of \( \mu_{max}, k_S \) and \( Y_{S/X} \) for each of the 100 GAs, where \( GA_1 \) corresponds to \( p_m = 0.001 \), \( GA_2 \) corresponds to \( p_m = 0.002 \), etc. These results are further processed in order to generate one main IM with elements the average values of the five criteria \( (J(C_1), T(C_2), \mu_{max}(C_3), k_S(C_4), Y_{S/X}(C_5)) \) of every 30 runs (estimates). The rows of the IM represent the five criteria and the columns are the 100 objects (100 differently tuned GAs – \( GA_1, GA_2, ..., GA_{99}, GA_{100} \)):

\[
IM_{average} = \begin{bmatrix}
C_1 & \text{average} & \text{values} \\
C_2 \\
C_3 \\
C_4 \\
C_5
\end{bmatrix}
\]

To investigate the influence of the objects number on ICrA results seven IMs are defined, as follows:

- **\( IM_1 \)** with 11 objects (the results from \( GA_1, GA_{10}, GA_{20}, GA_{30}, GA_{40}, GA_{50}, GA_{60}, GA_{70}, GA_{80}, GA_{90} \) and \( GA_{100} \) are included):

\[
IM_1 = \begin{bmatrix}
GA_1 & GA_{10} & GA_{20} & GA_{30} & GA_{40} & GA_{50} & GA_{60} & GA_{70} & GA_{80} & GA_{90} & GA_{100} \\
C_1 & 4.5166 & 4.5690 & 4.5198 & 4.5017 & 4.5134 & 4.4946 & 4.4935 & 4.4773 & 4.4999 & 4.5380 & 4.5257 \\
C_2 & 62.249 & 61.812 & 61.626 & 61.602 & 61.891 & 61.967 & 61.985 & 62.024 & 62.057 & 61.693 & 62.072 \\
C_3 & 0.4865 & 0.4949 & 0.4910 & 0.4888 & 0.4883 & 0.4878 & 0.4852 & 0.4857 & 0.4853 & 0.4937 & 0.4877 \\
C_4 & 0.0117 & 0.0133 & 0.0126 & 0.0121 & 0.0120 & 0.0119 & 0.0114 & 0.0116 & 0.0114 & 0.0130 & 0.0119 \\
C_5 & 2.0214 & 2.0216 & 2.0217 & 2.0218 & 2.0214 & 2.0215 & 2.0215 & 2.0212 & 2.0215 & 2.0223 & 2.0213
\end{bmatrix}
\]

- **\( IM_2 \)** with 25 objects (the results from \( GA_1, GA_5, GA_9, GA_{13}, GA_{17}, GA_{21}, GA_{25}, GA_{29}, GA_{33}, GA_{37}, GA_{41}, GA_{45}, GA_{49}, GA_{53}, GA_{57}, GA_{61}, GA_{65}, GA_{69}, GA_{73}, GA_{77}, GA_{81}, GA_{85}, GA_{89}, GA_{93} \) and \( GA_{100} \) are included):

\[
(11)
\]
following conclusions could be made: the variation of the number of objects does not affect the results substantially. For example:

- Criteria pair $C_1 - C_2$: the observed $\mu_{C,C'}$ values show that the criteria pair is in strong dissonance and only in the case of $IM_1$ the pair is in dissonance;
- Criteria pair $C_1 - C_5$: in this case the pair is in dissonance for all IMs;

The $IM_1 - IM_7$ are available at http://intercriteria.net/studies/gap/mutr/ (XLS file).

3.4 InterCriteria analysis of results

ICrA is applied on the 7 IMs. The obtained results are presented in the form of IM (12). They are analysed based on the scale proposed in [5], which defines the consonance and dissonance between the criteria pairs (see Table 1).

Based on the values of degree of “agreement”, $\mu_{C,C'}$, and degree of “disagreement”, $\nu_{C,C'}$, the following conclusions could be made:

- The existing relations and dependencies for some of the criteria pairs are so established that the variation of the number of objects does not affect the results substantially. For example:

- Criteria pair $C_1 - C_2$: the observed $\mu_{C,C'}$ values show that the criteria pair is in strong dissonance and only in the case of $IM_1$ the pair is in dissonance;
- Criteria pair $C_1 - C_5$: in this case the pair is in dissonance for all IMs;
|       | $IM_1$ | $IM_2$ | $IM_3$ | $IM_4$ | $IM_5$ | $IM_6$ | $IM_7$ |
|-------|--------|--------|--------|--------|--------|--------|--------|
|       | $\mu_{C,C}$ | $\nu_{C,C}$ | $\mu_{C,C'}$ | $\nu_{C,C'}$ | $\mu_{C,C}$ | $\nu_{C,C}$ | $\mu_{C,C'}$ | $\nu_{C,C'}$ | $\mu_{C,C}$ | $\nu_{C,C}$ | $\mu_{C,C'}$ | $\nu_{C,C'}$ | $\mu_{C,C}$ | $\nu_{C,C}$ |
| $C_1 - C_2$ | 0.40 0.60 | 0.56 0.44 | 0.54 0.45 | 0.53 0.47 | 0.52 0.47 | 0.51 0.49 | 0.49 0.51 |
| $C_1 - C_3$ | 0.80 0.20 | 0.70 0.30 | 0.72 0.28 | 0.73 0.27 | 0.74 0.26 | 0.74 0.26 | 0.75 0.25 |
| $C_1 - C_4$ | 0.80 0.20 | 0.70 0.30 | 0.71 0.29 | 0.73 0.27 | 0.74 0.26 | 0.74 0.26 | 0.75 0.25 |
| $C_1 - C_5$ | 0.62 0.38 | 0.64 0.36 | 0.66 0.34 | 0.63 0.37 | 0.61 0.39 | 0.62 0.38 | 0.63 0.37 |
| $C_2 - C_3$ | 0.24 0.76 | 0.31 0.69 | 0.30 0.70 | 0.33 0.67 | 0.34 0.66 | 0.33 0.67 | 0.31 0.68 |
| $C_2 - C_4$ | 0.24 0.76 | 0.32 0.68 | 0.31 0.69 | 0.33 0.66 | 0.34 0.65 | 0.34 0.66 | 0.32 0.67 |
| $C_2 - C_5$ | 0.20 0.80 | 0.32 0.68 | 0.43 0.57 | 0.31 0.68 | 0.37 0.62 | 0.35 0.65 | 0.34 0.66 |
| $C_3 - C_4$ | 0.96 0.04 | 0.97 0.03 | 0.97 0.03 | 0.97 0.03 | 0.96 0.04 | 0.96 0.04 | 0.96 0.04 |
| $C_3 - C_5$ | 0.75 0.25 | 0.68 0.32 | 0.59 0.41 | 0.62 0.38 | 0.60 0.40 | 0.62 0.38 | 0.63 0.37 |
| $C_4 - C_5$ | 0.71 0.29 | 0.66 0.34 | 0.59 0.41 | 0.60 0.40 | 0.58 0.42 | 0.60 0.40 | 0.61 0.39 |
Table 1: Consonance and dissonance scale [5]

| Interval of $\mu_{C,C'}$ | Meaning                        |
|--------------------------|--------------------------------|
| [0-0.5]                  | strong negative consonance (SNC) |
| (0.5-0.15]               | negative consonance (NC)        |
| (0.15-0.25]              | weak negative consonance (WNC)  |
| (0.25-0.33]              | weak dissonance (WD)            |
| (0.33-0.43]              | dissonance (D)                  |
| (0.43-0.57]              | strong dissonance (SD)          |
| (0.57-0.67]              | dissonance (D)                  |
| (0.67-0.75]              | weak dissonance (WD)            |
| (0.75-0.85]              | weak positive consonance (WPC)  |
| (0.85-0.95]              | positive consonance (PC)        |
| (0.95-1.00]              | strong positive consonance (SPC) |

- Criteria pair $C_3 - C_5$: the here observed $\mu_{C,C'}$ values show that the criteria pair is in dissonance and in the cases of $IM_1$ and $IM_2$ the pair is in weak dissonance;

- Criteria pair $C_4 - C_5$: the observed $\mu_{C,C'}$ values show that the criteria pair is in dissonance and again only in the case of $IM_1$ the pair is in weak dissonance.

As it can be seen, an alteration in the obtained $\mu$ values is observed exactly in the cases of $IM_1$ and $IM_2$, i.e. in the case of a small number of objects. To the contrary, the obtained results are stable, when a larger number of objects is considered.

- For other criteria pairs, the existing relations are not so established and the variation of the number of objects has impact on the results. For example:
  - Criteria pairs $C_1 - C_3$ and $C_1 - C_4$: the obtained $\mu_{C,C'}$ values show that the criteria pairs are in weak dissonance, while in the case of $IM_1$ the pairs are in weak positive consonance;
  - Criteria pairs $C_2 - C_3$, $C_2 - C_4$ and $C_2 - C_5$: the obtained $\mu_{C,C'}$ values show that the criteria pairs are mainly in dissonance (from weak dissonance to dissonance, see Table 1), while in the case of $IM_1$ weak negative consonance between the pairs is observed.

In this case, the use of a small number of objects could lead to some incorrect assumptions. For example, to conclude that a criteria pair is in consonance (dependence) to be made, while in fact the criteria are independent.

- Finally, there is a criteria pair $C_3 - C_4$ for which a high degree of “agreement” is observed no matter the number of used objects in the ICrA. For all IMs $\mu$ values, as 0.96 and 0.97, i.e. strong positive consonance, are obtained.
4 Conclusion

The influence of different number of objects on the InterCriteria analysis results is explored in this paper. The research is done on the basis of particular test case – parameter identification procedures of an \textit{E. coli} fed-batch fermentation model.

Data from 100 series of parameter identification procedures are used to construct several IMs with different number of objects. ICrA is applied on the so defined IMs.

The results show the importance of the number of objects for the reliability of the obtained ICrA results.

Acknowledgement

Work presented here is partially supported by the Bulgarian National Scientific Fund under Grants DFNI-I-02/5 “InterCriteria Analysis – A New Approach to Decision Making” and DN-02/10 “New Instruments for Knowledge Discovery from Data, and their Modelling”.

References

[1] Angelova, M., Roeva, O., & Pencheva, T. (2015) InterCriteria Analysis of a Cultivation Process Model Based on the Genetic Algorithm Population Size Influence. \textit{Notes on Intuitionistic Fuzzy Sets}, 21(4), 90–103.

[2] Atanassov, K. (2012) \textit{On Intuitionistic Fuzzy Sets Theory}, Springer, Berlin.

[3] Atanassov, K. (2010) On Index Matrices, Part 1: Standard Cases. \textit{Advanced Studies in Contemporary Mathematics}, 20(2), 291–302.

[4] Atanassov, K., Mavrov, D., & Atanassova, V. (2014) Intercriteria Decision Making: A New Approach for Multicriteria Decision Making, Based on Index Matrices and Intuitionistic Fuzzy Sets. \textit{Issues in IFSs and GNs}, 11, 1–8.

[5] Atanassov, K., Atanassova, V., & Gluhchev, G. (2015) InterCriteria analysis: ideas and problems. \textit{Notes on Intuitionistic Fuzzy Sets}, 21(1), 81–88.

[6] Bastin, G., & Dochain, D. (1991) \textit{On-line Estimation and Adaptive Control of Bioreactors}, Elsevier Scientific Publications.

[7] Capko, D., Erdeljan, A., Vukmirovic, S., & Lendak, I. (2011) A hybrid genetic algorithms for partitioning of data model in distribution management system. \textit{Information Technology and Control}, 40(4), 316–321.

[8] Goldberg, D. E. (2006) \textit{Genetic Algorithms in Search, Optimization and Machine Learning}, Addison Wesley Longman, London.
[9] Laboudi, Z., & Chikhi, S. (2012) Comparison of genetic algorithms and quantum genetic algorithms. *Int. Arab J Inf. Tech.*, 9(3), 243–249.

[10] Pencheva, T., Angelova, M., Atanassova, V., & Roeva, O. (2015) InterCriteria Analysis of Genetic Algorithm Parameters in Parameter Identification. *Notes on Intuitionistic Fuzzy Sets*, 21(2), 99–110.

[11] Pencheva, T., Roeva, O., & Hristozov, I. (2006) *Functional State Approach to Fermentation Processes Modelling*, Prof. Marin Drinov Academic Publishing House, Sofia.

[12] Prasad, G., Singh, D., Mishra, A., & Shah, V. H. (2017) Genetic Algorithm Performance Assessment by Varying Population Size and Mutation Rate in Case of String Reconstruction. *Journal of Basic and Applied Engineering Research*, 4(2), 157–161.

[13] Roeva, O., Pencheva, T., Hitzmann, B., & Tzonkov, S. (2004) A Genetic Algorithms Based Approach for Identification of *Escherichia coli* Fed-batch Fermentation. *International Journal Bioautomation*, 1, 30–41.

[14] Roeva, O., & Vassilev, P. (2016) InterCriteria Analysis of Generation Gap Influence on Genetic Algorithms Performance. *Advances in Intelligent Systems and Computing*, 401, 301–313.

[15] Roeva, O., Vassilev, P., Fidanova, S., & Paprzycki, M. (2016) InterCriteria Analysis of Genetic Algorithms Performance, *Studies of Computational Intelligence*, 655, 235–260.