



An Exploration for the Analysis of Ground Water Possibility in Case Based Recommendation System using Hybrid Cuckoo Search and Artificial Honey Bee Algorithm

¹Sheetal Jethi and ²Harsh Sadawarti

¹Research Scholar, I.K.G Punjab Technical University, Jalandhar

²Campus Director, Regional Institute of Management and Technology, MandiGovindgarh

Abstract: Groundwater Possibility Exploration is one of the prime problems in human life. Human and other social species are directly or indirectly dependent on water resources. But with increasing population and instant changing environmental conditions, groundwater resources are also becoming less day by day. One approach of manual exploration can be used but that is most laborious and time consuming task. So, there is need of some autonomous approach with which we can find the more groundwater resources without digging the bore well. This paper presents an autonomous approach in case based recommendation system using hybrid cuckoo search and artificial honey bee algorithm for groundwater possibility detection. The proposed concept is structured in the manner to have the input in the form six attributes of slope, geology, landuse, lineament, soil & landform and give output in the form of low, intermediate and high possibility. In this method, case based reasoning used in the manner to retrieve the previous knowledge of use cases. The cases are actually the host bird's nest and input is cuckoo's egg. CS originates from the behaviour of certain species of cuckoo which lay eggs in other birds nest in parasitic manner. If cuckoo's egg adapts the behaviour of host nests it will exist otherwise the host bird will discard cuckoo's egg. The minimum similarity value is calculated using ABC algorithm. CBR life cycle further uses the information from this metaphor. The overall method is evaluated using the parameters of sensitivity, specificity and accuracy.

Keywords: Case based Reasoning, Cuckoo Search, Artificial Honey Bee Algorithm, Groundwater possibility Exploration, Land Cover Attributes

I. INTRODUCTION

From existence of humans on earth it is water that guides the settlements that decide what to grow and what to sow, people need it, fight for it and above all can't survive without it [1]. We can say human existence on earth is directly proportional to water availability. It is fundamental right of every human, animal and plant to have access to the requisite quantity water [2]. But the resources of water are decreasing with time. There is no single "magic bullet" that can resolve this rising water deficiency problem. One of the major resources of fresh water on the earth surface is ground water. Due to unseasonal rain on earth surface, people usually depend upon groundwater. Groundwater constitutes only 0.6% of all the water on this earth planet, 97.4% accounts for seawater and 2% for snow and ice on the poles [3]. On the earth surface, groundwater is available in very less amount. Therefore, groundwater is an important commodity which we use for various purposes such as agricultural, industrial and domestic use but with the increase in population its resources are depleting and hence the necessity to find its resources arises. Here detection for possibility of ground water at a particular region is taken as an application [4].

This paper presents an autonomous approach in case based recommendation system using hybrid cuckoo search and artificial honey bee algorithm for groundwater possibility detection. The proposed concept is structured in the manner to have the input in the form six attributes of slope, geology, landuse, lineament, soil & landform and give output in the form of low, intermediate and high possibility. The overall method is evaluated using the parameters of sensitivity, specificity and accuracy.

Other sections of the paper are described as: Section II presents the related work of groundwater possibility detection. Section III discusses about the considered basic concepts of Cuckoo Search, Case Based Reasoning, Artificial Bee Forging etc. for proposed approach, Section IV presents the considered dataset, Section V discusses about proposed concept, Section VI shows the calculated results based on the considered parameters and Section VII concludes the paper.

II. RELATED WORK

In this section, existing concept for the groundwater possibility detection are explained. The considered concepts are explained based on the dataset as discussed above in section III for groundwater possibility detection.

Panchal et al. [5] has proposed case based reasoning system with cuckoo search algorithm for the prediction of groundwater possibility. Authors have calculated the possibility in the form of low, moderate and high solutions. The main assumption considered for this low, moderate and high concept is existence of host nest iteration as per cuckoo bird. Subramaniam et al. [6] proposed the case based reasoning model with semantic intelligence for deciding the excellent tuned resolution from the quantity of dispersed cases. They proposed the case model to incorporate the preferred data types and certificates through the restricted or comprehensive authorities for preferring the solution. The n-tier model of reasoning could be performed using the case agency and broker machinery to file and arbitrated the consequences for the submitted cases. Bisht et al. [7] have used the fuzzy logic and Artificial Neural Network for the simulation of water table elevation as an application. Authors

have used the dataset of Budaun district and developed the ANN based model for the training, testing and validation of proposed concept.

Further, Panchal et al. [8] proposed a new groundwater possibility detection system based on particle swarm optimization (PSO) and case based reasoning (CBR). A new concept of waves of swarm (WOS) derived from PSO is introduced in this study that operates on the problem case. The integration approach proposed here improves the retrieval accuracy of CBR using WOS. Panchal et al. [9] have also evaluated the groundwater possibility with Biogeography Based Optimization (BBO). The concept is BBO is used with the possible cases of case based recommender system. BBO bird is the intersection of multidimensional search space. Rules are developed to refine the BBO for each iteration.

III. BASIC CONCEPTS

This section describes the basic concepts of case based Reasoning, Cuckoo Search and Honey Bee Optimization algorithm

A. Case Based Reasoning (CBR)

Case-based reasoning is an Artificial Intelligence based approach but it works differently as compare to other AI based approaches in the manner that CBR uses previously experience based knowledge instead of solely dependent on problem domain, their description and available resources. The previous experienced based knowledge is considered as the cases for the problem solution. These cases are considered as the iterations to solve the problem. Another advantage of CBR approach is have the incremental solution for each time due to repetition of results improved which leads to overall higher efficient solution of each problem. [10].

The internal structure of CBR mechanism is categorized into components: case reasoner and the case retriever [11]. The appropriate cases in the case base can be retrieve by case retriever and further case reasoner uses the retrieved cases to find the solution of the problem [12]. This reasoning process generally involves both determining the differences between the cases retrieved and the current case, and modifying the solution to reflect these differences appropriately. The components of CBR system are as shown in figure 1.

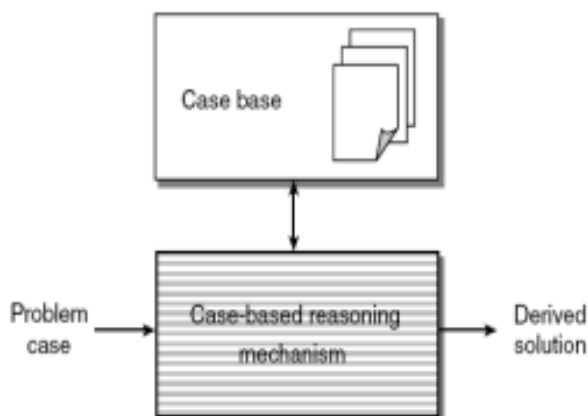


Figure 1: Case Based Reasoning

B. Cuckoo Search (CS)

CS originates from the behavior of certain species of cuckoo which lay eggs in other birds nest in parasitic manner [13]. If cuckoo's egg adapts the behavior of host nests it will exist otherwise the host bird will discard cuckoo's egg. The process of evolving to best lay parasitic egg is the main aim of CS [14]. The algorithm is described below:

- i) A comparison is done between cuckoo's egg and set of host nests.
- ii) Levy flight is used to introduce randomness to choose host nests.
- iii) Comparison produces two sets of solution; one is quality solution and the other discarded solution.
- iv) Depending on the ranking function best solution is found out from quality solution and worst nests are removed with probability p_a .

C. Honey Bee Optimization Algorithm

Honey Bee Optimization (HBO) is a meta-heuristic based swarm intelligence technique inspired from the natural/social agents namely bees (Honey bees). In 2005, karaboga, has introduced the nature inspired approach of HBO to optimize the solutions of computational problems [15]. The two main properties of HBO algorithm that supports it to fir in swarm intelligence approach are collective behavior and self organizing behavior.

HBO algorithm consists of mainly two components as mentioned: Food Sources and Honey Bees. Honey bees are further categorized into three categories as mentioned: Employed Bees, Onlooker Bees and Scout Bees.

Initially, employee bees search for the random food sources and gather the useful information of landscape. Further, this landscape information has been shared with the onlooker bees which further evaluate the information by using a probabilistic approach like Roulette Wheel Method [16]. These evaluation results have further used to initiate the neighbor search. Scout bees finally exploit the available food sources with random search and determine the quality of food sources. Scout bees confirm the sugar level quality in flower patches and inform the other bees available in hives waiting for quality food source information. This communication of scout bees with other available bees is performed with "wangle dance". With wangle dance, scout bees produce loud buzzing sound and share the information of food source quality, direction and distance of food source from hive.

At first, HBO was proposed to find the solution of optimization of combinatorial problems. HBO works in two stages as mentioned: forward stage and backward stage. In forward stage, solution for the combinatorial problems has been generated in individual meaner and collective experience has been shared. This partial generated solution has been further utilized in backward stage and quality has been assured with probability distribution function. Overall concept work in local and global search solution processes.

IV. DATASET CONSIDERED

The proposed integrated concept is used to explore the groundwater possibility. For this, we have used the expert dataset having six attributes of slope, landuse, landform, geology, soil type and lineament. The considered dataset can be used for any location to test the possibility of groundwater. An instance of expert dataset is given in figure 2.

V. PROPOSED CONCEPT

This section presents the integrated proposed concept of Case Based Reasoning with Cuckoo search and Artificial Honey Bee Optimization for groundwater exploration without digging the borewell. In this integrated approach, Case Based Reasoning uses previously experience based knowledge instead of solely dependent on problem domain, their description and available resources. Cuckoo Search and Artificial Bee Colony Optimization has been used for the local best and global optimization. Here, we have considered the integrated approach to optimize the solution upto the possible extent for groundwater possibility detection.

In this proposed concept, user query is considered as the input and groundwater possibility is determined as output. Initially, use cases are created using the expert dataset. The cases are actually the host bird's nest and input is cuckoo's egg. CS originates from the behaviour of certain species of cuckoo which lay eggs in other birds nest in parasitic manner. If cuckoo's egg adapts the behaviour of host nests it will exist otherwise the host bird will discard cuckoo's egg. The similarity of the new egg (user query) with host nest is calculated. The minimum similarity value is calculated using HBO algorithm. This value is stored in the optimized parameter. The nests (cases) which have similarity value more than optimized parameter are considered as quality solution. Rest of the nests (cases) are discarded as worst solutions. The quality solutions (cases) are ranked by calculating correlation. The nest (case) having maximum correlation is considered as best solution. After the number of iterations, we can get the possibility results of low, moderate or higher. The work flow of proposed work is shown in figure 4. Complete formulation is given below:

Input: User Query.

Output: Estimation of Groundwater (Low, Moderate, High).

ALGORITHM

Step 1: Consider the training dataset for the use cases (host nest).

Step 2: Insert the user query for groundwater detection with some attribute values

Step 3: Evaluate the Feature weights for the user query data and training dataset as shown in equation (1.1).

$$wt = \text{numberofoccurrence} \cdot \log_2 \left(\frac{\text{totalnumberoffeatures}}{\text{numberofsamplesinthedataoccuredin}} \right) \dots \text{Equation (1.1)}$$

Step 4: Initialize the population of HBO species and set the value of max_iteration as per total case bases.

Step 5: For max_iteration {

5.1 Calculate the Cosine similarity of input query with the available case bases by considering equation (1.2):

$$\text{Similarity} = \cos(\theta) = \frac{wt_A \cdot wt_B}{\|wt_A\| \|wt_B\|} = \frac{\sum_{i=1}^n wt_{Ai} * wt_{Bi}}{\sqrt{\sum_{i=1}^n (wt_{Ai})^2} * \sqrt{\sum_{i=1}^n (wt_{Bi})^2}} \dots \text{Equation (1.2)}$$

Where, wt_A is the query data weight and wt_B is the data weight of some particular case base.

Figure 2: Considered Training Dataset

Different attributes are described with their possible features. These six attributes are further subcategorized into their respective fields as shown in table 1.

Table 1: Attributes and their Subcategories

Attributes	Values
Lineament	Present, Absent
Slope	Steep, Gentle
Geology	Metamorphic, Igneous, Sedimentary, Older alluvium, Younger alluvium
Land use	Wasteland, Forest, Grass, Fallowland, Swampy land, Cultivatedland, Shrubs, Buildup, Agriculturalland, Urban, Waterbody, mixed vegetation etc.
Soil	Gravelsand, Sandygravel, Coarsesand, Sand, Clayloam, Alluvialsand, Gravel Sand Pebbles, Sandyloam, Rocky etc.
Landform	DeltaicPlain, Floodplain, Pediment, Bajada, Riverterraces, Alluvialfans, Pediplain, Buriedpediment, AlluvialPlain, Intermontanevalley, Wadi, Oldmeander etc.

5.2 Memorise the similarity for each case base.

- Step 6: Find the minimum similarity value using the artificial honey bee algorithm.
 Step 7: Store the minimum similarity value in optimized_parameter.
 Step 8: Nests having similarity value more than optimized_parameter are considered as quality solution.

- Step 9: Rest all are discarded as worst nest.
 Step 10: Rank the quality solution by finding the correlation.
 Step 11: The nest having maximum correlation is ranked as best solution.

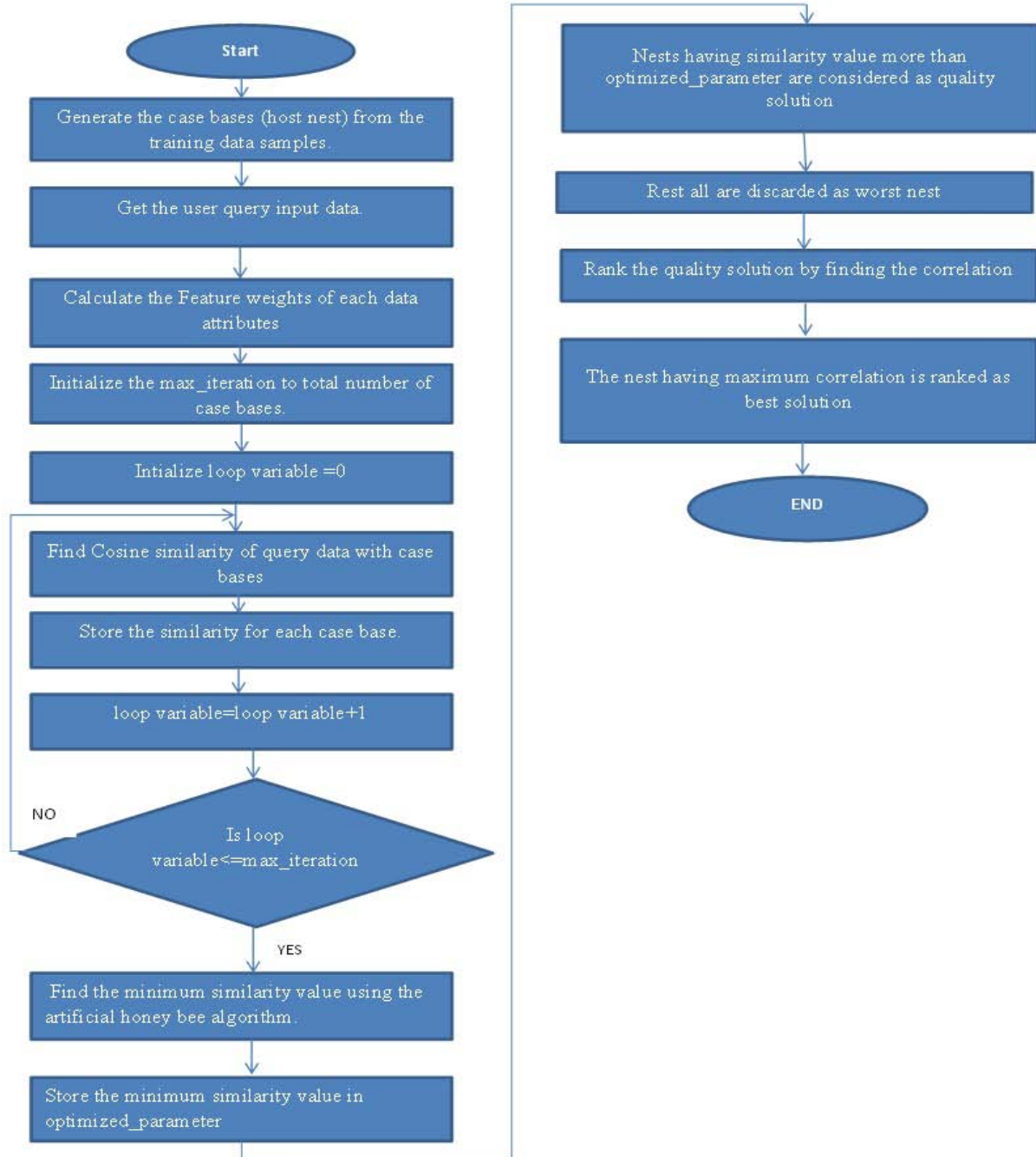


Figure 4: Work Flow for Groundwater Estimation using Proposed Algorithm

Further the explanation of this flow of work is explained as below:

Step 1: Initially, the number of case base (number of host nest) have been generated using the training dataset.

Step 2: Then, user enter the query values for the available attributes in generated graphical user interface. Further, this information is considered as the new cuckoo egg information.

Step 3: Here, the feature weights for each data attribute from training dataset is evaluated using the equation:

$$wt = \text{number of occurrence} \cdot \log_2 \left(\frac{\text{total number of features}}{\text{number of samples the data occurred}} \right)$$

Step 4: Here, initialization of maximum iterations is performed as per the number of cases in the case base.

Step 5: Here, similarity function has been evaluated to find the similarity between the use case and query data by making use of cosine similarity function. This similarity for each case is further stored. This similarity has been evaluated as defined below:

$$\text{Similarity} = \cos(\Theta) = \frac{wt_A \cdot wt_B}{\|wt_A\| \|wt_B\|} = \frac{\sum_{i=1}^n wt_{Ai} * wt_{Bi}}{\sqrt{\sum_{i=1}^n (wt_{Ai})^2} * \sqrt{\sum_{i=1}^n (wt_{Bi})^2}}$$

Step 6: The minimum similarity value is calculated using the Algorithm to predict ground water possibility in case based recommendation system using artificial Honey Bee method.

Step 7: The minimum similarity value is stored in optimized parameter which is used by cuckoo search algorithm.

Step 8: The nests (cases) which have similarity value more than optimized parameter are considered as quality solution.

Step 9: Rest of the nests (cases) are discarded as worst solutions.

Step 10: The quality solution (cases) are ranked by calculating correlation.

Step 11: The nest (case) having maximum correlation is considered as best solution.

VI. RESULTS & DISCUSSION

This section determines the evaluated results for the proposed algorithm in the form of groundwater possibility to be low, intermediate or high. Also the overall comparative parameters of specificity, sensitivity and accuracy are evaluated.

A. Results

The proposed integrated concept of CBR with CS and HBO is implemented in MATLAB with GUI (Graphical User Interface). The considered input attributes are landuse, lineament, soil type, landform, slope and geology. Output is shown with the groundwater possibility of Low, Moderate and higher. The low probability shows groundwater possibility upto 64%. Moderate probability shows groundwater possibility from 65% to 84% and higher values are from 85% to 100%. A sample GUI is shown in figure 5.

As per the shown figure 5, output is determined in low, moderate and high possible value solutions. For the output results, we have considered some test cases.

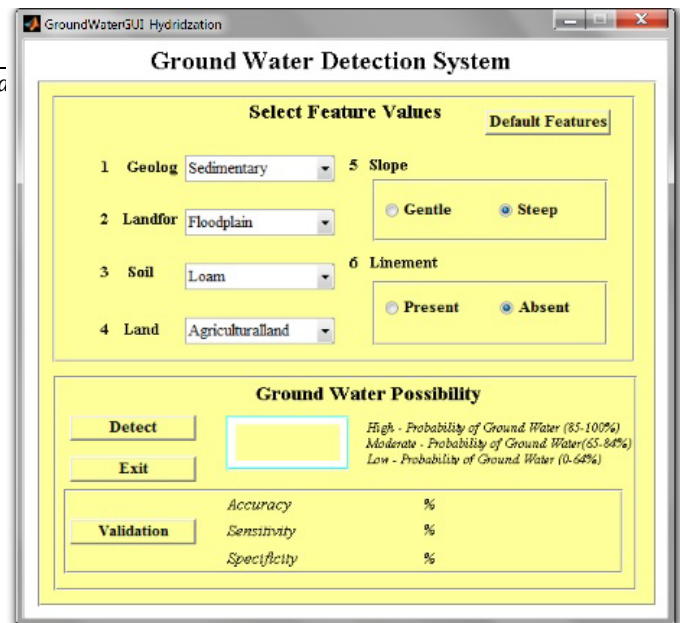


Figure 5: Sample GUI for Groundwater Exploration

As per the shown figure 5, output is determined in low, moderate and high possible value solutions. For the output results, we have considered some test cases which as shown as below.

1). Test Case 1

In test case 1, the considered attributes with their subcategories are shown in table 2.

Table 2: Test Case 1

Attribute	Attribute Value
Lineament	Present
Slope	Gentle
Land use	Cultivated land,
Soil	Coarse sand
Landform	Intermundane valley
Geology	Younger alluvium
Groundwater possibility	?

We have considered the above test case attributes to calculate the groundwater possibility. The output result with performance of the proposed concept is shown by figure 6 below:

The considered sub-attributes of Lineament: present, Slope: Gentle, Landuse: Cultivated land, Soil type: coarse sand, Landform: Intermundane valley and Geology: Younger alluvium. For these attributes, output results with performance evaluation parameters are shown in figure 6.

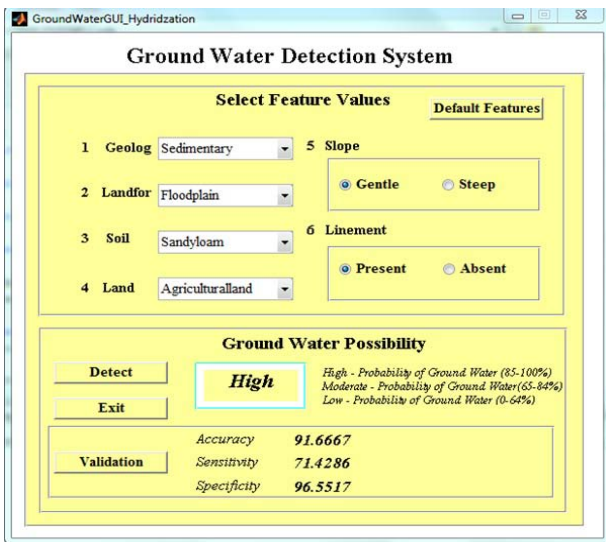


Figure 6: Test Case 1 for Groundwater Possibility

Figure 6 indicates the “High” probability of groundwater possibility.

2). Test Case 2

In test case 2, the considered attributes with their subcategories are shown in table 3.

Table 3: Test Case 2

Attribute	Attribute Value
Lineament	Present
Slope	Gentle
Land Use	Cultivatedland,
Soil	Coarsesand
Landform	Intermundanevalley
Geology	Younger alluvium
Groundwater possibility	?

The considered subattributes of Lineament: Present, Slope: Gentle, Landuse: Cultivated land, Soil type: coarse sand, Landform: Intermundane valley and Geology: Younger alluvium. For these attributes, output results with performance evaluation parameters are shown in figure 7.

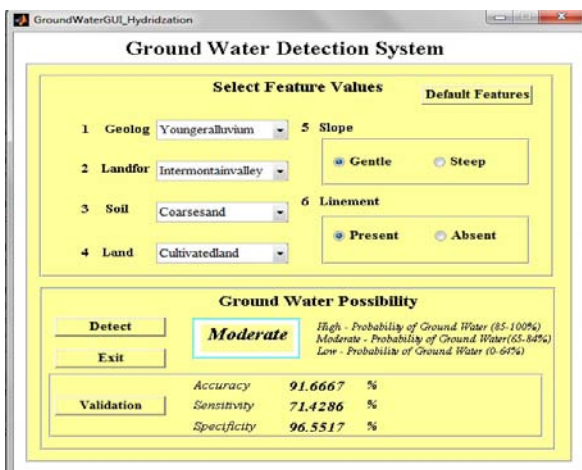


Figure 7: Test Case 2 for Groundwater possibility

Figure 7 indicates the “Moderate” probability of groundwater possibility.

3). Test Case 3

In test case 3, the considered attributes with their subcategories are shown in table 4.

Table 4: Test Case 3

Attribute	Attribute Value
Lineament	Absent
Slope	Steep
Land Use	Urban
Soil	Rocky
Landform	Bajada
Geology	Igneous
Groundwater possibility	?

The considered subattributes of Lineament: Absent, Slope: Steep, Landuse: Urban, Soil type: Rocky, Landform: Bajada and Geology: Igneous. For these attributes, output results with performance evaluation parameters are shown in figure 8.

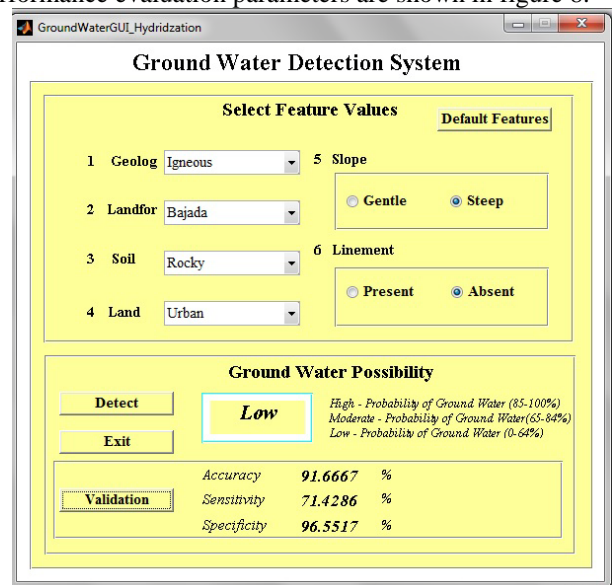


Figure 8: Test Case 3 for Groundwater possibility

Figure 8 indicates the “Low” probability of groundwater possibility.

From the considered test cases, we can say that groundwater possibility varies in the form of Low, Moderate and High. There are many more cases for the possibility of low, moderate and high. The considered dataset cases are compared with the results of proposed algorithm and calculated sensitivity, specificity and accuracy values.

B. Evaluation Parameters

In this research work, we have considered the performance parameters of Specificity, Sensitivity and Accuracy. These evaluation parameters have been discussed as below:

1). Sensitivity: It denotes the proportion to measure the negative outcomes of concept. In other words, we can say that it is the method to have the inaccurate value of groundwater possibility as per expert dataset. This can be calculated as:

$$\text{Specificity} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}$$

...Equation (1.3)

2). *Specificity*: It denotes the proportion to measure the positive outcomes of concept. In other words, we can say that it is the method to have the accurate value of groundwater possibility as per expert dataset. This can be calculated as:

$$\text{Specificity} = \frac{\text{True negatives}}{\text{True negatives} + \text{False positives}}$$

...Equation (1.4)

3). *Accuracy*: It defines the combined values of Sensitivity and Specificity. It can be calculated as below:

$$\text{Accuracy} = \frac{\sum \text{True positive} + \sum \text{True negative}}{\sum \text{Total Available Dataset Cases}}$$

...Equation (1.5)

From the output results and evaluation parameters as shown in figure 6 to figure 8, values are as mentioned in table 6.

Table 6: Parametric values

Parameter	Values (%)
Accuracy	91.6667
Sensitivity	71.4286
Specificity	96.5517

This can also be represented in the form of graphical representation as shown in figure 9.

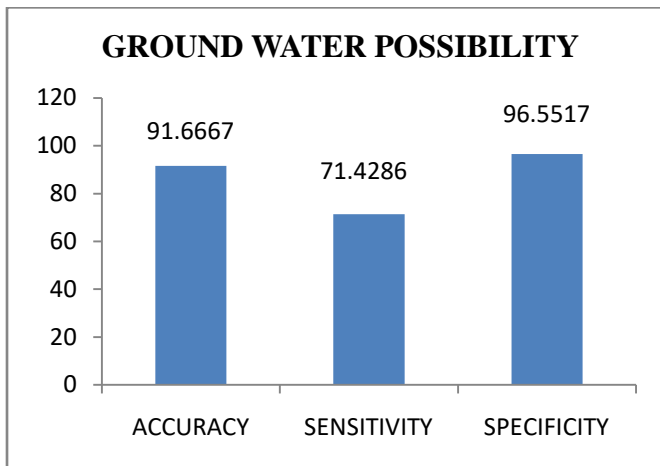


Figure 9: Ground water possibility

VII. CONCLUSION

Ground water is the precious natural resource of water available on the earth. But with increasing use of water, there is need to search more natural groundwater resources. In this paper, we have applied the Case Based Reasoning approach with integrated Cuckoo Search and Honey Bee Colony Optimization algorithm for the exploration of more groundwater resources under the soil. To generate the case

base for CBR system, training dataset has been used. The cases are actually the host bird's nest and input is cuckoo's egg. CS originates from the behaviour of certain species of cuckoo which lay eggs in other birds nest in parasitic manner. If cuckoo's egg adapts the behaviour of host nests it will exist otherwise the host bird will discard cuckoo's egg. The similarity of the new egg (user query) with host nest is calculated. The minimum similarity value is calculated using ABC algorithm. This value is stored in the optimized parameter. The nests (cases) which have similarity value more than optimized parameter are considered as quality solution. Rest of the nests (cases) are discarded as worst solutions. The quality solutions (cases) are ranked by calculating correlation. The nest (case) having maximum correlation is considered as best solution. Groundwater possibility depends upon the different attributes of earth surface like slope, geology, land use, lineament, soil & landform. The outcome of the experimented test cases has been shown in Figure 6 to 8. For the numerical evaluation of proposed concept, parameters of Sensitivity, Specificity and Accuracy have been considered. The calculated values of Specificity, Sensitivity and Accuracy are 96.5517%, 71.4286%, and 91.6667% respectively. From these calculated values, we can say that the proposed concept is efficient enough to estimate the possibility of ground water.

VIII. REFERENCES

- [1]. Daily, Gretchen. *Nature's services: societal dependence on natural ecosystems*. Island Press, 1997.
- [2]. Famiglietti, J. S. "The global groundwater crisis." *Nature Climate Change* 4, no. 11 (2014): 945-948.
- [3]. Gleeson, Tom, Yoshihide Wada, Marc FP Bierkens, and Ludovic PH van Beek. "Water balance of global aquifers revealed by groundwater footprint." *Nature* 488, no. 7410 (2012): 197-200.
- [4]. Teeuw, Richard M. "Groundwater exploration using remote sensing and a low-cost geographical information system." *Hydrogeology Journal* 3, no. 3 (1995): 21-30.
- [5]. Panchal, V.K., Bidisha Das, and Daya Gupta. "Applying case based reasoning in cuckoo search for the expedition of groundwater exploration." In *Proceedings of Seventh International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA 2012)*, pp. 341-353. Springer India, 2013.
- [6]. Subramaniam, Chandrasekaran, V. D. Dhandayudhapani, Niveditha Narendhran, and Mohammed Nazim Feroz. "Distributed Case Based Reasoning Model with Semantic Intelligence." *International Journal of Information and Electronics Engineering* 3, no. 2 (2013): 136.
- [7]. Bisht, Dinesh, Shilpa Jain, and M. Mohan Raju. "Prediction of water table elevation fluctuation through fuzzy logic & artificial neural networks." *International Journal of Advanced Science and Technology* 51, no. 7 (2013): 107-120.
- [8]. Panchal, V. K., Harish Kundra, and Navpreet Kaur. "A Novel Approach to Integration of waves of swarms with case based reasoning to detect groundwater potential." In *8th Annual Asian Conference &*

- Exhibition of Geospatial information technology & application, Map Asia, Singapore. 2009.*
- [9]. Panchal, V., Harish Kundra, and Amanpreet Kaur. "An integrated approach to biogeography based optimization with case based reasoning for retrieving groundwater possibility." *International Journal of Computer Applications* 1, no. 8 (2010): 975-8887.
- [10]. Richter, Michael M., and Rosina O. Weber. "Case-Based Reasoning." *A Textbook* (2013).
- [11]. Watson, Ian, and FarhiMarir. "Case-based reasoning: A review." *The knowledge engineering review* 9, no. 04 (1994): 327-354.
- [12]. deMantaras, Ramon Lopez. "Case-based reasoning." In *Machine Learning and Its Applications*, pp. 127-145. Springer Berlin Heidelberg, 2001.
- [13]. Yang, Xin-She, and Suash Deb. "Cuckoo search via Lévy flights." In *Nature & Biologically Inspired Computing, 2009. NaBIC 2009. World Congress on*, pp. 210-214. IEEE, 2009.
- [14]. Yang, Xin-She, and Suash Deb. "Engineering optimisation by cuckoo search." *International Journal of Mathematical Modelling and Numerical Optimisation* 1, no. 4 (2010): 330-343.
- [15]. Karaboga, Dervis. *An idea based on honey bee swarm for numerical optimization*. Vol. 200. Technical report-tr06, Erciyesuniversity, engineering faculty, computer engineering department, 2005.
- [16]. Karaboga, Dervis, and BahriyeBasturk. "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm." *Journal of global optimization* 39, no. 3 (2007): 459-471.