Issues in IFS and GNs, Vol. 12, 2015/2016, 139–149

Performance Optimization Approach in InterCriteria Analysis

Todor Kostadinov

Prof. Dr. Asen Zlatarov University
1, Prof. Y. Yakimov Blvd., 8010 Burgas, Bulgaria
e-mail: kostadinov_todor@btu.bg

Abstract: This study addresses the relation between air pollution factors and weather related parameters. By using the InterCriteria Analysis to the combined set of weather and atmosphere related parameters it is possible to discover and predict abnormal reading among the air pollution parameters and certain weather conditions. An approach to increase the computation capabilities of the ICA software has been proposed due to the fact that by applying new objects and criteria to ICA, the computational burden increases.
Keywords and phrases: Optimization algorithms, Performance optimization, InterCriteria analysis, Correlation, Weather analysis, Pollution analysis.
2000 Mathematics Subject Classification: 68Q85.

1 Introduction

The knowledge of weather is an aim since many centuries. It has an influence over many human activities and climate irregularities. It can provide information concerning the air flows and movements, temperature change and certain kind of phenomena such as tornadoes, blaze and hail. It influences the population of diseases and vermin infestation, but it also has an impact over the diffusion of gases, dust and other elements polluting air, water and environment. This means that by knowing the weather and pollution factors it is possible to predict or even restrain the pollution propagation which is useful for the early warning systems.

By applying an approach called InterCriteria Analysis (ICA) based on index matrices and fuzzy sets it is possible to obtain relations between many objects against many criteria, calculated for each criteria pair with values in the interval [0;1] over weather and pollution parameters it is possible to discover their correlation, the circumstances and phenomena due to the distribution of pollution factors. This approach is suitable in the process of decision making in data sets varying in size. The main advantage of ICA is that the input readings do not necessarily have to fulfill the requirement of equal number of criteria and objects. There is no limitation in the number of criteria and objects as well. Another advantage of ICA is the elimination of criteria that are expensive, hard to obtain or time consuming.

The application of operators over intuitionistic fuzzy sets have been presented in [1], where the geometrical representation of the results in terms of membership and non-membership in the intuitionistic triangle has been explored. An analysis of the graphical representation of the results in the intuitionistic triangle obtained by ICA has been proposed in [2]. A similar study on application of the InterCriteria Analysis over large weather datasets in order to obtain the correlation among diverse weather and weather related parameters along with an approach for data sets simplification has been implemented in [3]. The ICA application over large data sets of the Bulgarian universities rankings is proposed in [4].

In the present paper, an analysis of the correlation between weather and pollution factors by applying the ICA method is proposed. Thus influence of certain weather parameters over the propagation of small particle and gas emissions along with their dispersion could be predicted. Also the relation between pollution factors can be suggested, that could help to establish the pollution source due to the specific composition of gasses in every emission. It is also possible to locate the pollution source by knowing the parameters wind speed, humidity that relate with the pollution propagation. By applying the ICA approach it is possible to discover other interconnections between the parameters. In order to be able to provide an adequate estimation of the intercriteria correlation it is necessary to constantly monitor the parameters of interest. This requirement raises the need of the data storage and the computational power needed to apply the analysis. To reduce the computational requirements optimization approach that aims to reduce the amount of data input to ICA similar to [3] by estimating the initial correlation among the selected criteria and correlation among the objects is proposed. The higher the correlation is the more criteria/objects will be discarded. The aim is to remove as many criteria/object as possible without an impact over the relation functions.

The approach proposed in the present article extends the study performed in [3] by applying air pollution factors to the ICA along with weather parameters. It also aims to further reduce the data by ignoring object and / or criteria with great similarity.

2 InterCriteria Analysis Application over Weather and Atmosphere Parameters

The InterCriteria Analysis is a method that provides the correlation among each criterion in the data set that consists of objects and criteria. It is based on index matrices and intuitionistic fuzzy sets over the aforementioned data set. This method when applied over the current sampled data set provides the correlation between the weather parameters and the air pollution factors measured in the period of a week – from 1st of January to 7th of January. The Data is gathered from a station located in the city of Burgas. The data measurement period is 3 hours with a total of 8 samples per day. The presence of snowfall has to be accounted on 3rd of January, as well as rainfall on 5 – 6th of January.

The ICA approach is applied over 54 objects and the following criteria:

- 1 - minimal air temperature,
- 2 - minimal moisture,
- 3 - maximum air temperature,
- 4 - air temperature of dewing,
- 5 – temperature during maximum moisture
- 6 - moisture,
- 7 - wind direction in degrees,
- 8 - wind speed m/s,
- 9 - altitude in meters,
- 10 - clouds average height,
- 11 - visibility,
- 12 - amount of rain liters per km^2
- 13 Average amount of rain,
- 14 - atmospheric pressure,
- 15 - atmospheric pressure difference per 3h,
- 16 - Azote dioxide ,
- 17 - Benzene,
- 18 - Carbone oxide,
- 19 - Ozone,
- 20 - Sulfuric dioxide,
- 21 - Hydrogen sulfite,
- 22 - Styrene,
- 23 - Nano dust particles

Taking into account the sampled data the following parameters are expected to correlate: 6- 8, 12 - 23; 7 - 17-23; 13 – 18, 13 - 21, 22, 23, 25; 6, 17 – 23; 8, 17 -25. The results obtained by the analysis are illustrated in Table 1 a) and b).

μ	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	1,00	0,86	0,79	0,77	0,87	0,63	0,55	0,52	0,60	0,47	0,35	0,22	0,34	0,29	0,49	0,61	0,43	0,44	0,40	0,51	0,35	0,44	0,62
2	0,86	1,00	0,74	0,72	0,82	0,65	0,55	0,42	0,57	0,49	0,32	0,21	0,33	0,27	0,45	0,64	0,43	0,51	0,35	0,46	0,40	0,45	0,55
3	0,79	0,74	1,00	0,70	0,73	0,59	0,51	0,50	0,55	0,46	0,34	0,24	0,37	0,28	0,47	0,60	0,46	0,43	0,38	0,49	0,33	0,40	0,62
4	0,77	0,72	0,70	1,00	0,77	0,69	0,51	0,52	0,58	0,52	0,26	0,24	0,36	0,22	0,49	0,57	0,41	0,48	0,35	0,47	0,37	0,43	0,66
5	0,87	0,82	0,73	0,77	1,00	0,73	0,55	0,48	0,61	0,47	0,24	0,24	0,37	0,23	0,47	0,60	0,48	0,46	0,36	0,50	0,38	0,48	0,62
6	0,63	0,65	0,59	0,69	0,73	1,00	0,50	0,38	0,53	0,46	0,02	0,28	0,39	0,20	0,42	0,56	0,50	0,55	0,27	0,40	0,45	0,50	0,56
7	0,55	0,55	0,51	0,51	0,55	0,50	1,00	0,42	0,47	0,38	0,46	0,16	0,22	0,38	0,46	0,58	0,42	0,46	0,55	0,39	0,41	0,41	0,58
8	0,52	0,42	0,50	0,52	0,48	0,38	0,42	1,00	0,64	0,46	0,58	0,15	0,20	0,58	0,55	0,42	0,48	0,29	0,57	0,54	0,37	0,50	0,54
9	0,60	0,57	0,55	0,58	0,61	0,53	0,47	0,64	1,00	0,49	0,42	0,14	0,26	0,43	0,52	0,48	0,50	0,39	0,47	0,52	0,37	0,53	0,54
10	0,47	0,49	0,46	0,52	0,47	0,46	0,38	0,46	0,49	1,00	0,42	0,19	0,25	0,51	0,44	0,44	0,49	0,48	0,36	0,44	0,48	0,52	0,34
11	0,35	0,32	0,34	0,26	0,24	0,02	0,46	0,58	0,42	0,42	1,00	0,01	0,05	0,77	0,50	0,40	0,47	0,41	0,69	0,46	0,48	0,45	0,40
12	0,22	0,21	0,24	0,24	0,24	0,28	0,16	0,15	0,14	0,19	0,01	1,00	0,81	0,03	0,11	0,20	0,12	0,10	0,08	0,20	0,10	0,15	0,18
13	0,34	0,33	0,37	0,36	0,37	0,39	0,22	0,20	0,26	0,25	0,05	0,81	1,00	0,06	0,17	0,31	0,19	0,21	0,11	0,27	0,18	0,21	0,27
14	0,29	0,27	0,28	0,22	0,23	0,20	0,38	0,58	0,43	0,51	0,77	0,03	0,06	1,00	0,49	0,34	0,55	0,46	0,61	0,48	0,58	0,53	0,31
15	0,49	0,45	0,47	0,49	0,47	0,42	0,46	0,55	0,52	0,44	0,50	0,11	0,17	0,49	1,00	0,42	0,37	0,42	0,56	0,42	0,38	0,35	0,56
16	0,61	0,64	0,60	0,57	0,60	0,56	0,58	0,42	0,48	0,44	0,40	0,20	0,31	0,34	0,42	1,00	0,47	0,53	0,36	0,48	0,42	0,45	0,60
17	0,43	0,43	0,46	0,41	0,48	0,50	0,42	0,48	0,50	0,49	0,47	0,12	0,19	0,55	0,37	0,47	1,00	0,57	0,37	0,51	0,57	0,77	0,43
18	0,44	0,51	0,43	0,48	0,46	0,55	0,46	0,29	0,39	0,48	0,41	0,10	0,21	0,46	0,42	0,53	0,57	1,00	0,31	0,35	0,68	0,54	0,38
19	0,40	0,35	0,38	0,35	0,36	0,27	0,55	0,57	0,47	0,36	0,69	0,08	0,11	0,61	0,56	0,36	0,37	0,31	1,00	0,44	0,36	0,34	0,51
20	0,51	0,46	0,49	0,47	0,50	0,40	0,39	0,54	0,52	0,44	0,46	0,20	0,27	0,48	0,42	0,48	0,51	0,35	0,44	1,00	0,35	0,51	0,53
21	0,35	0,40	0,33	0,37	0,38	0,45	0,41	0,37	0,37	0,48	0,48	0,10	0,18	0,58	0,38	0,42	0,57	0,68	0,36	0,35	1,00	0,59	0,27
22	0,44	0,45	0,40	0,43	0,48	0,50	0,41	0,50	0,53	0,52	0,45	0,15	0,21	0,53	0,35	0,45	0,77	0,54	0,34	0,51	0,59	1,00	0,39
23	0,62	0,55	0,62	0,66	0,62	0,56	0,58	0,54	0,54	0,34	0,40	0,18	0,27	0,31	0,56	0,60	0,43	0,38	0,51	0,53	0,27	0,39	1,00

a)

ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	0,00	0,13	0,15	0,18	0,12	0,31	0,43	0,44	0,36	0,42	0,63	0,04	0,08	0,70	0,45	0,36	0,55	0,55	0,59	0,34	0,58	0,52	0,34
2	0,13	0,00	0,20	0,23	0,16	0,29	0,43	0,54	0,40	0,40	0,67	0,06	0,10	0,72	0,49	0,34	0,56	0,48	0,64	0,40	0,53	0,52	0,41
3	0,15	0,20	0,00	0,25	0,21	0,32	0,42	0,42	0,37	0,41	0,60	0,05	0,09	0,67	0,43	0,33	0,48	0,51	0,56	0,34	0,58	0,53	0,31

4	0,18	0,23	0,25	0,00	0,18	0,23	0,44	0,41	0,36	0,38	0,69	0,04	0,08	0,74	0,42	0,37	0,55	0,48	0,60	0,37	0,55	0,51	0,28
5	0,12	0,16	0,21	0,18	0,00	0,20	0,43	0,48	0,36	0,42	0,74	0,02	0,05	0,76	0,47	0,37	0,51	0,52	0,63	0,36	0,55	0,48	0,34
6	0,31	0,29	0,32	0,23	0,20	0,00	0,43	0,54	0,39	0,41	0,93	0,00	0,03	0,74	0,47	0,36	0,44	0,39	0,67	0,42	0,45	0,41	0,36
7	0,43	0,43	0,42	0,44	0,43	0,43	0,00	0,54	0,49	0,51	0,52	0,11	0,21	0,60	0,46	0,39	0,56	0,52	0,43	0,46	0,51	0,54	0,38
8	0,44	0,54	0,42	0,41	0,48	0,54	0,54	0,00	0,31	0,43	0,39	0,12	0,23	0,39	0,37	0,54	0,49	0,68	0,40	0,31	0,54	0,44	0,40
9	0,36	0,40	0,37	0,36	0,36	0,39	0,49	0,31	0,00	0,39	0,55	0,13	0,16	0,54	0,41	0,48	0,47	0,58	0,50	0,33	0,55	0,41	0,40
10	0,42	0,40	0,41	0,38	0,42	0,41	0,51	0,43	0,39	0,00	0,47	0,12	0,17	0,39	0,41	0,44	0,40	0,42	0,53	0,34	0,38	0,35	0,55
11	0,63	0,67	0,60	0,69	0,74	0,93	0,52	0,39	0,55	0,47	0,00	0,25	0,38	0,22	0,44	0,58	0,52	0,57	0,29	0,41	0,45	0,51	0,57
12	0,04	0,06	0,05	0,04	0,02	0,00	0,11	0,12	0,13	0,12	0,25	0,00	0,02	0,24	0,18	0,07	0,14	0,17	0,18	0,11	0,17	0,12	0,11
13	0,08	0,10	0,09	0,08	0,05	0,03	0,21	0,23	0,16	0,17	0,38	0,02	0,00	0,36	0,27	0,12	0,23	0,21	0,32	0,19	0,25	0,23	0,17
14	0,70	0,72	0,67	0,74	0,76	0,74	0,60	0,39	0,54	0,39	0,22	0,24	0,36	0,00	0,46	0,64	0,45	0,53	0,38	0,38	0,35	0,43	0,66
15	0,45	0,49	0,43	0,42	0,47	0,47	0,46	0,37	0,41	0,41	0,44	0,18	0,27	0,46	0,00	0,51	0,57	0,52	0,38	0,41	0,51	0,57	0,36
16	0,36	0,34	0,33	0,37	0,37	0,36	0,39	0,54	0,48	0,44	0,58	0,07	0,12	0,64	0,51	0,00	0,51	0,45	0,62	0,37	0,50	0,50	0,35
17	0,55	0,56	0,48	0,55	0,51	0,44	0,56	0,49	0,47	0,40	0,52	0,14	0,23	0,45	0,57	0,51	0,00	0,42	0,62	0,35	0,36	0,20	0,53
18	0,55	0,48	0,51	0,48	0,52	0,39	0,52	0,68	0,58	0,42	0,57	0,17	0,21	0,53	0,52	0,45	0,42	0,00	0,68	0,51	0,25	0,42	0,59
19	0,59	0,64	0,56	0,60	0,63	0,67	0,43	0,40	0,50	0,53	0,29	0,18	0,32	0,38	0,38	0,62	0,62	0,68	0,00	0,43	0,57	0,63	0,46
20	0,34	0,40	0,34	0,37	0,36	0,42	0,46	0,31	0,33	0,34	0,41	0,11	0,19	0,38	0,41	0,37	0,35	0,51	0,43	0,00	0,47	0,33	0,32
21	0,58	0,53	0,58	0,55	0,55	0,45	0,51	0,54	0,55	0,38	0,45	0,17	0,25	0,35	0,51	0,50	0,36	0,25	0,57	0,47	0,00	0,34	0,64
22	0,52	0,52	0,53	0,51	0,48	0,41	0,54	0,44	0,41	0,35	0,51	0,12	0,23	0,43	0,57	0,50	0,20	0,42	0,63	0,33	0,34	0,00	0,55
23	0,34	0,41	0,31	0,28	0,34	0,36	0,38	0,40	0,40	0,55	0,57	0,11	0,17	0,66	0,36	0,35	0,53	0,59	0,46	0,32	0,64	0,55	0,00

b)

Table 1: Degree of a) membership, and b) non-membership

As can be seen in Table 1, the strongest consonances are observed between criteria 1-2, 1-3, 1-4, 1-5, 2-1, 2-3, 2-4, 2-5, 11-14, 17-22 , which confirms the expectations. Another strong consonance can be noticed between criteria 6-11, 11-12, 11-13, 12-1:11, 12-13:23 13-1:11, 13-13:23.

The aforementioned consonances and dissonances can be observed in the intuitionistic triangle shown in Fig. 1.

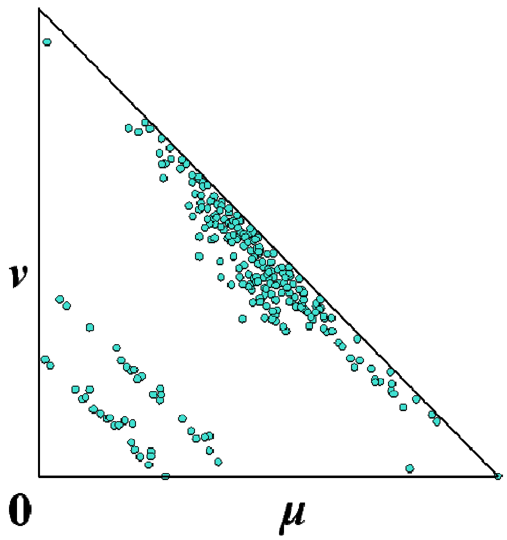

Figure 1: Intuitionistic triangle over the atmosphere data reading of interest

In Fig. 1 it can be observed that the level of uncertainty is reasonably high. This is due to the fact that some of the criteria do not correlate well with the other and have no significant meaning due to their slight change. The large number of criteria located in the uncertainty area in the intuitionistic triangle [1, 2] can be interpreted as presence of excessive data in the ICA. Its removal will not only improve the level of uncertainty, but will also increase the performance of the algorithm.

3 Performance optimization approach

As can be observed in Fig. 1 not all of the criteria are significant in terms of intercriteria relation. Furthermore not every strong consonance or dissonance is meaningful to the research. For example the strong relation between minimum, maximum temperature and the temperature of dewing is expected and every other criterion will be in the same degree of correlation with all of the temperature criteria. In order to improve the computational technique and thus the results it is necessary to reduce the meaningless data. The proposed optimization is a two-step approach:

- First step – estimation of constant criteria and objects. Slightly varying criteria and/or objects do not correlate good with the dynamically changing ones and increase the uncertainty level.
- Second step – estimate the correlation level between objects and/or criteria. Similar or equal data does not affect the ICA.

3.1 Estimation of constants in input data

In order to observe parameters that have different dimensions and magnitude it is required to perform normalization over each criterion. That way it is possible to equally visualize the criteria, as can be seen in Fig. 2.

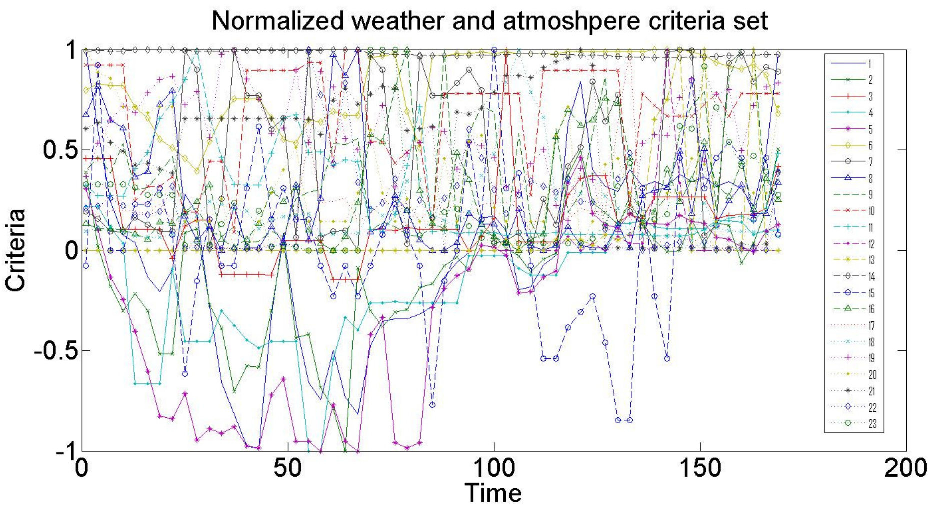

Figure 2: Normalized criteria set

To establish whether a criterion is changing slightly or not changing at all it is necessary to obtain the difference between minimal and maximal difference of each adjacent criterion. The lower the average value is – the less dynamically the criterion changes. A criterion that has an average magnitude difference value equal or less than 10% of the magnitude of the maximum value can be regarded as a constant and can be discarded.

After all the criteria has been calculated, the average minimum to maximum difference for criterion 1, 3, 9, 12, 13, 14 and 18 has a value less than the threshold of 10% and can be disregarded, as can be seen in Fig. 3.

Figure 4 shows that after the criteria proposed by this approach have been removed the uncertainty level is decreased and the following consonances between criteria 11 and 19 appear. It also impacts the existing relation by increasing the existing consonance level between the aforementioned criteria. The overall efficiency obtained by the proposed approach for the current data set is 30 percent reduction of data.

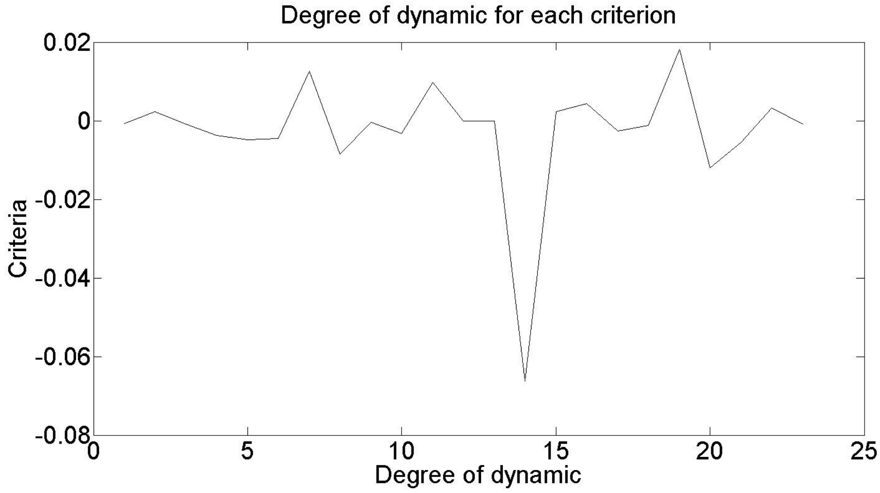

Figure 3: Average minimum to maximum criterion difference

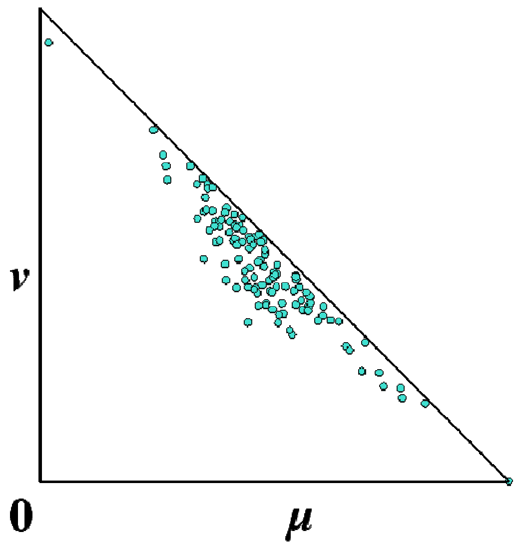

Figure 4: Intuitionistic triangle after ICA over the reduced data set

3.2 Estimation of correlation degree for objects/criteria

The proposed approach aims to discover similarities between separate criteria for every object and the one between every object for each criterion in the data set. If the correlation between certain objects/criteria is greater than 75% then the specified rows and/or pillars in the data set can be removed.

A correlation analysis among the criteria shows that the following criteria correlation 4-7, 7-4, 19-1, 21-4, 4-5, 4-19, 4-14, 5-4, 7-19, 7-21, 21-7 is over 75% as can be seen in Fig. 5 where all of the 240 correlations are illustrated. This means that criterion 4 correlates with most of the other criteria and can be removed. The same can be stated for criterion 7. The results of the analysis over the simplified data set can be observed in Fig. 6

The advantages of the proposed method are expressed by further reduce the input data and thus reduce the computation burden, improve the intercriteria correlation and decrease the uncertainty level.

The overall reduction is approximately 43.5 percent. Thus it is possible to choose between the highly correlating criteria the ones that are less expensive or the least time consuming without an impact over the decision making process. The objects removal can result improvement of the accuracy.

As can be noted by performing a comparison between Figs. 1, 4 and 6 the main consonances 2-4, 2-5, 4-5, 5-6, 17-22 and the dissonance 6-11 are still present in all of the three graphs. This means that the proposed approach reduces the uncertainty level as well as the amount of data.

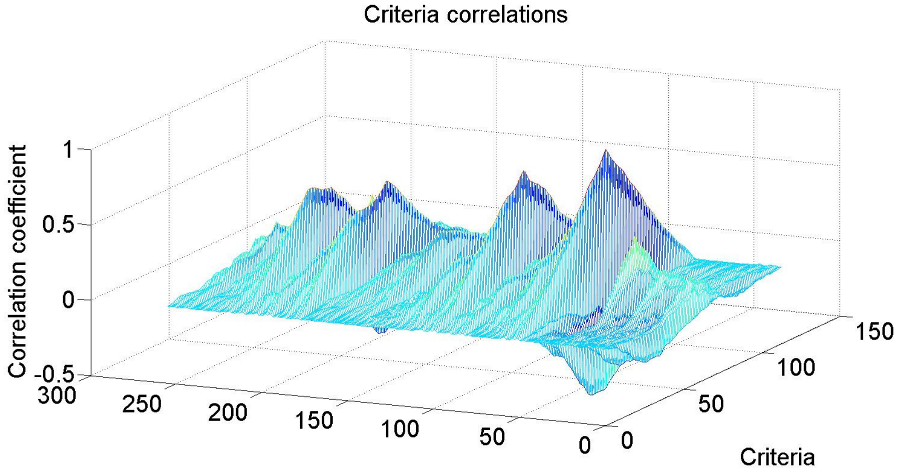

Figure 5: Correlation coefficient between criteria

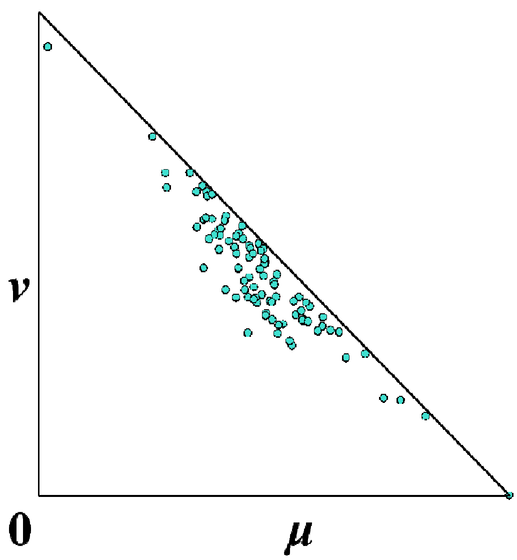

Figure 6: ICA over the simplified data analysis over the simplified data set can be observed

4 Conclusion

In this article an example of performance optimization analysis performed by the atmosphere and weather data from the city of Burgas has been implemented. Each step of the approach shows performance improvement without impact over the accuracy. The proposed optimization algorithms when combined provide up to 43.5 percent less data for analysis as well as reduction in the uncertainty level as can be seen in Fig. 1 before and after application in Figs. 4, 6. The results obtained for the current data are promising, but in order to prove the correctness of the algorithms further analyses over a diversity of type and size of the data, also including weather and atmosphere related data must be carried out. If the proposed algorithms are proven correct, this will lead to large benefits during the computation process of ICA.

Acknowledgements

The author is thankful for the support provided under Grant Ref. No. DFNI-I-02-4/2014 of the Bulgarian National Science Fund.

References

[1] Atanassov, K. A new topological operator over intuitionistic fuzzy sets. *Notes on Intuitionistic Fuzzy Sets*, Vol. 21, 2015, No. 3, 90–92.

[2] Atanassova, V. Interpretation in the Intuitionistic Fuzzy Triangle of the Results, Obtained by the InterCriteria Analysis. *IFSA-EUSFLAT 2015*, 30.06-03.07.2015, Gijon, Spain

[3] Kostadinov, T., T. Petkov. An example of intercriteria analysis application to weather parameters. *Notes on Intuitionistic Fuzzy Sets*, Vol. 21, 2015, No. 2, 126–133.

[4] Bureva, V., E. Sotirova, S. Sotirov, D. Mavrov. Application of the InterCriteria decision making method to Bulgarian universities ranking. *Notes on Intuitionistic Fuzzy Sets*, Vol. 21, 2015, No. 2, 111–117.

[5] http://www.stringmeteo.com/synop/bg_stday.php

[6] http://www.burgas.bg/bg/info/index/591