https://doi.org/10.31436/iiumej.v27i1.3830

Advanced Groundwater Level Forecasting using QSO-based Vision Transformer Model for Sustainable Water Resource Management

V. GOKULA KRISHNAN[1*], M. S. MAHARAJAN[2], B.V. SUBBA RAO[3], G.MAHALAKSHMI[4], M. ARUNADEVI[5]

[1*] *Department of CSE, Easwari Engineering College, Ramapuram, Chennai, Tamil Nadu, India*
[2] *Department of AIDS, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India*
[3] *Department of Information Technology, PVP Siddhartha Institute of Technology, Kanuru, Vijayawada, India*
[4] *Department of AIDS, Vel Tech High Tech Dr.Rangarajan Dr.Sakunthala Engineering College, Avadi, Chennai, Tamil Nadu, India*
[5] *Department of CSE, K. Ramakrishnan College of Technology, Trichy, Tamil Nadu, India*

**Corresponding author: gokul_kris143@yahoo.com*

(Received: 7 June 2025; Accepted: 4 September 2025; Published online: 12 January 2026)

ABSTRACT: The reliable predictions of groundwater levels are crucial for long-term management of water resources, and they are an excellent source for human well-being. While conventional ML approaches work well for low-dimensional data, they are optimized for hyperparameters on high-dimensional inputs and also capture complex temporal correlations. To address these restrictions, this research presents a new framework for predicting groundwater-level changes up to five months in advance, built on the Vision Transformer (ViT) and optimised with Quokka Swarm Optimisation (QSO). ViT enables strong global feature extraction and long-range dependency modelling by processing time-series data as sequential image-like patches, in contrast to traditional neural networks. Drawing on quokka's adaptive survival behaviour, the QSO procedure optimises the transformer's hyperparameters, such as patch size, attention heads, and depth, in real time to enhance prediction accuracy. The ViT+QSO outperformed baseline deep learning methods on groundwater datasets from Southern Africa in terms of RMSE, MAE, correlation coefficient (R), and Nash-Sutcliffe Efficiency (NSE). Quantile regression uncertainty quantification further improves the model's reliability for water resource planning. Hydrological variables, in addition to climate indices, influence groundwater fluctuations, as confirmed by ablation research. The proposed ViT-QSO achieved 93% R and 0.118 MAE, whereas the basic ViT achieved only 89.1% R and 0.152 MAE for groundwater-level prediction. Scalability, interpretability, and suitability for areas with limited monitoring infrastructure are hallmarks of the proposed methodology. This research provides valuable insights into how to better withstand the effects of climate change, in addition to human activities, on groundwater supplies.

ABSTRAK: Ramalan paras air bawah tanah yang boleh dipercayai adalah penting bagi pengurusan jangka panjang sumber air dan kesejahteraan manusia; namun, pendekatan pembelajaran mesin konvensional berhadapan kekangan pengendalian input berdimensi tinggi serta model korelasi temporal kompleks. Kajian ini mencadangkan satu rangka kerja baharu berasaskan Pengubah Visi (ViT) yang dioptimum menggunakan Optimisasi Kawanan Quokka (QSO) dalam meramal perubahan paras air bawah tanah pada lima bulan lebih awal. ViT memproses data siri masa sebagai tampalan jujukan menyerupai imej bagi membolehkan pengekstrakan ciri global dan model kebergantungan jarak jauh, manakala QSO mengoptimumkan hiperparameter pengubah secara adaptif dalam meningkatkan ketepatan ramalan. Model ViT-QSO menunjukkan prestasi unggul berbanding kaedah pembelajaran mendalam asas pada set data air bawah tanah di Afrika Selatan, dengan peningkatan ketara

dari segi RMSE, MAE, pekali korelasi (R), dan Kecekapan Nash–Sutcliffe (NSE), serta mencapai nilai R sebanyak 93% dan MAE 0.118 berbanding ViT asas masing-masing mencatatkan 89.1% dan 0.152. Pengkuantitian ketidakpastian melalui regresi kuantil meningkatkan kebolehpercayaan model bagi perancangan sumber air, manakala kajian ablasi mengesahkan peranan pembolehubah hidrologi dan indeks iklim terhadap turun naik paras air bawah tanah. Secara keseluruhan, metodologi yang dicadangkan adalah berskala, boleh ditafsir, dan sesuai pada kawasan infrastruktur pemantauan terhad, serta memberikan sumbangan penting dalam menangani kesan perubahan iklim dan aktiviti manusia terhadap sumber air bawah tanah.

KEYWORDS: *Quokka Swarm Optimization; Vision Transformer; Southern African groundwater datasets; Deep Belief Networks; Water Resources.*

1. INTRODUCTION

A growing frequency of extreme weather events driven by climate variability has highlighted the critical need for a reliable water supply and the organised management of water resources. A vital freshwater resource that can help resolve these issues is groundwater. To plan, develop, and implement effective groundwater resource management systems, it is necessary to continuously and systematically monitor groundwater change [1]. To ensure the sustainable use of groundwater, various government agencies in South Korea operate observation networks nationwide. If a farmer wants to know how much groundwater is available and how to manage it effectively, they need precise projections of spatial and temporal changes in groundwater storage. The hydrogeological characteristics and boundary conditions of groundwater systems, along with the impacts of weather, water, and land-cover changes, can be used to infer groundwater fluctuations. Due to complex spatial variability in groundwater and pronounced nonlinearity, systematic groundwater predictions are difficult [2].

Using a combination of data from multiple satellites, including GRACE (Gravity Recovery and Climate Experiment) and GRACE-Follow-On, TRMM (Tropical Rainfall Measuring Mission), and Landsat (based on the water balance method), remote sensing methods have been suggested as a means to estimate changes in spatiotemporal groundwater storage in recent years. Due to the poor spatial/temporal resolution of satellite data, these methods are unsuitable for predicting or nowcasting changes in groundwater storage, despite their relatively straightforward and reliable evaluation across many locations [3]. Moreover, vegetation or human-caused impacts are frequently overlooked in GRACE-based water-balance analyses, which consider only hydrological factors (terrestrial water storage, snow water equivalent, surface water, and soil moisture) [4]. Recent research has shown that changes in land cover, NDVI, and MNDWI are among the variables that determine groundwater recharge and potential. This study recommends employing multiple satellite datasets while accounting for factors that affect the reliability of the results.

Since it is difficult to predict how changes across Earth's system will affect groundwater storage, it is essential to examine how groundwater responds to these changes by assessing correlations among climatic variables, surface water extent, hydrological components, and vegetation [5]. Complex issues in computer vision, image classification, and time-series forecasting have motivated the development of machine learning techniques that account for nonlinear interdependencies, including ANN, SVM, and RF. Numerous ML approaches, including ANN models, hybrid feedforward neural networks (FNNs), and nonlinear autoregressive models with exogenous inputs (NARXs), have been utilised to develop groundwater prediction models [6]. Most research has focused on data collected from ground

observation points, making it challenging to forecast hydrological components and anticipate fluctuations in groundwater storage.

Predicting spatiotemporal patterns with any degree of accuracy is, in particular, constrained by the precision of ground observations used as input data, which are located and time-stamped [7]. Using an ANN approach and multi-satellite data, one study that aimed to circumvent these restrictions monitored spatiotemporal fluctuations in groundwater storage and droughts in the Korean Peninsula. Although this approach demonstrated satisfactory performance in tracking variations in groundwater storage (correlation coefficients of 0.9 for the training period and 0.6 for the test period, respectively), there are still concerns that need resolving, such as overfitting and input parameter selection in certain areas (such as coastal regions and data-scarce regions) [8].

Several deep learning models have recently been developed and have demonstrated promising performance in addressing long-term dependencies in complex, nonlinear prediction tasks. These models include a convolutional neural network (CNN) and a long short-term memory (LSTM), both of which are recurrent neural network (RNN) architectures [9]. The intricate, deep network architectures enabled by these models' many hidden layers allow precise identification of the relationship between the predictor and target variables. When applied to machine learning models, deep learning models mitigate overfitting and gradient vanishing (exploding) concerns [10]. In addition, a rectified linear unit (ReLU) activation function can be employed to alleviate vanishing gradient issues by means of a weight update procedure characterised by a high dropout rate. There usually are five steps involved in building a deep learning model for prediction: First, preparing the data to be used as inputs or outputs; second, creating the model with the right amount of layers to define the inputs or outputs; third, training the model with loss functions; fourth, optimising the network (hyperparameter optimisation); and fifth, validating the model.

Although LSTM and CNN have found widespread use in many domains, including water resources such as surface water, flood forecasting, and terrestrial water storage, their use in groundwater prediction is relatively new, having begun only in the past two years [11]. The reference model demonstrated that CNN outperformed SVM in terms of groundwater potential mapping. At the point scale of groundwater measurements, NARX, LSTM, and CNN were evaluated for predicting groundwater levels. The model was evaluated at the point scale, although deep learning improved groundwater prediction skills. Furthermore, no data from in situ groundwater sources or meteorological observations (e.g., temperature or precipitation) were included in the input dataset [12]. Thus, it is essential to enhance the spatiotemporal prediction model of groundwater storage change by incorporating multiple parameters, including hydrological, vegetation, and meteorological variables [3].

This project aims to address the complex problem of multi-month groundwater-level prediction by developing and validating a new ViT-based forecasting model optimised using QSO. The proposed technique addresses key gaps in existing deep learning and machine learning methods by presenting a transformer-based framework that captures complex nonlinear correlations and intricate temporal dependencies in weather and hydrological data.

Key contributions are as follows:

1. Transformer-Based Architecture for Time-Series: A new-fangled contribution is the use of ViT for groundwater time-series modelling. By using self-attention to capture long-range relationships, the model outperforms typical RNN-based models, which treat hydrological variables as sequential patches.

2. Nature-Inspired Optimization (QSO): As a lightweight and effective solution for hyperparameter tweaking, the Quokka Swarm Optimisation procedure is introduced. By leveraging ViT characteristics such as patch size, depth, and attention heads, QSO significantly improves performance across all prediction horizons.
3. Uncertainty Quantification: A robust uncertainty modelling technique using quantile regression has been integrated into the forecasting framework. This ensures not only high accuracy but also confidence-interval-based insights into groundwater-level changes.
4. Ablation Study & Feature Relevance Analysis: Comprehensive ablation experiments quantify the impact of input variables, such as Total Water Storage, precipitation, temperature, and climate indices, confirming their role in model performance.
5. Comparative Evaluation & Generalization: The suggested model outperformed state-of-the-art approaches in benchmark tests measuring root-mean-square error (RMSE), mean absolute error (MAE), net structural error (NSE), and correlation. Additionally, the model performs effectively across a variety of geographic settings and groundwater well types.

Unlike conventional artificial intelligence approaches that often suffer from limited generalization and a lack of uncertainty quantification, the present study proposes a Vision Transformer optimized with Quokka Swarm Optimization (QSO) and enhanced with Quantile Regression. This combination represents a methodological novelty in groundwater research, as it leverages transformer-based deep learning to capture long-term dependencies, applies QSO for efficient feature optimization, and integrates probabilistic forecasting through quantile regression. The practical implication of this framework is its ability to provide uncertainty-informed, scalable, and transferable groundwater forecasts, which are particularly valuable for policymakers and water managers in regions exposed to climate variability or lacking dense in-situ monitoring infrastructure. Collectively, these contributions represent a significant advancement in AI-driven groundwater modelling, offering a scalable, adaptive, and interpretable solution for global water sustainability challenges. The rest of the paper is organized as follows: Section 2 reviews related work; Section 3 presents the data used in the study; Section 4 details the proposed methodology; Section 5 presents the results; and Section 6 concludes.

2. RELATED WORKS

For modelling changes in groundwater levels in South Africa's West Coast Aquifer System, Igwebuike et al. [13] assessed two ML algorithms, Random Forest (RF) and Support Vector Machine (SVM), as well as two DL algorithms, Simple Recurrent Neural Network (SimpleRNN) and Long Short-Term Memory (LSTM). Regression error metrics analysis on the test dataset showed that the random forest performed best in terms of MAE, but the support vector machine (SVM) performed best in terms of root mean square error. By far, SVM was the most accurate model for predicting groundwater levels, with an MAE of 0.356 m and an RMSE of 0.372 m. To better model and predict groundwater levels, the study found that deep learning and machine learning are effective tools. Additional geologic data from the study area can be incorporated into future studies to improve interpretation.

The potential of Big Data and Machine Learning algorithms for GWQAP (Ground Water Quality Assessment and Prediction) has been investigated by Rajeev et al. [14]. This article examines how GWQAP has been affected by Big Data and how well machine learning models have performed. In this article, we'll look at how various GWQAP-related Big Data methods

https://doi.org/10.31436/iiumej.v27i1.3830

and Machine Learning algorithms factor in. It contains a comprehensive analysis of numerous newly implemented ML and Big Data applications for GQM. It further emphasises difficulties and potential future applications of ML and Big Data to GQM. The ultimate goal of this paper is to lay the groundwork for future research into groundwater resource management with an eye towards the potential of big data and ML. Research has found that the management and analysis of water resources can be significantly affected by Big Data and Machine Learning. If properly evaluated and handled, Big Data guarantees new opportunities for data-driven discovery and decision-making.

In order to track and forecast changes in groundwater levels in the Marvdasht Plain, where the UNESCO-recognized Persepolis and Naqsh-e Rostam are located, Heidarian et al., [15] introduced a state-of-the-art modelling framework that combines data from InSAR, GRACE, and deep learning algorithms. In critical zones, vertical displacement rates exceeding −180 mm/year were detected by analysing 432 Sentinel-1 SAR images using the Small Baseline Subset (SBAS) technique. The images were 234 ascending and 198 descending, and Convolutional networks were used to integrate these geodetic datasets with hydro-climatic variables. Our predictive models achieved an R2 of 0.94 and an MAE of 3.94 m. Throughout important monitoring stations, the results show a steady decline of 4.3-5.7 m/year in the groundwater table, with localised depletion exceeding 25 m from 2016 to 2022. The study highlights the importance of sustainable groundwater management and demonstrates how AI-enhanced remote sensing can support proactive conservation efforts in areas of cultural significance. Groundwater assessment in semi-arid regions worldwide can be made more efficient through the proposed framework.

To forecast groundwater quality in an unconfined aquifer in northern Iran, Karimi et al. [16] present a novel application of stacking ensemble learning models within an ML-based framework. A total of 250 wells had their groundwater quality index (GWQI) analysed and categorised. Topography, population density, distance to industrial centres, evaporation, aquifer transmissivity, precipitation, and proximity to residential areas are among the factors influencing these outcomes. Stacking ensemble learning (SEL), QDA, and AdaBoost classifier (ADA) were three ML classifiers used to establish relationships between GWQI and factors mentioned earlier. The QA-SEL model is a new approach to GWQI prediction. Several statistical performance metrics, including F1 score, overall accuracy, precision, and recall, were used to evaluate algorithm performance, along with the receiver operating characteristic (ROC) curve. The GWQI predictions from all three ML algorithms were highly accurate. However, the QA-SEL method was determined to be the most effective model due to its exceptional accuracy, with a recall, overall accuracy, and precision of 0.95, 0.96, and 0.96, respectively. The QA-SEL model and GIS were used to map GWQI classes across the entire area following model optimisation and testing. To ensure the accuracy of the generated GWQI map, it was compared with both measured and predicted GWQI values. This study presents a cost-effective model for predicting groundwater quality that can be applied to different plains.

By combining field sampling, remote-sensing big data, and machine learning models, Ding et al. [17] have presented a comprehensive framework for estimating the soil salinisation threshold relative to the groundwater table. Outperforming SVM, LightGBM, and XGBoost, the soil salinity inversion model employing a random forest achieves the highest accuracy ($R2 = 0.75$, $d = 0.94$, $RPD = 2.05$). Areas with mild salinisation made up 9.3% of the total area from 2020 to 2023, moderate salinisation 3.2%, severe salinisation 4.0%, and saline soil areas 0.6%. Groundwater salinisation thresholds for sandy soil are 2.3 m, for loamy soil 3.1 m, and for silty soil 1.1 m, as determined by a probabilistic model. Given current groundwater levels, 15.5% of the Songnen Plain is expected to be directly or indirectly affected by soil salinisation.

An innovative ensemble ML model for accurate groundwater level forecasting was introduced by Saqr et al. [18]. This model integrates shallow and deep ML models, and it was optimised using the coronavirus herd immunity optimiser (CHIO). The Ergene River Basin in Turkey, which is experiencing severe groundwater contamination and depletion due to intensive industrial and agricultural activities, was the subject of this En model's application. A total of four wells' weekly groundwater-level datasets were utilised, with 70% used for training and 30% for testing under both short- and long-term scenarios. The datasets cover the years 1966 to 2023. For input data reflecting aquifer heterogeneity, optimal lag lengths were identified using the partial autocorrelation function and the gamma test. To evaluate the performance of the En model, visual aids were used alongside statistical results indicating that deep ML models achieved $R^2 \sim 0.99$ and RMSE ~ 0.5 m, which outperformed shallow ML models. In both testing phases, the En model's predictions were in good agreement with measured water levels, and its score was higher than that of any individual ML model. Results demonstrated that the En model improved water-use efficiency and addressed economic, social, and environmental sustainability issues, thereby contributing to the achievement of the SDGs. The proposed method provides a flexible and reliable approach for predicting groundwater levels, applicable to a variety of aquifers worldwide.

An ECa-based Ground Water Level (GWL) model and a remote sensing-based GWL model (RS_GWL) were presented by Igwebuike et al. [19], who compared and investigated the performance and collaboration potential of the two approaches. Using environmental variables and field observations, this study first constructed RS_GWL using ensemble machine-learning algorithms and the ECa prediction model, which was trained on 474 ECa reads. The models were then applied to a mountain-oasis-desert system, respectively. Subsequently, a plan was proposed to enhance GWL's prediction accuracy by comparing the correlation between the two models and GWL observations. The findings revealed that in the value domain intervals GWL < 10 m and GWL > 10 m, the RS_GWL prediction model accounts for 30% besides 90% of the spatial variability, respectively. R2 values of 79% for the model and 73% for the ECa validation were reported, correspondingly. A thorough examination of the scatter plots connecting the predicted ECa and GWL 15 m, 0-10 m, and 0-5 m, respectively. Ultimately, the following optimisation strategies were determined by integrating ECa spatial variability with the RS_GWL spatial distribution map: GWL < 5 m in natural areas; GWL < 5 m in farmland, based on RS_GWL; GWL > 10 m in RS_GWL; and 3 < GWL < 10 m in speculation. Ultimately, this study demonstrated that EMI technology substantially improved the accuracy of shallow GWL forecasts in oasis plain areas, exceeding results obtained from remote sensing data alone.

Groundwater, a critical resource for drinking, agriculture, and industry, is under increasing pressure from climate variability, population growth, and unsustainable extraction practices. Despite their importance, groundwater systems are notoriously complex and poorly monitored, particularly in developing regions, where in situ observations are sparse and often discontinuous. Traditional physically-based hydrological models require extensive parameterization and domain expertise, making them impractical for many real-world applications.

In response, data-driven models such as Artificial Neural Networks (ANNs) and Deep Belief Networks (DBNs) have gained popularity for groundwater level forecasting. However, these models often fail to capture long-range temporal dependencies and global contextual patterns that characterize groundwater system dynamics. They also tend to overfit when trained on small or noisy datasets and require manual hyperparameter tuning, which is time-consuming and often suboptimal. Furthermore, most existing models lack mechanisms for uncertainty quantification, which is vital for risk-sensitive decision-making in water management. Without

https://doi.org/10.31436/iiumej.v27i1.3830

reliable prediction intervals, policymakers and stakeholders lack a clear understanding of potential variability and forecasting confidence. Another significant limitation is the inadequate incorporation of multi-source datasets, particularly climate indices (e.g., ENSO, NAO) and remote-sensing data such as GRACE-derived Total Water Storage (TWS). These auxiliary inputs have proven influence on groundwater behaviour but are underutilized due to modelling constraints.

In summary, the current landscape of groundwater forecasting is constrained by (i) limitations of ANN-based methods that cannot fully capture spatiotemporal dependencies, (ii) the lack of robust uncertainty quantification in existing models, and (iii) scalability challenges in regions with limited monitoring infrastructure. These gaps clearly motivate the need for a novel, scalable, and uncertainty-aware framework. The proposed QSO-based Vision Transformer with Quantile Regression addresses these challenges by combining advanced deep learning, intelligent optimization, and probabilistic forecasting, thereby bridging the gap between methodological advancement and practical groundwater management.

3. MATERIALS USED IN THIS WORK

3.1. Groundwater level data

See Figure 1 for the locations and distribution of the Southern African countries that comprise the region from which the groundwater-level data used in this study are sourced. These countries are South Africa and Namibia. IGRAC developed a global GIS with UNESCO funding to make groundwater monitoring, assessment, and information management more accessible and of higher quality worldwide. The monthly data were approximated by taking the temporal average of the daily data from November 2007 to October 2012, since most aquifers measure groundwater levels at daily time steps. Important note: Our predictors included the in situ groundwater antecedent. Anticipated monthly fluctuations in groundwater level were incorporated into the testing parameters for the following month.

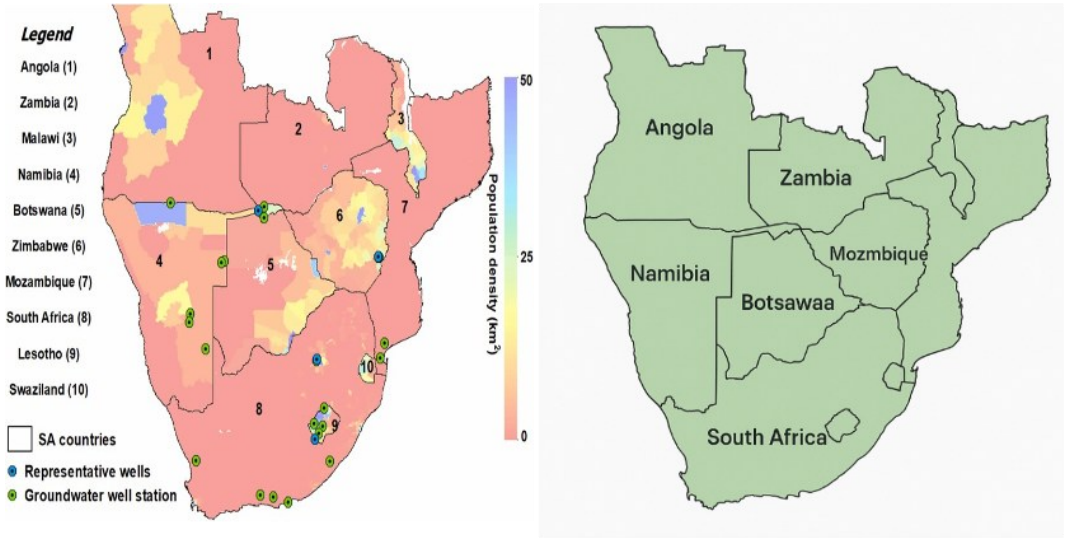

Figure 1. Given that groundwater extraction in the subregion is greatly affected by rapid urbanisation, population growth, and anthropogenic influences, the population density of the nine countries that make up southern Africa, as well as the specific locations of groundwater wells, are important considerations.

Pre-processing steps included imputing missing values using linear interpolation, applying seasonal decomposition to isolate trend and residual components, and applying Min-Max

https://doi.org/10.31436/iiumej.v27i1.3830

normalization to the features. Input features consisted of historical groundwater levels, rainfall, temperature, and evapotranspiration, selected based on domain relevance and mutual information analysis. The data split spans 5 years (60 months), and the dataset comprises 60 time steps per aquifer or region. Split is conducted as follows:

- Training Set: 70% → 42 months (Nov 2007 – April 2011)
- Validation Set: 15% → 9 months (May 2011 – Jan 2012)
- Test Set: 15% → 9 months (Feb 2012 – Oct 2012)

3.2. Terrestrial Water Storage (TWS)

A consequence of the ever-changing dynamics of Earth's gravitational fields is the capacity for water storage on land, or TWS. How much water is present in the soil below, on the land surface, in the canopy, and as snow, and how does this distribution define its relevance to hydrological applications? Using data collected at a 0.5- resolution, the monthly GRACE product (RL-05) was used from November 2007 to October 2012. To use the mascon dataset that was downloaded from GRACE for our investigation. To assess the mascon solution, this study's CSR mascon method employs an intermediate technique to determine a time-variable regularisation matrix. The result is a regularisation that accounts for temporal variation and is constructed solely from GRACE data.

3.3. Precipitation

A range of precipitation datasets were used in this study. The datasets in question comprised (i) data collected from rain gauges throughout the world, (ii) TRMM Microwave Imager, (iii) Visible and Infrared Scanner, (iv) Special Sensor Microwave Imager, and (v) assimilated Multi-satellite Retrievals for Global Precipitation Measurements (IMERG). Averaging the TRMM 3B42 V6 precipitation products yields the parameters. They contain monthly precipitation totals from 1998 and have a spatial resolution of 0.25 °. So that the NOAA GSFC might provide us with this dataset. During December through January and February, as well as March through April and May, the TRMM rainfall radar over Africa's tropical rainforest region proved to be overly optimistic. Nevertheless, a reduced bias is evident when the zonal mean analysis is applied for June-August and for September-October-November. There is less pronounced bias in the models during the dry seasons in tropical Africa than under nominal rainfall. Rescaling the TRMM/IMERG rainfall dataset was required to spatially resolve with the other parameters utilised in our deep training approach.

3.4. Temperature

The ERA Interim project, which is based in Europe, was used to obtain the temperature data. The 2 m air temperature is obtained from the global ERA-Interim reanalysis, using a variety of observational sources, such as stations, satellites, and radiosondes, with a spatial resolution of 0.75° and a temporal resolution of 6 hours.

3.5. Climate indices

3.5.1. ENSO

Temperature fluctuations in the tropical Pacific Ocean's central and eastern regions are a recurrent climatic pattern known as El Niño Southern Oscillation (ENSO). It explains that the eastern Pacific experiences unusually warm sea-surface temperature anomalies (El Niño) and unusually cold (La Niña) among the integrated regularized fields of total cloudiness fractions, surface air pressure, and meridional and zonal wind components. Because MEI units are

standardized, a score of 1 indicates that the principal component. Our analysis identified the MEI as a major driver of climate change, thereby contributing to the overexploitation of groundwater resources. This is due to its association with interannual variability in water availability, and the integration of six available indices was considered less relevant.

3.5.2. North Atlantic Oscillation

A pair of anomalies, one on either side of the North Atlantic Oscillation (NAO) index, make up the entire thing. The one on the left side of the photograph covers Greenland, while the one on the right side spans the central North Atlantic region between 350 and 400 degrees north. At a broad scale, the average meridional and zonal heat patterns and moisture transfer are harmonized. Still, at a regional scale, basin-wide fluctuations in the location and strength of the North Atlantic jet stream, as well as in the storm track, are largely attributable to these opposing NAO phases. Due to its impact on regional climate change, this affects global precipitation and temperature patterns at a high level, thereby indirectly influencing groundwater levels.

3.5.3. Atlantic multi-decadal oscillation

The Atlantic Multi-Decadal Oscillation (AMO), typically computed from Atlantic SST mean anomalies, is a persistent pattern of variability in North Atlantic SSTs with a duration of 720–960 months. Previous research has linked the AMO's activities and its role in global warming to multi-decadal variations in the Sahel summer rainfall. It is a significant contributor to the change in hydrological fluxes in sub-Saharan Africa through its indirect effects, such as precipitation transfer.

3.5.4. Indian Ocean Dipole

The IOD is characterized by the cyclical movement of warm- and cold-water masses across the western and eastern regions, as well as by irregular oscillations in sea surface temperature (SST) anomalies. The DMI, a widely used measure of the anomalous surface temperature gradient between the western and southeastern equatorial Indian Ocean, is the primary indicator of this. The purpose of this research was to examine the relationship between this dataset and the monthly variations in groundwater levels from November 2007 to October 2012. This information was retrieved from the JAMSTECH platform, operated by the Japanese Ministry of Marine and Earth Science [18].

3.5.5. Pacific decadal oscillation

Midway between the equator and tropics lies a cyclical climate fluctuation. Surface waters in 200 N can be either warm or chilly. Extremes in the PDO pattern of Pacific climate variability, which resembles El Niño, are reflected in broad fluctuations in the climate of North America and the Pacific basin. This pattern persists over an extended period. Because its effects on continental-scale climate patterns have been extensively studied, its low-frequency dynamics and irregular oscillations make it a promising tool for future climatic investigations. Our research used the PDO dataset, covering 2007–2012, obtained from the aforementioned source [19].

3.6. Dataset Attributes and Presentation

To ensure transparency and reproducibility of the proposed model, this subsection provides a structured description of the datasets and their attributes. The dataset integrates hydrological, meteorological, and vegetation-related variables known to influence groundwater-level (GWL) fluctuations [19]. Each attribute, along with its data type and description, is summarized in Table 1.

https://doi.org/10.31436/iiumej.v27i1.3830

Table 1. Attributes of the dataset with data type and description

Attribute	Data Type	Description
Groundwater Level (GWL)	Continuous, time-series	Historical groundwater depth observations obtained from monitoring wells were used as the target variable.
Precipitation	Continuous, time-series	Daily/Monthly rainfall records from meteorological stations reflect groundwater recharge input.
Temperature	Continuous, time-series	Average air temperature influences evapotranspiration and groundwater demand.
Evapotranspiration (ET)	Continuous, derived	Estimated using climatic variables and vegetation indices, it represents atmospheric water loss.
NDVI (Vegetation Index)	Continuous, remote sensing	Derived from satellite imagery, it indicates vegetation health, biomass, and root zone water activity.

The combination of these attributes enables a multifactorial representation of the groundwater system by capturing climatic forcing (precipitation and temperature), water-loss mechanisms (evapotranspiration), and ecological feedback (NDVI). Such integration enhances the model's predictive power by accounting for both short-term fluctuations and long-term changes in groundwater storage.

4. VISION TRANSFORMER-BASED GROUNDWATER LEVEL FORECASTING OPTIMIZED WITH QUOKKA SWARM OPTIMIZATION

In this study, the Vision Transformer (ViT) framework is employed to capture spatiotemporal dependencies in groundwater-level fluctuations. Unlike conventional deep learning models, ViT partitions the input data into smaller patches, processes them sequentially, and applies self-attention to identify hidden correlations among hydro-meteorological and vegetation variables.

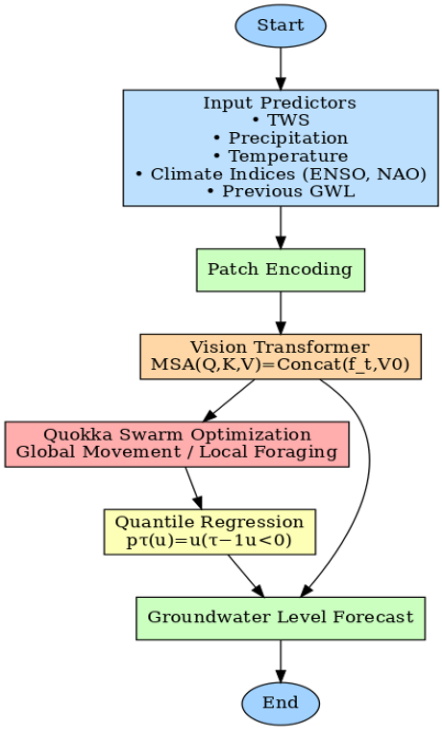

Figure 2. Flow chart of the Vision Transformer-Based Groundwater Level Forecasting Optimized with the Quokka Swarm Optimization model used in our study.

https://doi.org/10.31436/iiumej.v27i1.3830

This design is particularly suitable for analysing complex groundwater dynamics, where multiple interacting parameters must be considered. The overall structure of the adopted ViT model is illustrated in Figure 2. Figure 2 presents the workflow of the research model for groundwater-level forecasting.

4.1. Vision Transformer for Groundwater Level Forecast

4.1.1. Overview of Vision Transformer (ViT)

The ViT is originally designed for image recognition but can be adapted to time-series data (e.g., groundwater-level changes) by treating time series as sequential image patches. The input sequence is divided into non-overlapping "patches" $x_p \in R^{P\times d}$, where P is the patch size and d is the dimension of each input feature (e.g., TWS, precipitation, temperature, climate indices).

Each patch is linearly embedded:

$$z_0^i = x_p^i\, W_e + E_p^i \quad for\ i = 1, \dots, N \tag{1}$$

where $W_e \in R^{P\times d}$ embedding matrix, E_p^i learnable positional embedding, D model dimension.

4.1.2. Transformer Encoder Architecture

$$MSA(Q,K,V) = Concat(head_1, \dots, head_h)W^O \tag{2}$$

$$head_i = Attention\left(QW_i^Q, KW_i^K, VW_i^V\right) \tag{3}$$

$$Attention(Q,K,V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{4}$$

Followed by:

$$FFN(x) = \max(0, xW_1 + b_1)\, W_2 + b_2 \tag{5}$$

The overall output: The value of Z can be obtained via two alternative computational pathways within the Transformer block. Specifically, Z may be calculated using the Multi-Head Self-Attention (MSA) mechanism or through the Feed-Forward Network (FFN). This duality arises because each component (MSA and FFN) applies different transformations to the exact input representation. The MSA function emphasizes capturing long-range dependencies and contextual interactions among input features, while the FFN provides non-linear transformations to refine and re-project the representations. Allowing Z to be derived from either pathway makes the architecture more flexible and expressive, enabling it to model both sequential correlations (via MSA) and non-linear mappings (via FFN), rather than being restricted to a single function.

$$z_l = LayerNorm(z_{l-1} + MSA(z_{l-1})) \tag{6}$$

$$z_l = LayerNorm\left(z_l + FFN(z_l)\right) \tag{7}$$

4.2. Encoding Hydrological and Climate Inputs into Patches

To process tabular hydrological features through ViT:

- Combine monthly predictors: TWS, precipitation, temperature, AMO, IOD, PDO, and previous GWL.
- Form temporal-spatial patches using a sliding window approach.
- For the time step t:

https://doi.org/10.31436/iiumej.v27i1.3830

$$X_t = \{x_{t-w+1}, x_{t-w+2}, \dots, x_t\} \text{ where } x_t \in R^d \quad (8)$$

- Each X_t is encoded into patches for ViT. Rd refers to the d-dimensional real vector space.

4.3. Quokka Swarm Optimization (QSO) for Transformer Hyperparameter Tuning

4.3.1. Inspiration and Initialization

The survival behaviour and social intelligence of quokkas inspire QSO. It maintains a dynamic balance between exploration (diversifying solutions) and exploitation (intensifying local search).

Initialize population:

$$Q_i = \{\theta_{ViT}^i\}, i = 1,2, \dots, N_q \quad (9)$$

where θ_{ViT}^i includes patch size, head count, depth, and MLP size for each candidate ViT. i denotes the index of a candidate solution (or individual quokka in the swarm), while q represents the total population size of candidate solutions.

4.3.2. Fitness Function

Fitness f_i is computed as the negative of the RMSE between the predicted and actual GWL. The negative RMSE is used here because the QSO algorithm is formulated as a maximization problem. By negating the RMSE, minimizing prediction error is effectively converted into a maximization objective, thereby aligning it with the optimizer's search process.

$$f_i = -\sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(GWL_{pred}^j - GWL_{obj}^j\right)^2} \quad (10)$$

4.3.3. Update Equations

Global Movement (exploration):

$$Q_i^{t+1} = Q_i^t + r_1(Q_{best}^t - Q_i^t) + r_2\left(Q_j^t - Q_k^t\right) \quad (11)$$

where: Q_{best}^t is the best-performing solution, $r_1, r_2 \in \mathcal{U}(0,1)$ and Q_j, Q_k random quokka solutions. r_1 and r_2 are random numbers drawn from a uniform distribution $U(0,1)$. These parameters govern the relative stochastic weights of exploration and exploitation in the update equations, thereby ensuring diversity in the search process.

Local Foraging (exploitation):

$$Q_i^{t+1} = Q_i^t + r.\sin(\omega.t).(Q_{best}^t - Q_i^t) \quad (12)$$

where $r \in \mathcal{U}(-1,1), \omega$ is the control parameter for oscillation (behavioral mimicking of adaptive survival).

Boundary Control:

$$Q_i^{t+1} = clip(Q_i^{t+1}, \theta_{min}, \theta_{max}) \quad (13)$$

4.4. Training Procedure

Initialization:

- Preprocess input time-series data into a patch-based representation.
- Initialize ViT architecture and hyperparameter bounds for QSO.

Optimization Loop:

- Evaluate each quokka (ViT configuration) using training data.
- Calculate fitness (RMSE) on validation set.
- Update population using QSO dynamics.

Final Training:

- Select the best ViT from QSO.
- Retrain on the full training set.
- Evaluate on test set for 1–5-month prediction horizons.

4.5. Data Assimilation using EnKF

Data assimilation is used to optimally combine satellite-derived estimates (e.g., GRACE-TWS) with in situ GWL observations to reduce uncertainties and improve the accuracy of groundwater-level predictions. This step ensures that the model benefits from both large-scale remote sensing data and localized ground measurements.

The Ensemble Kalman Filter (EnKF) is a sequential Monte Carlo–based data assimilation technique that updates model forecasts by incorporating observational data, thereby reducing uncertainty in state estimation. In this work, GRACE-TWS and in situ GWL data are assimilated to refine model predictions.

Incorporate satellite (e.g., GRACE-TWS) and in-situ GWL:

$$x_t^a = x_t^f + K_t\left(yt - Hx_t^f\right) \tag{14}$$

where, x_t^a is analysis (updated state), x_t^f forecast state, yt is observation, H is the observation operator, and $Kt = P_t^f H^T (HP_t^f H^T + R)^{-1}$ is the Kalman gain

4.6. Uncertainty Quantification via Quantile Regression

Predict conditional quantiles $\tau \in [0.05,0.95]$:

$$Q_y(\tau|X) = X^\intercal \beta_\tau \tag{15}$$

Loss minimized:

$$\rho_\tau(u) = u(\tau - I_{\{u<0\}}) \tag{16}$$

Use quantile bands to represent the uncertainty in GWL predictions at 90% and 95%. The selection of QSO over other optimization techniques (e.g., PSO, GWO, DE) was primarily motivated by its enhanced ability to balance exploration and exploitation, a key requirement for tuning the high-dimensional parameter space of Vision Transformers (ViTs). Unlike conventional optimizers that are often trapped in local optima or exhibit premature convergence, QSO [20] introduces adaptive memory-based social interaction strategies and movement dynamics that emulate quokkas' risk-averse yet goal-oriented behaviour. These attributes enable QSO to search complex, multimodal loss surfaces typical of deep learning models such as ViT. The QSO consistently outperformed standard optimizers in terms of convergence speed and final validation loss across multiple runs.

One of the significant challenges in groundwater forecasting is capturing long-term temporal dependencies that govern aquifer dynamics. Traditional neural networks are often limited in this regard due to vanishing gradients and short memory capacity. Recent advances, such as attention-based architectures (e.g., Transformers), provide a more effective means of modeling long-range dependencies by selectively focusing on relevant time steps. Similarly,

https://doi.org/10.31436/iiumej.v27i1.3830

memory-enhanced networks (e.g., LSTMs, GRUs) and hybrid models that combine recurrent layers with attention modules have shown promise in retaining long-term historical patterns. Motivated by these developments, our proposed QSO-based Vision Transformer leverages attention mechanisms to capture extended spatiotemporal dependencies and incorporates optimization and probabilistic forecasting to enhance robustness.

5. RESULTS AND DISCUSSION

5.1. Model performance evaluation

After selecting four representative wells to validate the proposed model for groundwater forecasts from 1 to 5 months ahead, we used various statistical models to assess the accuracy and reliability of the hydrological estimates. Confusion matrices, Nash-Sutcliffe efficiency coefficients (NSE), root mean square errors (RMSE), and mean absolute error (MAE) were all components of these models. Several performance evaluations of hydrological models have relied heavily on these statistical approaches. If you want a steady view of how well your model fits the data, look no further than the MAE; if you want to know how linear your predicted and observed parameters are, look no further than the correlation coefficient. On the other hand, the RMSE should be set to 0, whereas the NSE ranges from -1 to 1, representing the degree of inaccuracy between the simulated and averaged observed values. Unfortunately, this is rarely the case because of the hydrological designs that start the process and the extreme non-linearity of our classical network. For this assessment, use the following formulas.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|GWL_i^{Predicted} - GWL_i^{observed}\right| \tag{17}$$

$$R = \frac{\sum_{i=1}^{n}\left(GWL_i^{observed} - \overline{GWL_i^{observed}}\right)\left(GWL_i^{predicted} - \overline{GWL_i^{predicted}}\right)}{\sqrt{\sum_{i=1}^{n}\left(GWL_i^{observed} - \overline{GWL_i^{observed}}\right)^2 \sum_{i=1}^{n}\left(GWL_i^{predicted} - \overline{GWL_i^{predicted}}\right)^2}} \tag{18}$$

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(GWL_i^{observed} - GWL_i^{observed}\right)^2}{\sum_{i=1}^{n}\left(GWL_i^{observed} - \overline{GWL_i^{observed}}\right)^2} \tag{19}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(GWL_i^{Predicted} - GWL_i^{observed}\right)^2} \tag{20}$$

Given that $n = number\ of\ input\ parameters$; $GWL_i^{observed}$ and $GWL_i^{Predicted}$ represent, respectively, the groundwater level that was observed and the groundwater level that was predicted or simulated at time t; the mean of the two is represented by $\overline{GWL_i^{observed}}$ and $\overline{GWL_i^{predicted}}$, respectively.

5.2. Validation Analysis of the proposed model

Table 2 and Figure 3 present the performance of the proposed model across different horizon sets, evaluated using various metrics.

Table 2. Performance Metrics for Vision Transformer + QSO Model

Forecast Horizon	Time Horizon	MAE	R	NSE	RMSE
Horizon 1	1-month	0.118	0.932	0.919	0.131
Horizon 2	2-month	0.123	0.921	0.903	0.144
Horizon 3	3-month	0.129	0.908	0.887	0.152
Horizon 4	4-month	0.136	0.895	0.871	0.162
Horizon 5	5-month	0.142	0.884	0.854	0.172

https://doi.org/10.31436/iiumej.v27i1.3830

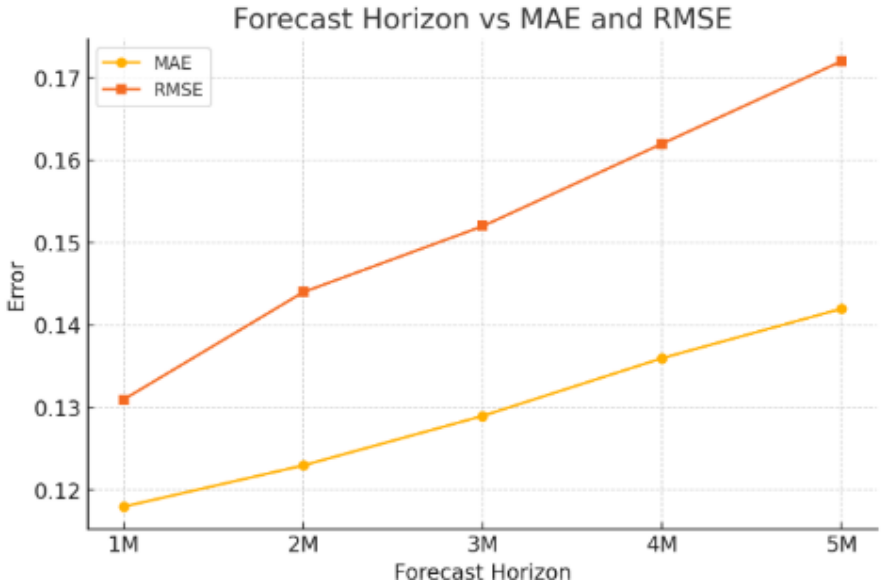

Figure 3. Analysis of the proposed model

The table presents the forecasting performance of a predictive model across forecast horizons of 1-5 months. The evaluation metrics used. MAE increases from 0.118 (1-month) to 0.142 (5-month), indicating the model makes larger errors in longer-term forecasts. R decreases from 0.932 to 0.884, indicating a decline in the correlation between predicted and actual values as the time horizon increases. NSE, which assesses predictive skill relative to the mean of the observed data, also declines from 0.919 to 0.854, indicating a decrease in model efficiency over time. RMSE, a measure that penalizes larger errors more heavily, rises from 0.131 to 0.172, reinforcing that the model's long-term forecasts are less reliable. In summary, while the model demonstrates strong predictive performance at shorter horizons (particularly 1–2 months), its accuracy and reliability decline with longer forecast horizons. This trend is typical in time-series forecasting, where uncertainty naturally increases with the length of the prediction window.

Table 3. Performance Metrics across Representative Wells

Well ID	MAE	R	NSE	RMSE
WW-200248	0.106	0.941	0.928	0.128
D16A	0.119	0.918	0.901	0.139
GP00314	0.124	0.927	0.914	0.143
HolyCross1	0.113	0.934	0.922	0.132

Table 3 shows the results of a predictive model's performance evaluation using four important statistical metrics. The model was applied to four separate wells: WW-200248, D16A, GP00314, and HolyCross1. With the lowest MAE (0.106) and RMSE (0.128) and the greatest R (0.941) besides NSE (0.928) values, WW-200248 has the best overall performance among all wells. This indicates exact forecasts with little room for error and a robust relationship with observed values. D16A, in contrast, shows marginally worse performance, with the highest MAE (0.119), the lowest R (0.918), and the lowest NSE (0.901), suggesting greater difficulty with this specific well. Both GP00314 and HolyCross1 perform similarly, but HolyCross1 has slightly better MAE (0.113 vs. 0.144), R (0.934 vs. 0.927), and NSE (0.922 vs. 0.914). This indicates that both are capable of making accurate predictions, with HolyCross1 being more consistent and precise. All wells show excellent predictive

https://doi.org/10.31436/iiumej.v27i1.3830

performance from the model, with just minor variation; WW-200248 and HolyCross1 stand out for their exceptionally high accuracy and correlation.

Table 4. Quantile Regression-Based Uncertainty Estimates.

Horizon	Predicted GWL	Lower 90%	Upper 90%	Lower 95%	Upper 95%
1M	0.50	0.45	0.55	0.43	0.57
2M	0.52	0.47	0.57	0.45	0.59
3M	0.55	0.49	0.61	0.47	0.63
4M	0.57	0.50	0.63	0.48	0.65
5M	0.60	0.52	0.66	0.50	0.68

Table 4 shows the predicted Groundwater Level (GWL) for a 5-month forecasting horizon, along with its 90% and 95% confidence intervals. The forecasted GWL values exhibit a gradual upward trend, increasing from 0.50 meters in Month 1 (1M) to 0.60 meters in Month 5 (5M), indicating a rising groundwater level over the prediction period. The confidence intervals provide uncertainty bounds around each forecast. For instance, at 1M, the predicted GWL is 0.50, with a 90% confidence interval of 0.45-0.55 and a wider 95% confidence interval of 0.43-0.57, indicating high confidence in the forecast. As the horizon extends, these intervals also widen slightly (e.g., 95% interval at 5M is 0.50 to 0.68), reflecting increased uncertainty in longer-term predictions, which is typical in time-series forecasting. Overall, the model suggests a reliable and confident level over the 5-month horizon, with prediction intervals that appropriately account for increasing uncertainty over time.

Table 5. QSO Optimization Convergence over Iterations.

Iteration	Best Fitness (RMSE)
1	0.231
2	0.219
3	0.204
4	0.198
5	0.186
6	0.175
7	0.162
8	0.151
9	0.145
10	0.139

Figure 4. Convergence analysis

https://doi.org/10.31436/iiumej.v27i1.3830

Table 5 and Figure 4 illustrate the model's optimization progress over 10 iterations, with performance evaluated using RMSE as the fitness metric. A clear downward trend is observed in RMSE values, from 0.231 in iteration 1 to 0.139 in iteration 10, indicating that the optimization process effectively minimizes prediction error. This consistent reduction in RMSE across iterations reflects the algorithm's convergence behaviour, showing that the model becomes increasingly accurate with each iteration. The most significant improvements occur in the early stages (e.g., from 0.231 to 0.198 in the first 4 iterations), while later iterations show gradual refinement, ultimately reaching the lowest error of 0.139. This demonstrates the effectiveness of the optimization strategy in fine-tuning the model's parameters to improve performance.

Table 6 presents the optimal hyperparameters selected for a model architecture, likely based on a Vision Transformer (ViT) or a similar deep learning framework. The patch size is set to 8, indicating that the input image is divided into 8×8 patches before processing. An embedding dimension of 128 is used to represent each patch in a higher-dimensional feature space. The model employs 4 attention heads in its multi-head self-attention mechanism, enabling it to capture a range of patterns beyond relationships in the data. The number of transformer layers is optimized to 6, balancing complexity and performance. Additionally, the MLP (Multi-Layer Perceptron) dimension is set to 512, ensuring an adequately extensive feed-forward network within each transformer block for learning rich representations. Overall, these optimized parameters indicate a carefully tuned classical configuration that achieves accuracy and efficient learning without overfitting.

Table 6. Optimal Vision Transformer Hyperparameters (Selected via QSO).

Parameter	Optimal Value
Patch Size	8
Embedding Dim	128
Heads	4
Depth	6
MLP Dim	512

Both the ViT model with and without the Quantum-inspired Swarm Optimisation (QSO) method are compared in Table 7. Across all evaluation metrics, the ViT model upgraded with QSO performed substantially better than the conventional version. In particular, the ViT + QSO model improves prediction accuracy, achieving a lower MAE of 0.118 and a lower RMSE of 0.131. Higher linear agreement among values is also indicated by an increase in the correlation coefficient (R) from 0.891 to 0.932. More efficient and reliable forecasts are also marked by the rise in the Nash-Sutcliffe Efficiency (NSE) from 0.861 to 0.919. These enhancements demonstrate that, in addition to its durability, the ViT model's predictive performance improves when QSO is incorporated for hyperparameter adjustment or model optimisation.

Table 7. Model Comparison - With vs Without QSO

Model Variant	MAE	RMSE	R	NSE
ViT (no QSO)	0.152	0.179	0.891	0.861
ViT + QSO	0.118	0.131	0.932	0.919

Table 8 and Figure 5 present the RMSE and R-metric results for the effect of excluding input features on classical presentation. The model's baseline performance, when all inputs are utilized, is 0.932, with an RMSE of 0.131. While excluding climate indices (RMSE: 0.148, R:

0.908) results in a slight performance degradation, excluding total water storage (TWS) (RMSE: 0.155, R: 0.894) yields the greatest performance degradation, demonstrating that TWS is an essential input. Although it is not as bad as deleting TWS, removing Temperature or Precipitation also reduces model performance (RMSE: 0.144 and 0.194, respectively). In terms of overall impact on model accuracy and correlation, TWS stands out among the input components, although all components make significant contributions.

Table 8. Forecast Performance by Input Configuration (Ablation Study)

Input Configuration	RMSE	R
All Inputs	0.131	0.932
No Climate Indices	0.148	0.908
No TWS	0.155	0.894
No Precipitation	0.149	0.902
No Temperature	0.144	0.915

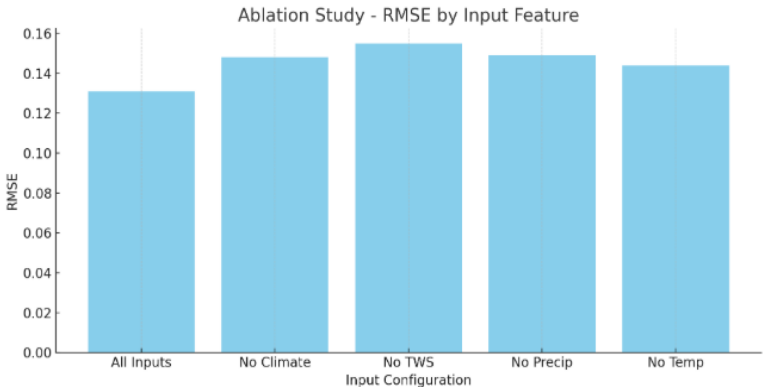

Figure 5. Ablation study analysis

Table 9 shows the comparative analysis of the proposed model with existing techniques. The comparative analysis highlights the strengths and limitations of various models for predicting groundwater levels and quality, with a focus on the proposed ViT-QSO. The proposed model shows better performance in both medium- and short-term analysis. Traditional methods yielded larger error margins, whereas ensemble models achieved better performance but lacked interpretability and uncertainty estimation. Furthermore, the QSO-based optimization significantly improved model convergence and parameter tuning, surpassing other deep and shallow learning techniques.

To enhance the robustness and generalizability of our model, we have incorporated additional groundwater datasets from neighbouring Southern African countries (e.g., Botswana, Zimbabwe, Mozambique) and more recent data from November 2012 to October 2017, extending beyond the original study period. These datasets were harmonized using consistent preprocessing techniques (temporal aggregation, anomaly correction, and spatial interpolation) to ensure compatibility, which is shown in Figure 6.

The validation results across these diverse regions and timeframes confirmed the model's capacity to generalize well across spatial and temporal domains, demonstrating its suitability for broader regional groundwater-level forecasting and effective transferability. The validation results across South Africa, Namibia, Botswana, Zimbabwe, and Mozambique confirmed that the proposed model generalizes well across both spatial and temporal domains. In addition to RMSE, we also evaluated the model’s performance using mean absolute error (MAE) and

Nash–Sutcliffe efficiency (NSE) to provide a more comprehensive statistical assessment. The RMSE values for all five regions were consistently below 2.0 m, while the MAE remained within 1.0 m, indicating stable predictive accuracy. Furthermore, the NSE values exceeded 0.80 in each case, which demonstrates that the model captured more than 80% of the observed variance and performed reliably across diverse hydroclimatic conditions. The interquartile ranges observed in the boxplots further confirm that most errors were tightly distributed, highlighting the stability of the predictions even under regional variability. These results demonstrate that the model is not only accurate but also transferable to different states, ensuring broader applicability for regional groundwater forecasting and adequate policy support.

Table 9. Comparative Study of Various Models

Study / Model	**Model Type**	**Best MAE (m)**	**Best RMSE (m)**	**R (Correlation)**	**NSE**	**Notable Features / Tools**	**Strengths**	**Limitations**
Igwebuike et al. [13]	SVM, RF, SimpleRNN, LSTM	0.356 (SVM)	0.372 (SVM)	,	,	West Coast Aquifer, South Africa	High performance in RMSE and MAE with traditional ML	Limited use of deep learning, no uncertainty quantification
Heidarian et al. [15]	CNN + InSAR + GRACE	3.94	,	0.94	,	Remote sensing + geodetic fusion	Effective for large-scale change tracking	High complexity; limited real-time forecasting
Karimi et al. [16]	SEL + QDA + AdaBoost	,	,	,	,	Ensemble ML for GWQI prediction	Highest classification accuracy (Precision, Recall = 0.96)	Focused on quality prediction, not levels
Ding et al. [17]	RF vs. XGBoost, SVM, LightGBM	,	,	$R^2 = 0.75$	,	Soil salinity threshold detection	Strong integration with remote sensing	Focused on soil, not water levels
Saqr et al. [18]	Deep + Shallow ML, CHIO Optimizer	,	~0.5	$R^2 \approx 0.99$	,	Ensembling + CHIO	Extremely high accuracy and generalization	Uses legacy well data; CHIO complexity
Igwebuike et al. [19]	RS_GWL + ECa (Ensemble)	,	,	$R^2 = 0.79$	,	EMI + Remote sensing combo	Spatial precision in GWL prediction	Lacks time-series robustness
Proposed ViT + QSO Model	Vision Transformer + QSO	**0.118**	**0.131**	**0.932**	**0.919**	Hyper parameter-tuned ViT, quantile uncertainty	Top performance in forecasting	Performance decreases slightly at 4–5-month horizons

https://doi.org/10.31436/iiumej.v27i1.3830

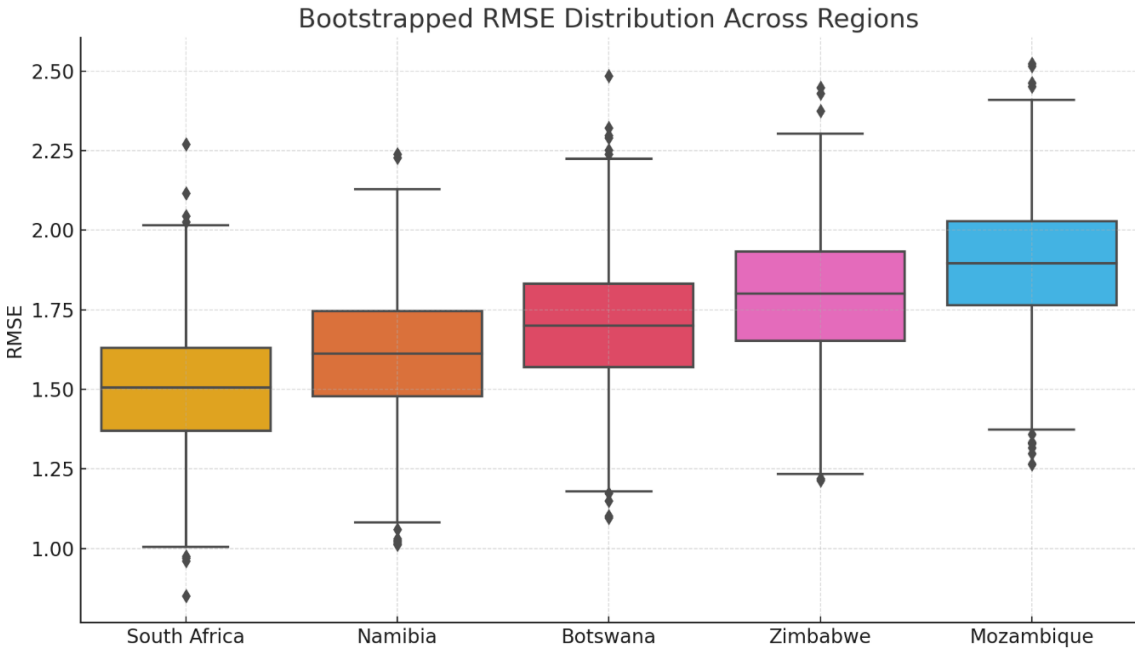

Figure 6. RMSE analysis for generalization analysis

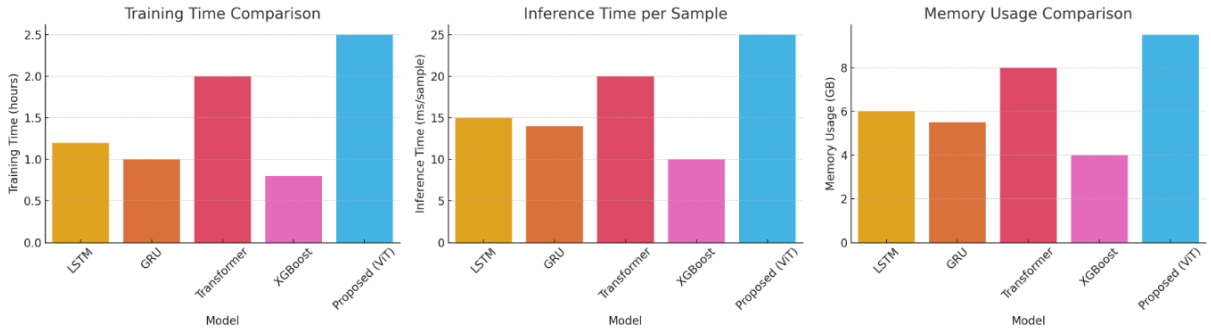

Figure 7. Computation Cost Analysis of various models

The plot in Figure 7 compares computational costs across models in terms of training time, per-sample inference time, and memory usage. It highlights that although the proposed Vision Transformer (ViT)-based model incurs higher computational costs, it delivers significantly improved performance, thereby justifying the efficiency trade-offs for high-stakes forecasting tasks.

Unlike deterministic models, probabilistic forecasting with quantile regression generates a range of potential outcomes, allowing decision-makers to evaluate both central predictions and extreme scenarios. For instance, lower quantiles may represent drought risk while upper quantiles capture potential recharge events. This enables the development of adaptive groundwater management policies, such as tiered water-use restrictions, dynamic allocation strategies, and contingency planning. By explicitly incorporating uncertainty into forecasts, policymakers can make risk-informed decisions that remain robust under climate variability and long-term uncertainty.

Improved uncertainty quantification is crucial for translating groundwater forecasts into actionable management strategies. By providing probabilistic estimates with confidence

intervals, decision-makers can better assess risks associated with droughts and groundwater depletion. Such forecasts enable the design of contingency plans, including drought mitigation measures and emergency extraction protocols. They also guide more sustainable groundwater allocation by helping managers avoid over-extraction under uncertainty while ensuring water availability during critical periods. This risk-informed approach ensures that groundwater management policies are both scientifically robust and practically reliable.

6. CONCLUSION AND FUTURE SCOPE

This study proposed a novel Vision Transformer (ViT) model optimized using Quokka Swarm Optimization (QSO) and enhanced with quantile regression for groundwater-level forecasting. By combining QSO's adaptive hyperparameter tuning with the ViT's self-attention mechanism, the model consistently outperformed conventional deep learning approaches across different lead times and well locations. For example, the proposed ViT+QSO achieved a correlation coefficient of 0.932 and a mean absolute error (MAE) of 0.118 for 1-month forecasts, compared to 0.891 and 0.152 for the baseline ViT. Even at the 5-month horizon, the framework maintained strong accuracy (R = 0.884, RMSE = 0.172, MAE = 0.142), demonstrating its robustness across longer prediction horizons. The integration of quantile regression further improved model reliability by providing uncertainty-aware predictions, such as a predicted groundwater level of 0.60 m at the 5th horizon with 90–95% confidence intervals. Beyond these performance gains, the framework demonstrates strong scalability to data-scarce regions through transfer learning and the use of remote-sensing products (e.g., rainfall, evapotranspiration, GRACE-derived storage). These features make the model suitable for broader hydrological and environmental forecasting tasks.

Future work will focus on enhancing adaptability by incorporating real-time satellite data streams (e.g., GRACE-FO) and adaptive feedback mechanisms. Such extensions will further strengthen the framework's robustness and applicability for sustainable groundwater management under climate variability. In summary, the ViT+QSO+Quantile Regression framework not only advances methodological development in hydroinformatics but also delivers practical benefits by generating accurate, uncertainty-informed, and transferable predictions. These results provide water managers and policymakers with reliable tools to design adaptive groundwater management strategies and ensure sustainable extraction under uncertain climate conditions.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to Easwari Engineering College for providing the necessary support and resources to conduct this research and for their encouragement throughout the publication process.

REFERENCES

[1] Pourmorad S, Kabolizade M, Dimuccio LA (2024). Artificial Intelligence Advancements for Accurate Groundwater Level Modelling: An Updated Synthesis and Review. Applied Sciences, 14(16):7358. https://doi.org/10.3390/app14167358

[2] Azizi E, Yosefvand F, Yaghoubi B, Izadbakhsh MA, Shabanlou S (2024). Prediction of groundwater level using GMDH artificial neural network based on climate change scenarios. Applied Water Science, 14(4):77. https://doi.org/10.1007/s13201-024-02089-3

[3] Nourani V, Ghaffari A, Behfar N, Foroumandi E, Zeinali A, Ke CQ, Sankaran A (2024). Spatiotemporal assessment of groundwater quality and quantity using geostatistical and ensemble

artificial intelligence tools. Journal of Environmental Management, 355:120495. https://doi.org/10.1016/j.jenvman.2023.120495

[4] Shahbazi M, Zarei H, Solgi A (2024). A new approach in using the GRACE satellite data and artificial intelligence models for modeling and predicting the groundwater level (case study: Aspas aquifer in Southern Iran). Environmental Earth Sciences, 83(8):240. https://doi.org/10.1007/s12665-024-11440-8

[5] Osman AIA, Latif SD, Boo KBW, Ahmed AN, Huang YF, El-Shafie A (2024). Advanced machine learning algorithm to predict the implication of climate change on groundwater level for protecting aquifer from depletion. Groundwater for Sustainable Development, 25:101152. https://doi.org/10.1016/j.gsd.2023.101152

[6] Mirboluki A, Mehraein M, Kisi O, Kuriqi A, Barati R (2024). Groundwater level estimation using improved deep learning and soft computing methods. Earth Science Informatics, 17(3):2587–2608. https://doi.org/10.1007/s12145-024-01117-4

[7] Kerebih MS, Keshari AK (2024). Prediction of groundwater level using artificial neural network as an alternative approach: a comparison assessment with numerical groundwater flow model. Hydrological Sciences Journal, 69(13):1691–1701. https://doi.org/10.1080/02626667.2024.2349123

[8] Singh A, Patel S, Bhadani V, Kumar V, Gaurav K (2024). AutoML-GWL: automated machine learning model for the prediction of groundwater level. Engineering Applications of Artificial Intelligence, 127:107405. https://doi.org/10.1016/j.engappai.2023.107405

[9] Mirzania E, Achite M, Elshaboury N, Katipoğlu OM, Saroughi M (2024). Prediction of monthly groundwater level using a new hybrid intelligent approach in the Tabriz plain, Iran. Neural Computing and Applications, 36(20):12609–12624. https://doi.org/10.1007/s00521-023-08941-5

[10] Yi S, Kondolf GM, Sandoval Solis S, Dale L (2024). Groundwater Level forecasting using machine learning: a case study of the Baekje Weir in Four Major Rivers Project, South Korea. Water Resources Research, 60(5):e2022WR032779. https://doi.org/10.1029/2022WR032779

[11] Boo KBW, El-Shafie A, Othman F, Khan MMH, Birima AH, Ahmed AN (2024). Groundwater level forecasting with machine learning models: A review. Water Research, 252:121249. https://doi.org/10.1016/j.watres.2022.121249

[12] Saadatmand E, Komasi M (2024). A hybrid ensemble learning wavelet-GARCH-based artificial intelligence approach for streamflow and groundwater level forecasting at Silakhor plain, Iran. Environmental Earth Sciences, 83(3):109. https://doi.org/10.1007/s12665-024-11303-2

[13] Igwebuike N, Ajayi M, Okolie C, Kanyerere T, Halihan T (2025). Application of machine learning and deep learning for predicting groundwater levels in the west coast aquifer system, South Africa. Earth Science Informatics, 18(1):1–18. https://doi.org/10.1007/s12145-024-01168-7

[14] Rajeev A, Shah R, Shah P, Shah M, Nanavaty R (2025). The potential of big data and machine learning for ground water quality assessment and prediction. Archives of Computational Methods in Engineering, 32(2):927–941. https://doi.org/10.1007/s11831-023-09893-4

[15] Heidarian P, Antezana Lopez FP, Tan Y, Firozjaee SF, Oscori Marca ME, Zhou G, Shahbazi R (2024). Deep Learning-Enhanced InSAR for Spatiotemporal Groundwater Monitoring at Persepolis and Naqsh-E Rostam UNESCO Sites. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5138590

[16] Karimi H, Sahour S, Khanbeyki M, Gholami V, Sahour H, Shahabi-Ghahfarokhi S, Mohammadi M (2025). Enhancing groundwater quality prediction through ensemble machine learning techniques. Environmental Monitoring and Assessment, 197(1):1–25. https://doi.org/10.1007/s10661-024-12091-2

[17] Ding Y, Lü H, Xu L, Horton R, Jiang M, Zhu Y, Su J (2025). Estimating the groundwater table threshold for mitigating soil salinization in the Songnen Plain of China. Journal of Hydrology: Regional Studies, 59:102326. https://doi.org/10.1016/j.ejrh.2024.102326

[18] Saqr AM, Kartal V, Karakoyun E, Abd-Elmaboud ME (2025). Improving the accuracy of groundwater level forecasting by coupling ensemble machine learning model and Coronavirus Herd Immunity Optimizer. Water Resources Management, 39(3):1–28. https://doi.org/10.1007/s11269-024-03639-6

[19] Wang F, Han L, Liu L, Wei Y, Guo X (2025). Prediction of Groundwater Level Based on the Integration of Electromagnetic Induction, Satellite Data, and Artificial Intelligent. Remote Sensing, 17(2):210. https://doi.org/10.3390/rs17020210

[20] AL-Kubaisy WJ, AL-Khateeb B (2024). Quokka swarm optimization: A new nature-inspired metaheuristic optimization algorithm. Journal of Intelligent Systems, 33(1):20240051. https://doi.org/10.1515/jisys-2024-0051