DS Journal of Artificial Intelligence and Robotics
Volume 3 Issue 1, 1-8, Jan - Mar 2025
ISSN: 2583-9926 / https://doi.org/10.59232/AIR-V3I1P101

Original Article

Rules-Based Expert System for Determining Groundwater Quality in Tropical Areas in Africa

Humberto Cuteso Matumueni

Institute Superior Polytechnic of Soyo, Department of Computer Science Soyo, Angola.

hkuteso@hotmail.com

Received: 05 January 2025; Revised: 01 February 2025; Accepted: 03 March 2025; Published: 28 March 2025

Abstract -This article proposes a rule-based expert system to determine groundwater quality in tropical Africa, where communities rely heavily on groundwater for drinking, agriculture, and domestic use. In tropical regions, the consumption and use of unsafe water is common; it is a cause of disease, and communities face water-related health problems. The criteria and methods for determining water assessment parameters are based on international standards established by the World Health Organization, with particular attention paid to Class A parameters corresponding to drinking water quality. After compiling the acquired knowledge or parameters, a decision grid was developed to structure and simplify repetitive knowledge. Based on this knowledge, rules were created to provide a knowledge base for determining water quality. Implementing an expert system using Corvid Exys software to determine water quality in water quality decision-making in regions facing water-related health problems. A comparative test was conducted with Corvid Exsys software, and based on the results obtained, the experts validated the developed expert system. Finally, the results were evaluated using Cohen's Kappa coefficient to measure the reliability of the expert system. The results of this comparative study highlighted the usefulness of a rule-based expert system in determining groundwater quality and thus reducing the risk of emerging diseases in tropical areas of Africa. The results indicated good agreement with the experts' manual assessments and revealed areas at high risk of contamination.

Keywords - Artificial Intelligence, Expert system, Rule-based system, Groundwater, Exsys Corvid.

1. Introduction

Today, access to quality drinking water is a major public health issue worldwide, particularly in tropical Africa, where people rely heavily on groundwater. More often than not, this precious liquid is exposed to contamination from both anthropogenic and natural sources, which promotes the emergence of waterborne diseases such as diarrhoea, cholera, and typhoid fever. Particularly, the lack of conditions makes tropical areas vulnerable to these risks due to climatic conditions conducive to the proliferation of pathogens in water resources. Human survival is seriously compromised without access to safe drinking water, especially in vulnerable regions such as tropical Africa. According to the World Health Organization, consuming contaminated water has serious consequences for human health, and nearly two billion people worldwide use contaminated water, which significantly increases the risk of infectious diseases [1].The World Health Organization states that infectious diseases cause millions of deaths each year, particularly in countries with limited health resources. Water is essential to human life, and safe drinking water is a basic human right and a vital need [2]. Nearly two billion people worldwide drink contaminated water without treatment, causing infectious diseases, and this situation is more often exacerbated in rural and remote areas where sanitation infrastructure is inadequate. In this context, a rules-based expert system for testing groundwater quality could be crucial in preventing emerging diseases. Despite scientific advances in water quality monitoring, several gaps remain. Methods for determining and analyzing water quality do not meet field needs, and much remains to be done. Among the existing systems, in China, Huang's fuzzy sets, an expert system based on fuzzy logic for water quality classification; in Spain, Rodriguez's Bayesian approaches, who developed a Bayesian model for the detection of chemical contaminants in groundwater, offer advanced modelling, but are limited by their complexity and the need for large amounts of data; and finally, Patil's IoT and artificial intelligence-based systems, which require expensive infrastructure and constant connectivity.The ExpertAgua Expert system offers a difference, such as simplicity, technological independence, and adaptation to local conditions, thus ensuring greater accessibility and more efficient implementation in tropical African areas.

2. Literature Review

2.1. Expert Systems

Expert systems are a subset of artificial intelligence and first appeared in the late 1960s [3]. The general idea is to store human expertise in a computer, allow users to access it, and provide advice [4].

2.2. Rule-Based Expert System

A rule-based expert system is an approach that uses the expertise and knowledge of experts to solve problems that require human intelligence. To solve these problems, it uses these two elements as rules to explain the reasoning to the computer. The rules take the form of If-Then to form the phases or conditional statements that form elements of the knowledge base [5].

2.3. Decision Table

A decision table is a logical tool for easily representing an exhaustive set of conditional and mutually exclusive expressions. Instead of obtaining a series of conditions nested by a succession of IF..., THEN..., ELSE..., creating a table containing them is possible. It simplifies the decision-making process by associating specific quality conditions with the results [6].

2.4. Knowledge base from Decision Tables

Decision tables offer simple validation and verification or a mechanism for quickly performing validation and verification. The approach is also conditionally oriented, except that we are no longer interested in a single answer but in an overview of possible combinations and conclusions, which allows us to check for completeness, correctness, and consistency [7].

2.5. Introduction to Expert Systems for Water Quality Assessment

Humans cannot be replaced, especially in terms of creativity and the use of general knowledge in a well-defined domain. Expert systems are artificial intelligence applications that mimic human reasoning in specialized fields. Expert systems have one advantage over humans: human knowledge eventually loses its effectiveness, especially if it is rarely used. Expert systems memorize human knowledge and allow users to answer specific or hypothetical questions, which can lead to obtaining accurate and relevant information [8].The composition of groundwater is determined by its interactions within the hydrological cycle, and these interactions can lead to the chemical activity of undesirable constituents in groundwater, thereby affecting water quality.When water does not meet relevant quality criteria according to international standards, it is considered unsafe to drink and poses a threat to human health and ecological integrity [9].Water quality assessment is generally based on chemical and microbiological analyses. However, compliance with drinking water quality guidelines remains a recommendation of the World Health Organization for risk management of hazards that may compromise the safety of drinking water. These recommendations should be considered in the context of risk management of other sources of exposure to these hazards, such as waste, air, food and consumer products [1].

2.6. Water Quality

Water quality is essential for disease prevention. The presence of chemical, biological, and physical parameters can make water unsafe for human consumption [2]. Assessing these parameters is essential to ensure water safety. A combination of physicochemical and microbiological analyses determines water quality. These tests measure contaminant levels and classify water according to its suitability for human consumption [10].

2.7. Parameters of Each Water Class According to the WHO

Various parameters used to classify water include turbidity, fecal coliform concentration, pH, heavy metal levels, and the concentration of elements such as iron and arsenic. For each class the World Health Organization has defined maximum permissible limits for each parameter, allowing water to be classified into a class. The World Health Organization classifies water into different classes based on its microbial quality to determine and classify it. These categories are generally A, B, and C, depending on the degree of contamination. Each class has its own set of rules, which determine the need for different levels of treatment (light treatment, full treatment, or no treatment). If water quality indicators are significantly outside safe limits, classify it as Class C, requiring comprehensive treatment methods to ensure water quality.

Table 1. Summary of rules for each class

Class	Microbiological	Chemical	Physical	Treatment Requirement
Class A	No coliforms, no E.coli, No Enterococci	Nitrates < 50 mg/L, Arsenic < 0.01 mg/L, Lead < 0.01 mg/L, Fluoride < 1.5 mg/L	pH 6.5-8.5, turbidity < 5 NTU	No treatment required
Class B	1- 10 Total Coliforms, no E.coli, 0 -10 Enterococci	Nitrates < 50 mg/L, Arsenic < 0.01 mg/L, Lead < 0.01 mg/L, Fluoride < 1.5 mg/L	pH 6.5-8.5, Turbidity 5 - 10 NTU	Light treatment (disinfection, filtration)
Class C	> 10 Total Coliforms or any E. coli, > 10 Enterococci	Nitrates > 50 mg/L, Arsenic > 0.01 mg/L, Lead > 0.01 mg/L, Fluoride > 1.5 mg/L	pH < 6.5 or > 8.5, Turbidity > 10 NTU	Full treatment (advanced filtration, chemical treatment)

The classification and regulatory system assesses water in different areas based on its level of contamination. By classifying groundwater A, B, or C, the expert system can determine water quality and the appropriate level of treatment, thereby reducing the risk of waterborne diseases and ensuring access to safe drinking water.

3. Methodology

The research methodology adopted in this work was defined as bibliographic and experimental. The bibliographic research is based on previously published material, consisting mainly of books and journal articles that are currently available on the Internet. The bibliographic research was based on consultations of bibliographic sources and theoretical references: articles, books, theses, dissertations, websites with content on groundwater, Artificial Intelligence and Expert Systems.

3.1. Testing Tools and Platforms

The applications used in this work were:

The Exsys Corvid tool was used to implement the Expert System. Office suite applications, including Excel, are also used for data collection and validation and access for creating and standardizing the database. From the analysis of these records, the variables necessary for the elaboration and approval of the rules of the Expert System were extracted, and Exsys Corvid was used for this purpose. The objective of Expert Sinta is to simplify the steps for creating a complete expert system as much as possible. The experimental methodology followed the following phases: knowledge acquisition and implementation of the expert system. The experiments were carried out in the following order:

First, knowledge acquisition.

In this work, the data acquisition task was performed using spreadsheets and prior knowledge of the institution's experts and followed the following steps:

(a) Pre-processing data collection using the Excel application based on the expertise of water quality technicians or water quality analysts through direct interviews and literature;
(b) Evaluation and validation of the data collected with Excel by water quality technicians or water quality analysts;

Table 2. Decision table to monitor the rules obtained by the Laboratory technicians

	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14
No Coliforms	Y													
Total Coliforms >1 and <10	Y	Y	N		N	Y	N	Y		Y	N		Y	Y
No E.coli	Y	Y												
No Enterococci	Y		N		Y	Y	Y	Y		Y	Y		Y	y
Enterococci >0 and <10	Y	Y	N		Y	Y	N	Y		Y	Y		Y	Y
Nitrates<50	Y	Y	N		Y	N	Y	N		N	Y		Y	N
Arsenic<0.01	Y	Y	N		Y	Y	Y	Y		Y	Y		Y	N
Lead<0.01	Y	Y	N		Y	N	Y	Y		N	Y		Y	N
Fluoride<1.5	Y	Y	N		Y	N	Y	Y		N	Y		Y	N
pH >6.5 and <8.5	Y	Y	N		Y	Y	Y	Y		Y	Y		N	Y
Turbidity >5 and <10	Y	Y	N		N	Y	Y	N		Y	Y		N	Y
No Treatment is Required (Drinking Water)	X													
Light Treatment		X									X			
Full Treatment			X		X	X	X	X		X				
Water is Not Suitable for Consumption													X	
Water as Unsafe														X

(c) Conversion and storage of the processed data into variables to be used in the construction of the Expert System in an Access database;
(d) Insertion of the acquired variables and their respective values into the Exsys Corvid knowledge base;

Expert System Flowchart:

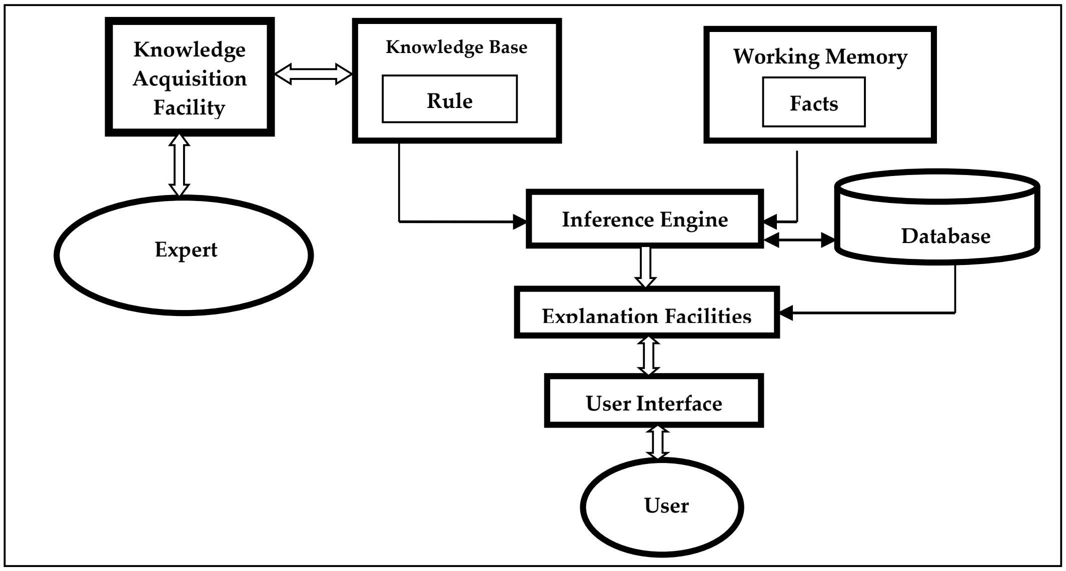

Fig. 1 Proposed model for ExpertAgua [12]

(e) Submission of the implemented knowledge base in the form of production rules that were produced in the If... Then style, the evaluation by water quality technicians or water quality analysts, who analyzed each case and made the necessary changes to the initial rules.

Table 3. Rule-base for the proposed Expertagua

Rule 1.	If no coliforms, no E.coli, no enterococci, nitrates<50,Arsenic<0.01,Lead<0.01,Fluoride<1.5, pH >6.5 and <8.5,Turbidity >5 and <10 Then No Treatment required (Drinking water)
Rule 2.	If Total Coliforms >1 and <10 , no E.coli, Enterococci>0 and <10, Nitrates < 50, Arsenic < 0.01, Lead < 0.01, Fluoride < 1.5, pH > 6.5and <8.5, Turbidity> 5 and< 10 Then Light treatment (disinfection, filtration)
Rule 3.	If Total Coliforms>10 or any E. coli, Enterococci>10, Nitrates > 50, Arsenic > 0.01, Lead > 0.01, Fluoride > 1.5, pH < 6.5 or > 8.5, Turbidity > 10 Then Full treatment (advanced filtration, chemical treatment)
Rule 4.	If the water quality indicators (microbiological, chemical, physical) are significantly outside safe limits, Then classify it as Class C, requiring full treatment methods to ensure the water is safe for drinking.
Rule 5.	If Turbidity exceeds 10 NTU, Then water must undergo full treatment such as filtration, and possibly chemical coagulation, to reduce suspended solids and pathogens.
Rule 6.	If Nitrates, Arsenic, Lead, or Fluoride exceed safe limits, Then classify water as Class C, requiring full treatment.
Rule 7.	If Total Coliforms exceed 10 CFU/100mL, or E. coli is detected, Then classify the water as Class C and apply full treatment
Rule 8.	If Turbidity exceeds 10 NTU, Then water should be classified as Class C and requires full treatment.
Rule 9.	If the Total Coliform count is above acceptable levels, Then apply light treatment methods.
Rule 10.	If Nitrates, Arsenic, Lead, or Fluoride exceed safe limits, Then water is not suitable for consumption and requires full treatment
Rule 11.	If Total Coliforms are between 1-10 CFU/100mL, Then water is classified as Class B, requiring light treatment such as disinfection
Rule 12.	If the above parameters are within the limits, Then classify water as Class A, indicating no treatment is necessary.
Rule 13.	If Turbidity exceeds 5 NTU or pH is outside 6.5 - 8.5, Then classify water as not suitable for consumption without treatment.
Rule 14.	If levels of Nitrates, Arsenic, Lead, or Fluoride exceed the WHO recommended limits, Then classify the water as unsafe (requiring treatment). If Total Coliforms, Fecal Coliforms/E. coli, or Enterococci are detected above 0 CFU/100mL, Then Class A is not applicable, and further treatment is required.

3.2. Implementation of the Expert System

The process of implementing the Expert System from the extracted and validated data followed the following steps:

(a) Conceptualization: in this phase, it was defined how the extracted information would be used and how it could be represented in the knowledge base;
(b) Interpretation and use of the data collected in the construction of the Expert System;
(c) Formalization: in this phase, the acquired knowledge was transferred for representation in the knowledge base, the modeling of the knowledge acquisition, in a system with a rule base for this purpose, the Shell Exsys Corvid was used, the knowledge was organized in the form of rules. At this time, the hardware and software to be used were also tested;
(d) Once all the variables used and all their respective values have been created, you can move on to the next step, which is creating rules. Exsys Corvid uses production rules to model human knowledge, making it ideal for selection problems where a given solution must be reached from a set of selections. A query from an Expert System aims to find the answer to a given problem. The "problems" are represented by variables, and it is necessary to define which variables will control the way the inference engine behaves.
(e) Submission of rules, evaluation by water quality technicians or water quality analysts, who analyzed each case and made the necessary changes to the initial rules;
(f) Tests in this phase were carried out on the system by simulating real day-to-day situations. The results were evaluated and validated by experts, and all components were reviewed and revised components;
(g) Implementation at this stage, the Expert System was implemented in the environment itself, that is, in use within the domain;
(h) Validation and testing in real cases by Specialists, water quality technicians and water quality analysts, both from within the institution and those hired as external consultants. Approval and implementation of the Expert System in a production environment.

3.3. System Testing with Software Exsys Corvid [14]

Fig. 2 Presentation of Exsys Corvid

Fig. 3 Presentation of Expertagua

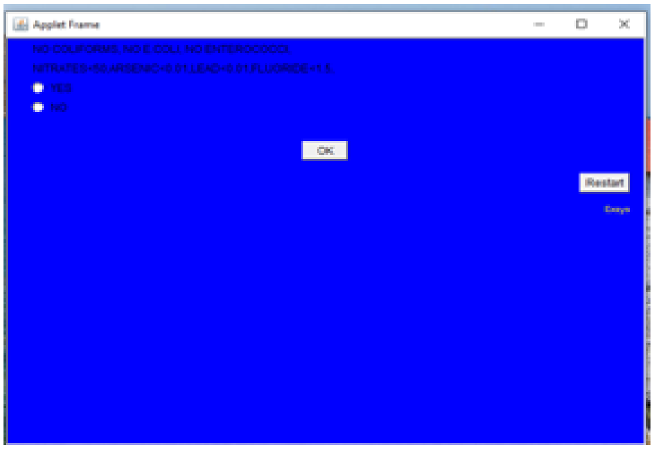

Fig. 4 Patient registration page

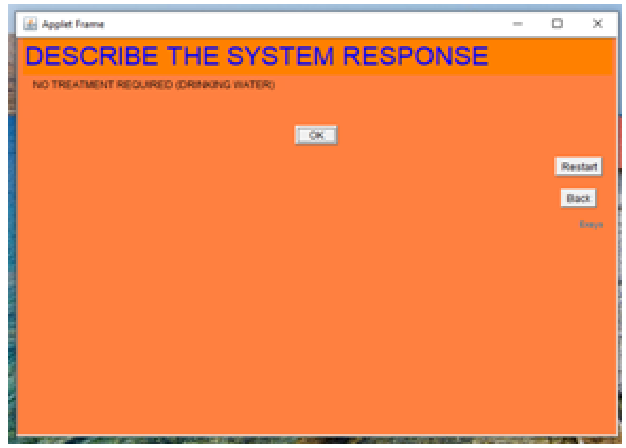

Fig. 5 Report page for the proposed ExpertAgua expert system

4. Results and Discussion

4.1. Application to Water Quality Expert Systems

In our study, we compared the classifications our rule-based expert system provided with those performed by a group of hydrology experts.

Evaluation Steps:

1. A set of 200 groundwater samples was selected and analyzed in the laboratory for parameters such as pH, nitrates, and fecal coliforms.
2. Independent classification by human experts according to WHO standards [1].
3. Predictions generated by the expert system and comparison with human assessments.
4. Calculation of the kappa coefficient to measure agreement.

Table 4. The Kappa coefficient

		Expert System	Expert
		Agree	Disagree
Expert System	Agree	6	3
Expert	disagree	2	4

$$P_e = \left(\frac{9 \times 8}{15^2}\right) + \left(\frac{7 \times 6}{15^2}\right)$$

$$P_o = \frac{6+4}{6+3+2+4} = \frac{10}{15} = 0.6667$$

The formula gives the Kappa coefficient:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \qquad \kappa = \frac{0.6667 - 0.5067}{1 - 0.5067} = \frac{0.16}{0.4933} \approx 0.324$$

Interpretation: a kappa coefficient of 0.324 indicates moderate agreement between the two assessments, indicating near-perfect agreement between our system and human experts. This demonstrates that our rule-based approach is reliable for assessing water quality [15].

5. Conclusion

Le système expert basé sur des règles s'est avéré efficace dans la classification de la qualité des eaux souterraines, fournissant une analyse rapide et accessible. Le coefficient Kappa de Cohen a validé la précision et la fiabilité de notre système expert par rapport aux évaluations humaines. Avec un Kappa de 0.324, notre approche se distingue comme une alternative efficace et accessible aux modèles plus complexes, tout en restant robuste aux réalités de l'Afrique tropicale. Les résultats de cette étude soulignent l'utilité d'un système expert basé sur des règles pour améliorer la gestion de la qualité des eaux souterraines et ainsi réduire les risques de maladies émergentes dans les zones tropicales d'Afrique. Avec de petites adaptations aux règles et aux paramètres, le système peut être reproduit dans d'autres régions tropicales d'Afrique, en respectant les variations géologiques et socio-environnementales. L'outil peut aider les autorités locales et les ONG à identifier les domaines critiques qui nécessitent une intervention immédiate ou un suivi continu. Bien qu'efficace, le système dépend de la qualité des données d'entrée. Les versions futures devraient intégrer l'apprentissage automatique pour affiner les règles en fonction des données historiques et intégrer des alertes en temps réel.

References

[1] World Health Organization, Guidelines for Drinking-Water Quality, 4th Edition, Incorporating the the First and Second Addenda, 2022**.** [Online]. Available: https://www.who.int/publications/i/item/9789240045064

[2] World Health Organization, Guidelines for Drinking-Water Quality, 4th Edition, Incorporating the 1st Addendum, 2017. [Online]. Available: https://www.who.int/publications/i/item/9789241549950

[3] Bruce G. Buchanan, and Edward H. Shortliffe, *Rule-Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley Longman Publishing Co., Inc., United States, 1984. [Google Scholar] [Publisher Link]

[4] Shu-Hsien Liao, "Expert System Methodologies and Applications — A Decade Review from 1995 to 2004," *Expert Systems with Applications*, vol. 28, no. 1, pp. 93-103, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ajith Abraham, *Rule-based Expert Systems*, Handbook of Measuring System Design, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[6] Jan Vantilenen, "Knowledge Acquisition and Validation Using a Decision Table Engineering Workbench", *The World Congress on Expert Systems*, Pergamon Press, pp. 1861-1868, 1991. [Google Scholar]

[7] M.R. Verhelst, *De Praktijk van Beslissingstabelle*, Kluwer,1980. [Google Scholar]

[8] John Durkin, *Expert Systems: Catalog of Applications*, Intelligent Computer Systems Inc., 1993. [Google Scholar]

[9] Subhasis Giri, "Water quality prospective in Twenty First Century: Status of water quality in major river basins", contemporary strategies and impediments: A review," *Environmental Pollution*, vol. 271, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] World Health Organization, Guidelines for Drinking-Water Quality, 4th Edition, 2011. [Online]. Available: https://iris.who.int/bitstream/handle/10665/44584/9789241548151_eng.pdf

[11] World Health Organization, Guidelines for Drinking-Water Quality: Small Water Supplies, 2024. [Online]. Available: https://www.who.int/publications/i/item/9789240088740

[12] K.S. Adewole, M.A. Hambali, and M.K. Jimoh, "Rule-Based Expert System for Disease Diagnosis," *Book of Proceedings, International Science, Technology, Engineering, Arts, Management and social Sciences (iSTREAMS) Multidisciplinary Conference*, vol. 7, 2015.

[13] Jacob Cohen, "A Coefficient of Agreement for Nominal Scales. Educational and Psychological Measurement," *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37-46, 1960. [CrossRef] [Google Scholar] [Publisher Link]

[14] Exsys,"Exsys Corvid Software for Expert System Development Tool, 2011. [Online]. Available: http://www.exsys.com/exsyscorvid.html

[15] J. Richard Landis, and Gary G. Koch, "The Measurement of Observer Agreement for Categorical Data," *Biometrics*, vol. 33, no. 1, pp. 159-174, 1977. [CrossRef] [Google Scholar] [Publisher Link]