An Artificial Intelligence-Based Environment Quality Analysis System

Mihaela Oprea[1], Lazaros Iliadis[2]

[1] University Petroleum - Gas of Ploiesti, Department of Informatics,
Bd. Bucuresti nr. 39, Ploiesti, 100680, Romania
mihaela@upg-ploiesti.ro
[2] Democritus University of Thrace, Greece
liliadis@fmenr.duth.gr

Abstract. The paper describes an environment quality analysis system based on a combination of some artificial intelligence techniques, artificial neural networks and rule-based expert systems. Two case studies of the system use are discussed: air pollution analysis and flood forecasting with their impact on the environment and on the population health. The system can be used by an environmental decision support system in order to manage various environmental critical situations (such as floods and environmental pollution), and to inform the population about the state of the environment quality.

1 Introduction

The continuous improvement of life quality is related directly, among others, to the improvement of environmental quality. This can be achieved by using adequate intelligent monitoring, analysis, forecasting, decision and control systems, based on intelligent tools and techniques. Such integrated intelligent systems are very useful decision support tools for the management of environmental critical situations (events or incidents): floods, severe air/water/soil pollution, earthquakes, tsunami, storms, land sliding, avalanches and others. Also, they can be used in everyday life in order to inform the population of a certain region: urban region (city/town), metropolitan area, or wider areas such as a whole county, continental or intercontinental areas. The complexity of such systems is high due to the interdependencies of various environment parameters. For example, the meteorological factor of precipitations has a direct impact to the evolution of the flow in a specific hydrographic basin as well as to the degree of environmental pollution (air, water or soil). If the volume of precipitation is very high it can contribute to a flood in a certain region of a river and also it can lead to soil pollution, if the water of the river has been polluted.

   An environmental decision support system (DSS) capable of handling environmental crises requires an efficient forecasting and analysis approach that can guide the development of a proper strategy. The implementation of such a system can be obtained by employing artificial intelligence (AI) methods that might be combined in an optimal and hybrid manner. AI approaches that can be employed include rule-based expert systems, machine learning techniques like artificial neural networks

(ANN) and inductive learning, data mining algorithms, case based reasoning and optimization methods like genetic algorithms.

As the majority of the environmental processes have a certain degree of uncertainty, and a variety of interdependencies between environmental factors and as we do not always have complete datasets (with recorded timeseries or real time accurate measurements) the environmental forecasting and analysis systems have to deal with uncertain knowledge.

A hybrid method is proposed in this paper that combines two AI techniques: a feed forward artificial neural network (doing the forecasting) and a rule-based expert system (doing the analysis). The INTELLEnvQ system (**INTELL**igent **Env**ironment **Q**uality) was developed as a tool for environmental quality analysis, that provides information to environmental decision factors and to the population, about the state of the environment of a specific region in a certain period of time, with the possible consequences on human health and on the whole environment (air, soil, water, vegetation, fauna). The system provides also alert codes associated to the specific environmental event/incident.

2   Environmental Forecasting and Analysis

The quality of life is affected by various factors and critical environmental situations. It is important to have forecasting intelligent systems capable of alerting the population in order to prevent or to reduce the effects of an extreme environmental incident.

Figure 1 shows the possible effects of some environmental critical situations to other parts of the environment. For example, a severe air pollution problem would affect the soil, and the water. Also, if two problems are combined, the impact on the environment could increase significantly. For example, if there is a severe air pollution and flood in a given region, than the impact on the water and soil is greater, increasing the pollution degree.
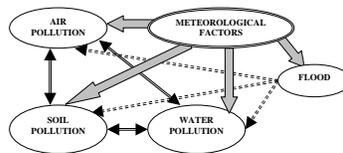
Fig. 1. Interdependencies between different environmental critical situations

Figure 2 presents an overview of the dependencies between the meteorological system and the environment (air, water, soil, vegetation, humans). The system is linked to a network of meteorological stations (MSN) which provides measurements and observations from specific locations. Also, the environment can be associated to specific networks of air, water or soil quality monitoring stations capable of providing measurements and observations of various environmental parameters (e.g. concentration of air pollutants). For example, the air stations network (ASN) can provide measurements of air pollutants in urban regions (SO_2, CO, NO_2, NH_3, O_3,

PM_{10}, $PM_{2.5}$) whereas the water stations network (WSN) can measure the level of water pollution in a hydrographic basin or hydrometric station and soil stations network (SSN) can offer information on soil pollution in the agricultural (farming) regions. Also, WSN can include hydrometric stations networks which monitor and analyze various hydrologic parameters in specific hydrographic basins in order to allow management of torrential phenomena. Some events or incidents that have to be predicted and managed (through alerting systems) are also specified in Figure 2. For example, rainfalls and storms in case of the meteorological system, air pollution in case of air monitoring system, soil pollution and land sliding in case of the soil monitoring system, and flood and tsunami in case of the water monitoring system, just to name some of the most important.
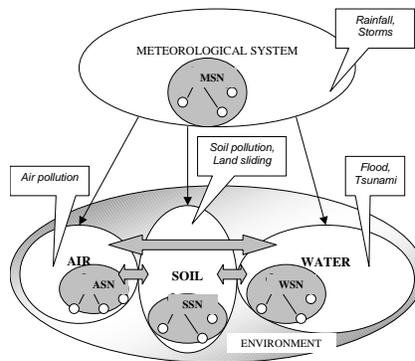
Fig. 2. Overview of the dependencies between the meteorological system and the environment (air, soil, water)

3   Artificial Intelligence and Environment Quality Analysis

3.1   Literature Review

The potential benefits of using various artificial intelligence techniques and in particular, machine learning (ML) techniques for solving different environmental problems were emphasized by various research efforts that have been reported in the literature [1-7] and by new research opportunities such as those mentioned in [8]. Several environmental problems (monitoring, analysis, diagnosis, forecasting, planning, control) can be solved efficiently by using AI techniques. In this paper we shall focus on the analysis problem of two particular cases: air pollution and flood forecasting. The solution that was adopted is the use of a rule-based system that receives the values of the forecasted parameters by an ANN.

The main AI techniques that can be applied to forecasting problems solving are ANN, fuzzy logic, case-based reasoning, Fuzzy Inference Systems (ANFIS), expert

systems, adaptive neuro-fuzzy systems. Other hybrid techniques that can be used are those based on Genetic Algorithms (GA), Hidden Markov Models (HMM), Support Vector Machines (SVM).

Air Pollution Forecasting. In the last years, several AI-based forecasting systems were presented in the literature for the prediction of different air pollutants. In [9], the authors describe a novel feature scaling method, named neighbor-based feature scaling that is combined with an ANN model and an ANFIS model for the prediction of sulfur dioxide (SO_2) concentration. The performance of the proposed method was evaluated by using four statistical indicators, including the index of agreement (IA). The results were very good showing the benefits of using the proposed method for sulfure dioxide concentration forecasting in a specific region of Turkey (the Konyia province).

In [4] it is presented an overview of several soft computing approaches like ANN, SVM, fuzzy logic as a feasible solution for urban air forecasting. Another work that reported a successful use of ANN and ANFIS models for the prediction of daily carbon monoxide (CO) concentration in the atmosphere of Tehran is described in [10]. The performances of the models were analyzed by taken into account the uncertainty in forecasting (a Monte-Carlo simulation was used). In [11] it is presented knowledge modelling in an air pollution control DSS.

There are several research projects that deal with air pollution forecasting in various regions worldwide, on all the continents. Some examples of such projects and research initiatives are reported in [12-15] (SATURN for urban air monitoring), [16-17]. From this brief review, it can be concluded that the use of ANN and ANFIS models provides one of the best forecasting results.

Flood Forecasting. Flood forecasting became (during the last 20 years) an intensive research topic for most of the countries. Most of the recent flood forecasting systems that were reported in the literature are based on AI. In [18] it is presented an analysis of using ANFIS models in forecasting river floods. The results showed that such models are accurately and reliable. The models forecast the discharge at a downstream station by using the flow information at various upstream stations. The research work described in [19] focuses on three models of flood forecasting used in India, ANN, ANFIS and ANGIS (Adaptive Neuro-GA Integrated System). The conclusion of the research was that ANFIS performs better than ANN, whereas ANGIS provides maximum accuracy in predicting flood events. The ANGIS model uses GA to optimize the ANN weights. In [20] it is presented a hybrid solution, based on ANN and kernel algorithms for the water resources management in the Thasos islands. In this case, SVM were used to optimize the ANN. The most significant result of this research is that the optimal ANN was validated by using a kernel approach. In [21] it is described an expert system prototype for flood forecasting and river monitoring. The knowledge base of the system consists in expert rules that take into account the specific hydrographic basin parameters, and some meteorological factors. The system was experimented as a simulation for the Prahova hydrographic basin from Romania.

Finally, we have to mention that there are several research projects and initiatives to develop flood monitoring, forecasting and management systems ([12], [13], [16], [22]).

From the brief review presented in this section we can conclude that the most succesfully AI techniques that were used both in air pollution forecast and flood forecasting are ANN and ANFIS. In this paper we shall focus our discussion on the rule-based system considering that the ANN provides the forecasting values for the parameters that are analyzed by the rule-based system in a specific environmental event/incident scenario.

4 The INTELLEnvQ System

The architecture of the INTELLEnvQ system is presented in Figure 3. The main components of the system are a knowledge base under the form of rule base, an inference engine, both components being parts of a rule-based expert system and a forecasting module that is composed by a feed forward artificial neural network. The rules from the knowledge base provide the environmental quality analysis result by taking into account the forecasting result provided by the ANN and other factors (such as the meteorological ones) that influence the environment quality.
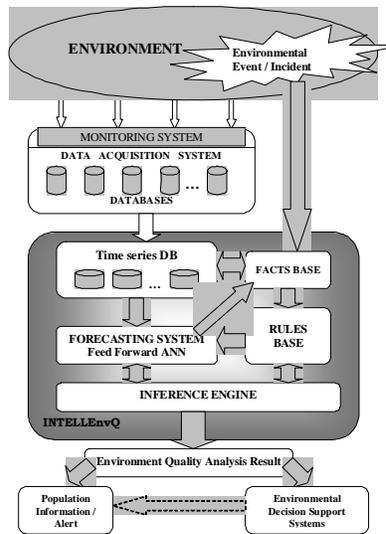

Fig. 3. The architecture of the INTELLEnvQ system

The information about a certain environmental incident is described by the facts from the facts base. The forecasting system uses the meteorological timeseries and the timeseries with the values of the parameters specific to the environmental incident (real measurements, taken from publicly available databases), and provides the values of the forecasted parameters that are taken by the facts base of the rule-based system, and that are used during the reasoning process of the environmental quality analysis. Therefore, the integration of ANN with the rule-based system is facilitated by the

facts base, which apart from the current context of a certain environmental incident will contain the values of specific forecasted parameters that will fire certain rules during the reasoning process. The INTELLEnvQ system provides the environmental quality analysis results that might be announced to the population of a region affected by an environmental incident and to environmental DSS. Also, the system can be used as a daily information tool about the state of the environment (air, water, soil) and the potential effects on the population's health in regions with environmental problems, such as urban regions where air pollution can affect seriously the health of people.

The data acquisition system is included in an environmental monitoring system that records the environmental data (meteorological, air, water, soil) in databases. As most of the EU countries have implemented the air pollution monitoring system and some of them have implemented also the hydrometric stations network, we have developed the knowledge base of the INTELLEnvQ system for air quality analysis and flood forecasting. In the next section we focus on these two case studies.

The rules from the knowledge base include standard checking rules, that make comparisons with predefined limits of pollutant concentration (i.e. admissible values in a certain interval of time), dependences analysis rules, that make an extensive analysis by taking into account other parameters that can influence the analyzed environment quality parameter (e.g. meteorological parameters), alert coding rules, and impact analysis rules (e.g. the impact on the environment or on the population health). The standard checking rules were established by taking into account real measurements of the environmental parameters and the environmental quality standards with pre-defined limits of pollutants concentrations (e.g. the maximum admissible concentration in a certain period of time). The dependences analysis rules were established by taking into account the timeseries with real measurements of the environmental parameters (their evolution in time) and the experts' knowledge (about the interdependences between different environmental parameters derived from their experience) that provides some heuristic rules. The alert coding rules were established by taking into account the environmental norms in critical situations. The impact analysis rules were derived from experts' knowledge. The dependences analysis rules and the impact analysis rules depend also on the characteristics of the region for which it is made the environment quality analysis. The problem of combinatorial explosion was phased by developing decision tables in a rather heuristic manner. This minimized the number of rules to a rational extend. Some examples of rules from the knowledge base of INTELLEnvQ system are given in Figure 4 (rules for air quality analysis), and in Figure 5 (rules for flood analysis).

```
RULE 62
IF AP_PM10 = greater_increase AND AP_SO2 = greater_increase
   AND AP_NO2 = greater_increase
THEN AQ = severe_air_pollution;

RULE 102
IF AQ = possible_air_pollution THEN  Alert_Code = yellow;

RULE 103
IF AQ = severe_air_pollution THEN  Alert_Code = red;

RULE 201
IF Alert_Code = yellow THEN
  Impact = possible_impact
  DISPLAY "Attention - keep monitoring and analyzing very carefully";

RULE 202
IF Alert_Code = red THEN
  Impact = real_impact
  DISPLAY "ALERT - persons with asthma and other respiratory diseases";
```

Fig. 4. Rules for air quality analysis from the knowledge base of the INTELLEnvQ system

```
RULE R22
IF Temp = high AND PrecipitationS = high
   AND Type_precip = snow   AND Season = winter
THEN Flood_Warning = Yes;

RULE R23
IF Temp = normal AND PrecipitationS = high
   AND Type_precip = rains   AND Season = summer
THEN Flood_Warning = Yes;

RULE T27
IF Flood_Warning = Yes AND Flow >= (Attention_TH)
   AND Flow < (Flood_Danger_TH)
THEN Flood_Forecast = possible_flood;

RULE T28
IF Flood_Warning = Yes AND Flow >= (Flood_Danger_TH)
   AND Flow < (Real_Flood_TH)
THEN Flood_Forecast = very_possible_flood;

RULE A1
IF Flood_Forecast = no_flood THEN   Alert_Code = green;

RULE A2
IF Flood_Forecast = possible_flood THEN   Alert_Code = yellow;

RULE A3
IF Flood_Forecast = very_possible_flood THEN  Alert_Code = orange;

RULE A4
IF Flood_Forecast = severe_flood THEN   Alert_Code = red;
```

Fig. 5. Rules for flood forecasting from the knowledge base of the INTELLEnvQ system

5   Case Studies

For the first version of the INTELLEnvQ system we have used the Neural Network Toolbox from Matlab 7.7.0 (2008b) for the development of the forecasting system (i.e. the ANN), and we have implemented the rule-based expert system in VP-Expert, an expert system generator. We have tested the system on two case studies: flood forecasting and air quality analysis in urban regions from Romania. We have used the timeseries with public available data from the National Meteorology Agency (http://www.meteoromania.ro), the Romanian Statistical Institute (http://www.inss.ro), and the National Air Quality Monitoring Network (http://www.calitateaer.ro). The ANN employed in both application cases were back propagation multi layer feed forward using the Extended Delta Bar Delta algorithm and one hidden sub layer. The number of hidden layers was kept as minimum as possible in order to avoid complexity.

5.1   Flood Forecasting

The first case study is a flood forecasting problem. We have extracted the timeseries from the publicly available databases with precipitations and rivers flows data for some Romanian hydrographic basins.

   Each hydrographic basin has a number of hydrometric stations that provide measurements and observations for precipitations, rivers flows and other specific parameters (e.g. the three flow thresholds: attention, danger, flood). We have run several scenarios for each hydrographic basin. Figure 6 presents a screenshot of the INTELLEnvQ system run in the case of a scenario with real flood and possible pollution of the soil, while Figure 7 shows a screenshot of a system run in the case of a scenario with a possible flood with no additional impact on the environment pollution.

```
██▼                                                               _ □ ✕
                      ─[ KBS: FLOOD_FO ]─
   summer                      winter ◄

Hydrometric station: SV-1
The result of the analysis is severe flood
  FLOOD ALERT CODE >>>>  red
How is the water pollution?
  higher_pollution ◄        no_pollution

Analysis of impact on the environment: impact on soil pollution

         ─[ RULES ]─                    ─[ FACTS ]─
Testing I3                      Flood_Warning = Yes CNF 100
RULE I3 IF                      Flood_Forecast = severe_flood CNF 100
Alert_code = red AND
Water_pollution = higher_pollution   Alert_Code = red CNF 100
THEN                            Water_pollution = higher_pollution CN
Impact = impact_on_soil_pollution CNF 1  F 100
00                              Impact = impact_on_soil_pollu CNF 100
Finding Water_pollution
─────────────────────────────────────────────────────────
1Help     2        3WhatIf   4Variable 5Rule    6Set     7Edit    8Quit
1Help 2How? 3Why? 4Slow 5Fast 6Quit
```

Fig. 6. Screenshot of the INTELLEnvQ system run (flood forecasting - alert code - red)


```
██▼                                                               _ □ ✕
                      ─[ KBS: FLOOD_FO ]─
   summer ◄                    winter

Hydrometric station: SV-1
The result of the analysis is possible flood
  FLOOD ALERT CODE >>>>  yellow
How is the air pollution?
  severe_pollution         small_pollution ◄        no_pollution

Analysis of impact on the environment: insignificant impact

         ─[ RULES ]─                    ─[ FACTS ]─
Testing I11                     Flood_Warning = Yes CNF 100
RULE I11 IF                     Flood_Forecast = possible_flood CNF 1
Alert_code = yellow AND         00
Precipitations = high AND       Alert_Code = yellow CNF 100
Air_pollution = small_pollution   Air_pollution = small_pollution CNF 1
THEN                            00
Impact = insignificant_impact CNF 100  Impact = insignificant_impact CNF 100
Finding Air_pollution
─────────────────────────────────────────────────────────
1Help     2        3WhatIf   4Variable 5Rule    6Set     7Edit    8Quit
1Help 2How? 3Why? 4Slow 5Fast 6Quit
```

Fig. 7. Screenshot of the INTELLEnvQ system run (flood forecasting - alert code - yellow)

5.2   Air Quality Analysis

The second case study is an air quality analysis in urban regions. We have considered three air pollutants: PM10, SO_2 and NO_2. The possible impact of these pollutants to human health was also analyzed. Figure 8 presents a screenshot of the INTELLEnvQ system run in the case of a scenario with serious air pollution with great impact on the population health, while Figure 9 shows a screenshot of a system run in the case of a scenario with no air pollution and no impact on the population health.


```
██▼                                                               _ □ ✕
                      ─[ KBS: AIR_PM10 ]─
   other_pollutant

What is the concentration of PM10 [ug/mc]?
  29.49

  Air quality analysis result: increased air pollution
  Alert code >>>> red
ALERT - persons with asthma and other respiratory diseases
  The impact on population health: real impact

         ─[ RULES ]─                    ─[ FACTS ]─
Alert_Code = green              Station = PH-1 CNF 100
THEN                            Air_pollutant = PM10 CNF 100
Impact = no_impact CNF 100      AP_Conc_PM10 = 29.49 CNF 100
Testing IT3                     AP_PM10 = increase CNF 100
RULE IT3 IF                     AQ = increased_air_pollut CNF 100
Alert_Code = red                Alert_Code = red CNF 100
THEN                            Impact = real_impact CNF 100
Impact = real_impact CNF 100
─────────────────────────────────────────────────────────
1Help     2        3WhatIf   4Variable 5Rule    6Set     7Edit    8Quit
1Help 2How? 3Why? 4Slow 5Fast 6Quit
```

Fig. 8. Screenshot of the INTELLEnvQ system run (air pollution - alert code - red)

Fig. 9. Screenshot of the INTELLEnvQ system run (air pollution - alert code - green)

6   Conclusion

The paper presented an AI-based environment quality analysis system, INTELLEnvQ, that can be mapped to different scenarios of environmental events or incidents with the adaptation of its knowledge base accordingly, by taken into account the specific parameters and environmental context that need to be instantiated. The system uses an artificial neural network for short term forecasting and a rule-based expert system that makes the analysis or forecasting, depending on the type of environmental problem that need to be solved. The main purpose of the system is to give a quick answer to the population or to the environmental decision factors in order to reduce the impacts on the environment and on the population. In this paper we have presented two case studies of the system use: a flood forecasting problem and an air quality analysis problem. Such a system can be integrated in an environmental decision support system.

References

1. Dunea, D., Oprea, M.: Fuzzy-APA: Employing Fuzzy and Neural Networks Techniques in Data Analysis of Industrial Wastewaters Monitoring. WSEAS Transactions on Environment and Development, 6(8), 581--590 (2010)
2. Fdez-Riverola, F., Corchado, J.M.: Improved CBR System for Biological Forecasting. Proceedings of the ECAI WS – Binding Environmental Sciences and Artificial Intelligence, Valencia, (2004)
3. Karatzas, K.D.: Artificial Intelligence Applications in the Atmospheric Environment: Status and Future Trends. Environmental Engineering Management Journal, 9(2), pp. 171--180 (2010)
4. Li, M., Hassan, M.R.: Urban air pollution forecasting using artificial intelligence-based tools. Chapter 9 (2009)
5. Núñez, H., Sànchez-Marrè, M., Cortés, U., Comas, J., Martinez, M., Rodríquez-Roda, I., Poch, M.: A comparative study on the use of similarity measures in case-based reasoning to

improve the classsification of environmental system situations. Environmental Modelling & Software, 9(9), pp. 809--819 (2004)

6. Sànchez-Marrè, M., Gibert, K., Sevilla B.: Evolving GESCONDA to an Intelligent Decision Support Tool. Proceedings of the International Congress on Environmental Modelling and Software Modelling for Environment's Sake, Fifth Biennial Meeting, Ottawa, Canada (2010)

7. Tsataltzinos, T., Iliadis, L., Spartalis, S.: An intelligent Fuzzy Inference System for Risk Estimation Using Matlab Platform: the Case of Forest Fires in Greece. In: IFIP International Federation for Information Processing, Vol. 296, Artificial Intelligence Applications and Innovations III, pp. 303--310 (2009)

8. Dietterich, T.: Machine Learning in Ecosystem Informatics and Sustainability. Proceedings of IJCAI 2009, pp. 8--13 (2009)

9. Polat, K.: A novel data preprocessing method to estimate the air pollution (SO_2): neighbor-based feature scaling (NBFS). Neural Computing & Applications, Springer (2011)

10. Noori, R., Hoshyaripour, G., Ashrafi, K., Araabi, B.N.: Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. Atmospheric Environment, 44(4), pp. 476--482 (2010)

11. Oprea, M.: A case study of knowledge modelling in an air pollution control decision support system. AiCommunications, IOS Press, 18(4), pp. 293--303 (2005)

12. Ebel, A., Davitashvili, T., (Eds): Air, Water and Soil Quality Modelling for Risk and Impact Assessment, Springer (2007)

13. Kim, Y.J., Platt, U. (Eds): Advanced Environmental Monitoring. Springer, (2008)

14. Kolehmainen, M., Martikainen, H., Ruuskanen, J.: Neural networks and periodic components used in air quality forecasting. Atmospheric Environment, 35, pp. 815--825 (2001)

15. Moussiopoulos, N. (Ed): Air Quality in Cities. Springer, Berlin (2003)

16. Von Storch, H., Tol, R.S.J., Flöser (Eds): Environmental Crises – Science and Policy. Springer (2008)

17. Xu, H, Xue, H.: Improved ANFIS to Forecast Atmospheric Pollution. Management Science and Engineering, 2(2), pp. 54--61 (2008)

18. Ullah, N., Choudhury, P.: Flood Forecasting in River System Using ANFIS. AIP Proceedings of the International Conference on Modeling, Optimization, and Computing, Vol. 1298, pp. 694--699 (2010)

19. Mukerji, A., Chatterjee, C., Raghuwanshi, N.S.: Flood Forecasting Using ANN, Neuro-Fuzzy, and Neuro-GA Models. Journal of Hydrologic Engineering, 14(6) (2009)

20. Iliadis, L.S., Spartalis, S.I., Tachos, S.: Kernel Methods and Neural Networks for Water Resources Management. Environmental Engineering Management Journal, 9(2), pp. 181--187 (2010)

21. Matei, A.: Flood forecasting and river monitoring expert system in the Prahova river basin. Proceedings of the IFAC Workshop on Intelligent Control Systems, Vol. 8, Part 1, pp. 82--86 (2010)

22. The European Flood Alert System (EFAS) EFAS-IS Portal, http://efas-is.jrc.ec.europa.eu