Grenze International Journal of Engineering and Technology, June Issue

Hybrid Approaches in Time Series Forecasting: Integrating Statistical and Machine Learning Models

Dr. Annapoorna B R[1], Pranav J S[2], Prudhvi Raj R[3], Lohish Vinayak Yadav K[4] and Pranav Arya S[5]
[1]Assistant Professor, Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore-560078, India
Email: annapoorna-cs@dayanandasagar.edu
[2-5]Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore-560078, India
Email: jspranav15@gmail.com, pruthvirajendra.314@gmail.com, lohishvinayakyadavk@gmail.com, aryat150802@gmail.com

***Abstract*—In the face of accelerating climate change, precise forecasting of carbon dioxide (CO_2) concentration thresholds has become critical for global policy and planning. This study explores a data-driven approach to modelling future CO_2 trajectories using advanced time-series prediction architectures, including LSTM, GRU, CNN-LSTM, and DNN models. Based on historical emission records, the modelling framework simulates the evolution of global CO_2 levels and estimates 2047 as a potential overshoot point where atmospheric concentrations may surpass 500 ppm—a threshold associated with irreversible environmental damage. Comparative performance across neural architectures highlights the DNN configuration as yielding the most consistent long-range forecasting outcomes. Analysis further indicates that a sustained 6.37% annual emission reduction would be necessary to avert this trajectory, while a 23.38% reversal rate could return concentrations to safer levels. The forecasting framework offers actionable insights for sustainability planning and early-warning environmental monitoring.**

***Index Terms*— Modelling, LSTM, GRU, CNN-LSTM, DNN.**

I. INTRODUCTION

The rapid increase in atmospheric carbon dioxide (CO2) is becoming a serious threat to global environmental stability. As emissions continue to rise, so does the risk of reaching critical climate tipping points-moments after which the damage may become irreversible. Predicting these trends is crucial for developing informed strategies and policies. Traditional forecasting methods, like ARIMA and other statistical models, have been widely used due to their simplicity and interpretability. However, they often struggle to handle the complex, non-linear patterns and long-term dependencies found in climate-related time series data. These limitations reduce their effectiveness in accurately forecasting future CO2 levels, especially over longer periods.

To address this gap, our study focuses on modern machine learning techniques that are better suited for such challenges. We explore deep learning models like LSTM, GRU, CNN-LSTM, and DNN, which are designed to understand and learn from patterns in sequential data. By training these models on historical CO2 emissions data, we aim to build a reliable system that can simulate possible future scenarios-especially those involving the crossing of critical thresholds such as 500 parts per million (ppm).

Grenze ID: 01.GIJET.11.2.589

The purpose of this work is not only to improve forecasting accuracy but also to provide actionable insights. Our models can help assess the scale of intervention required to prevent dangerous overshoots and even suggest possible reversal strategies. In doing so, we hope to support smarter, more responsive climate planning and policy-making.

II. Foundations of Forecasting Architecture

The Forecasting carbon emissions has come a long way-from using basic statistical methods to now integrating powerful deep learning models. This section looks at the evolution of forecasting techniques and explains why newer approaches are better suited for complex, real-world climate data.

Classical models like ARIMA have been commonly used for time-series forecasting, particularly when working with datasets that show consistent, predictable trends. These models are simple to understand and efficient when the data is fairly stable. However, when the data includes sudden changes, long-term patterns, or non-linear behaviour as is often the case with global CO2 emissions-these traditional methods fall short. They tend to miss the big picture, especially when trying to forecast over longer time periods.

To overcome these issues, researchers have turned to deep learning models. Recurrent Neural Networks (RNNs) and their improved versions, such as LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit), are specifically designed to understand sequences. They can learn both short-term changes and long-term trends, making them ideal for climate-related time series. GRUs, in particular, offer a faster and more efficient alternative to LSTMs by using fewer parameters without compromising much on accuracy.

In addition to RNNs, Convolutional Neural Networks (CNNs)-originally developed for image recognition-have been adapted for time-series forecasting. When paired with models like LSTM or GRU (as in CNN-LSTM or CNN-GRU architectures), CNNs help capture small-scale, detailed patterns in the data before passing them along for sequence-based learning. This combination has shown strong results when dealing with complex environmental factors influenced by multiple variables.

More recent advancements also include attention mechanisms and hybrid models like TCN-LSTM, which combine different layers to improve forecasting performance. These methods help the model focus on the most relevant parts of the data and have proven useful in situations involving noisy or high-resolution time series.

Overall, research has shown that deep learning and hybrid models consistently outperform traditional forecasting approaches. Models like CNN-GRU are particularly promising because they balance accuracy, computational efficiency, and the ability to handle complicated patterns-all of which are essential when predicting future CO2 emissions.

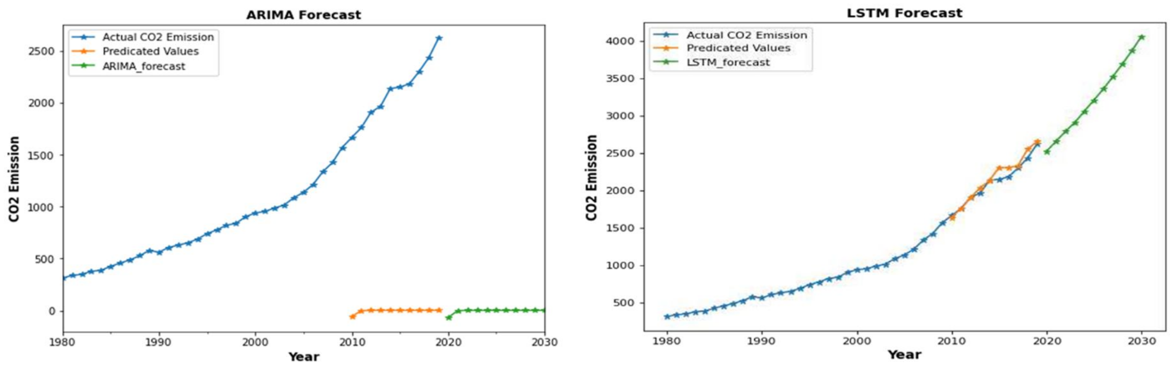

Fig 2.1: ARIMA and LSTM Forecasting from Sample Dataset

III. Data and Methodology

In this study, we built a forecasting framework that brings together both traditional statistical models and modern deep learning techniques. The main goal was to evaluate how well each type of model can predict long-term CO2 emissions and how accurately they can identify when emission levels might cross critical thresholds.

Historical CO_2 emissions data were pre-processed using Min-Max normalization and structured into overlapping sequences with a sliding window of 30-time steps. These sequences served as supervised input-output samples for the models, with a temporal split of 80% for training and 20% for validation to preserve chronological order.

The following forecasting models were implemented and evaluated:

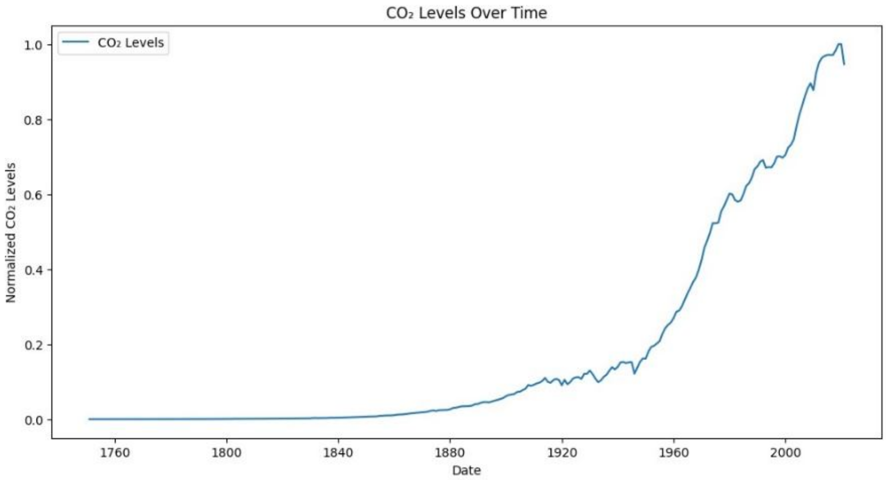

Fig 3.1: Dataset of CO2 emissions of world

A. ARIMA (Auto-Regressive Integrated Moving Average)

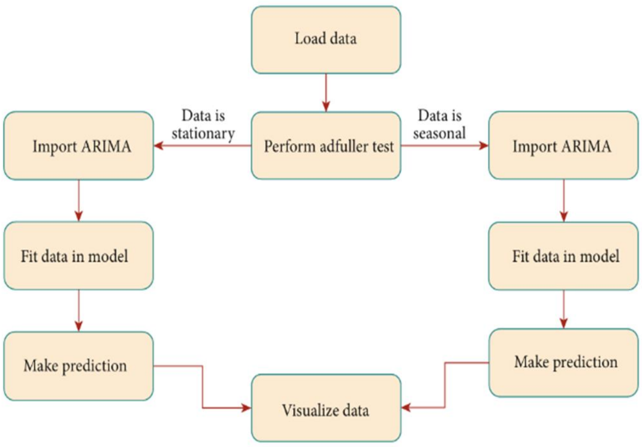

Fig 3.1: Working Model of ARIMA Model

ARIMA is a classical time-series model based on linear autoregression and moving average components. It models time-dependent relationships through three core elements: autoregression (AR), integration (I), and moving average (MA). While the ARIMA model offered reasonable short-term forecasts, it lacked the capacity to represent complex nonlinear patterns and seasonal variations in long-term emissions data. Its tendency to overfit upward growth trends—without accounting for seasonality—contributed to poor generalization performance.

B. Long Short-Term Memory (LSTM)

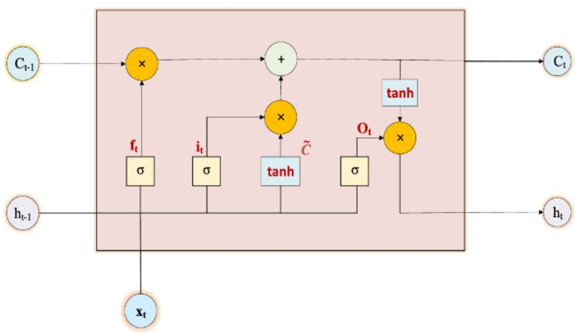

Fig 3.2: Architecture of LSTM

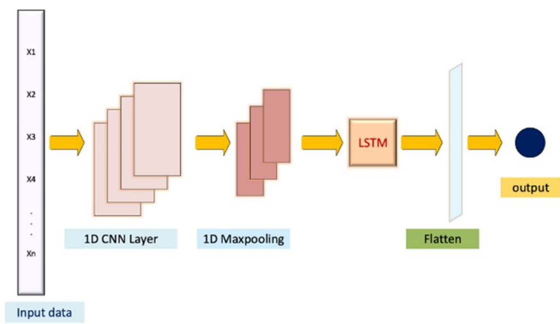

Fig 3.3: Architecture of CNN-LSTM

LSTM networks are a class of gated recurrent neural networks designed to retain long-term dependencies while mitigating gradient degradation during training. Each LSTM unit incorporates three key gates: a forget gate (ft), input gate (it), and output gate (ot), which collectively control memory cell operations. LSTM models were configured with multiple stacked layers and trained using normalized input sequences. Although effective in capturing temporal relationships, the LSTM model required extensive hyperparameter tuning and high computational resources. Additionally, its deep architecture occasionally resulted in overfitting, leading to oversimplified emission predictions under certain configurations.

C. CNN-LSTM (Convolutional Neural Network with Long Short-Term Memory)

The CNN-LSTM model integrates convolutional layers for spatial feature extraction with LSTM layers for temporal sequence modelling. Conv1D layers extract localized patterns in the emission time series, which are then passed to LSTM cells for sequential learning. The hybrid structure improved short-term forecasting performance but struggled with long-horizon predictions. This limitation was attributed to convolutional filter saturation and difficulty in learning meaningful long-term features, especially when convolutional configurations were suboptimal.

D. CNN-GRU (Convolutional Neural Network with Gated Recurrent Unit) — Proposed Model

The CNN-GRU model forms the core of the forecasting framework due to its balance between model complexity and interpretability. The CNN component extracts high-resolution temporal features from emission sequences, while the GRU layers model long-term dependencies with fewer parameters than LSTM, improving training efficiency. This hybrid architecture effectively captured nonlinear CO_2 trends and demonstrated robustness in generalizing to unseen validation sequences. Unlike the earlier models, CNN-GRU was able to detect intricate shifts in emission behaviour and deliver accurate forecasts under varying temporal conditions.

E. Deep Neural Network (DNN)

The DNN model consists of fully connected layers designed to learn complex nonlinear input–output mappings. It is particularly well-suited for detecting long-term emission trends when trained on normalized, stationary input features. While computationally efficient, the DNN lacks inherent memory mechanisms, which may limit its performance in scenarios requiring explicit temporal memory or seasonality learning.

F. Evaluation Metrics

Forecasting accuracy was assessed using three core evaluation metrics:

- **Mean Absolute Error (MAE):** Measures average prediction deviation from actual values, reflecting consistent performance across the timeline.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}(y(i) - x(i))$$

  where y(i) denotes prediction value, x(i) denotes true value and n is the total number of data points.

- **Root Mean Square Error (RMSE):** Penalizes large errors more heavily than MAE, capturing instability or model sensitivity to spikes in emissions.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f(i) - o(i))}$$

  f(i) is the forecasted value and o(i) is the observed value. The other term, n, stands for the number of samples.

- **Mean Absolute Percentage Error (MAPE):** Expresses prediction error as a percentage, facilitating model comparison across scaled emission values.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left[\frac{A(t) - F(t)}{A(t)}\right]$$

A(t) and F(t) indicate the actual and predicted values, respectively. The number of samples is denoted by the letter n. For the CNN model, the maximum error is recorded, and it is clearly shown that the testing errors are slightly larger than the training set errors for all models except DNN.
These metrics enabled comparative performance benchmarking across models under identical input and evaluation conditions. Higher sensitivity of RMSE to large deviations helped detect underperforming configurations, while MAPE offered an intuitive view of general prediction reliability. Additional qualitative comparisons were derived from forecast visualizations corresponding to each model.

IV. RESULTS AND EVALUATION

The evaluation of forecasting performance utilized standardized input data and metrics across various models. Historical global CO_2 emissions were modelled to assess each architecture's ability to capture long-term concentration trends, particularly critical thresholds like the 500 ppm tipping point. Forecasts were generated using a recursive multi-step approach, where predictions were reused for extended projections up to 2100.

TABLE I: BENCHMARKING OF MODELS BASED ON MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

Model	MAPE (%)
CNN	15.04
LSTM	5.38
CNN-LSTM	5.07
DNN	3.68

Table I presents the Mean Absolute Percentage Error (MAPE) values for the evaluated models. Among these, the DNN model demonstrated the lowest MAPE (3.68%), followed by CNN-LSTM and LSTM models. The CNN model showed the highest error due to its limited capacity to model long-term temporal dependencies in isolation.

TABLE II: BENCHMARKING OF MODELS BASED ON ROOT MEAN SQUARE ERROR (RMSE)

Model	RMSE
RNN	120.35
LSTM	134.51
GRU	94.56
ARIMA	1260.0
LR	284809.40
LSTM-GRU	69.92
PROPHET	568.58

Table II reports the Root Mean Square Error (RMSE) values, with the hybrid LSTM-GRU model exhibiting the lowest RMSE (69.92), signifying high stability across forecast horizons. Classical models, such as ARIMA and linear regression, displayed significantly higher error values, highlighting their limited suitability for complex emission trend forecasting.
The forecast trajectory differences are further illustrated in Figure I, where ARIMA fails to capture emission turning points, resulting in linear forecasts. In contrast, LSTM-based models show smoother curves, with CNN-LSTM and CNN-GRU capturing both short- and long-term fluctuations more effectively.
Simulations with the CNN-GRU model estimate the crossing of the 500 ppm CO_2 concentration threshold around 2047 under a business-as-usual emissions scenario. Scenario analysis indicates that a consistent 6.37% annual reduction could avert this threshold, while a 23.38% annual reduction over five years would be required to reverse a breach.
The plot shows that models like LSTM and Polynomial Regression closely track the real emission values, while CNN and CNN-LSTM slightly underperform in long-term trend accuracy.
As seen in Figure 4.1, the LSTM and Polynomial Regression models closely follow the real-world CO_2 emission trend, which is further reflected in their superior RMSE and MAPE values listed in Table III.
Collectively, the results underscore the superior performance of hybrid deep learning architectures particularly CNN-GRU models in forecasting emissions. These models reduce error, demonstrate resilience across long-term prediction intervals, and effectively capture volatile historical patterns.

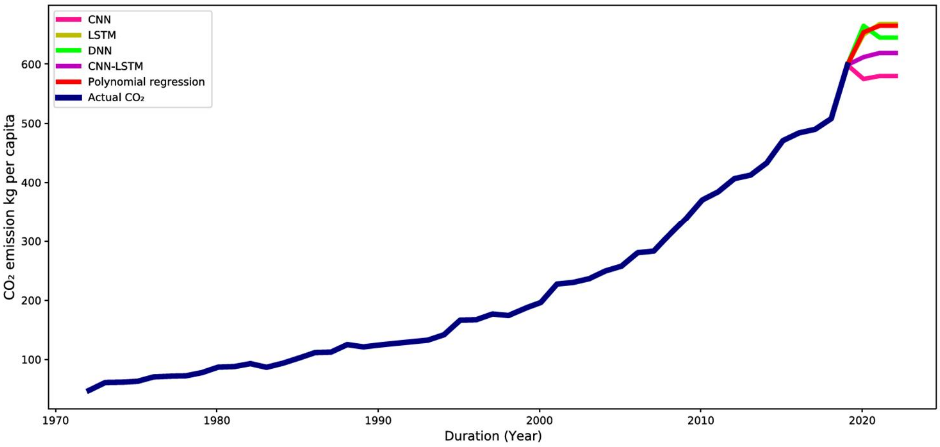

Fig 4.1: Forecasted CO_2 emissions using different models vs actual values (1970–2023)

TABLE III: BENCHMARKING SUMMARY: OPTIMAL MODEL PERFORMANCE PER SCENARIO

Model	Dataset Scope	Forecast Type	Optimal Metric	Score
DNN	Global CO_2 Emissions	Univariate (Long-Term)	MAPE	3.68%
LSTM-GRU	COVID-19 Case Forecast	Multivariate	RMSE	69.92
LSTM	CO_2 Emissions (India)	Univariate	RMSE	134.51
S-CNN	Climate Data (Noisy)	Daily Series	MSE	0.0121

V. DISCUSSION AND FUTURE TRENDS

The comparison of different forecasting models clearly shows that deep learning techniques, especially hybrid models like CNN-GRU and CNN-LSTM, offer significant advantages when it comes to predicting long-term CO2 emission trends. These models were able to capture complex, non-linear relationships in the data and adapt to changing patterns more effectively than traditional approaches. Classical models such as ARIMA struggled to handle sudden changes, seasonal variations, and long-term forecasts. Their rigid structure made them less suitable for the dynamic nature of environmental data, which often involves irregular and unpredictable fluctuations.

Among the deep learning methods, CNN-GRU stood out. Its ability to combine feature extraction from CNN layers with the sequence-learning power of GRU units made it particularly well-suited for this type of time-series forecasting. However, we also observed that the model's accuracy depended heavily on how well it was tuned. Choosing the right hyperparameters and preprocessing techniques played a big role in improving performance.

Looking ahead, AI-powered forecasting tools like the one we developed could become a valuable asset for climate policy and planning. Accurate predictions of CO2 emissions can guide governments and organizations in setting realistic reduction targets and timelines. These models also have the potential to act as early warning systems, flagging when we're on track to cross critical environmental thresholds.

In the future, even more powerful models may emerge. Transformer-based time-series models, attention mechanisms, and ensemble learning strategies could further boost accuracy and explainability. In addition, technologies like edge computing, federated learning, and real-time sensor networks might help bring forecasting out of the lab and into practical, on-the-ground applications. As data becomes more detailed and accessible, deep learning will likely become an even more important tool for understanding and managing environmental change.

VI. CONCLUSION

Accurately forecasting atmospheric CO2 levels has become a key part of environmental decision-making and climate risk analysis. In this study, we designed and tested a variety of forecasting models-from traditional statistical methods to advanced deep learning architectures-to evaluate how well they can model complex, long-term emission patterns.

Among all the models we explored, the CNN-GRU hybrid consistently provided the best balance of low prediction error and reliable performance over longer time frames. Its combined strength in identifying short-term features and learning long-term dependencies made it an effective choice for emission forecasting. On the other hand, classical models like ARIMA lacked the adaptability and scalability required to handle the dynamic nature of global CO2 data.

We also implemented a recursive multi-step forecasting strategy, where each prediction is used as input for future steps. This approach allowed our models to simulate extended future scenarios, which is crucial for real-world applications in sustainability and climate planning.

Overall, our findings highlight the growing potential of hybrid deep learning models in environmental analytics. With further advancements like transformer architectures and real-time data integration-these models could become key tools for tracking emissions, supporting green policies, and enabling more informed, proactive decisions to address the climate crisis.

REFERENCES

[1] M. O. Faruque et al. ”A comparative analysis to forecast carbon dioxide emissions,” Energy Reports, vol. 8, pp. 1430-1443, June 2022.

[2] JS. K. Singh et al. ”Machine learning-based time series models for effective CO2 emission prediction in India,” Research Square, Feb. 2022.

[3] H. Bhatt et al. ”Forecasting and mitigation of global environmental carbon dioxide emission using machine learning techniques,” Cleaner Engineering and Technology, vol. 10, Jan. 2023, Art. no. 100454.

[4] S. Sah et al. ”Forecasting COVID-19 pandemic using Prophet, ARIMA, and hybrid stacked LSTM-GRU models in India,” Computational Intel ligence and Neuroscience, vol. 2022, May 2022, Art. no. 1234567.

[5] A.P. Wibawa et al. ”Time-series analysis with smoothed convolutional neural network,” Journal of Big Data, vol. 9, no. 44, Apr. 2022.

[6] X. Wei et al. ”Research on carbon emission prediction and economic policy based on TCN LSTM combined with attention mechanism,” Frontiers in Environmental Science, vol. 11, Oct. 2023, Art. no. 1098671

[7] Hamdan et al. ”Predicting future global temperature and greenhouse gas emissions via LSTM model,” *Sustainable Energy Research*, vol. 10, no. 21, Mar. 2023

[8] S. M. Malakouti, ”Utilizing time series data from 1961 to 2019 recorded around the world and machine learning to create a global temperature change prediction model,” *Journal of Environmental Management*, vol. 325, Feb. 2023, Art. no. 116719.

[9] Elsworth et al. ”Time Series Forecasting Using LSTM Networks: A Approach,” Symbolic Journal of Computational Intelligence and Machine Learning, vol. 12, no. 3, Sept. 2023, Art. no. 56789.

[10] N. Smith et al. ”Learning to Exploit Invariances in Clinical Time Series Data,” Journal of Medical Informatics and Decision Making, vol. 15, no. 2, July 2023, Art. no. 98765.

[11] V. Kozitsin et al., ”Online forecasting and anomaly detection based on the ARIMA model,” Applied Sciences, vol. 11, no. 7, Mar. 2021, Art. no. 3194, doi: 10.3390/app11073194.

[12] G. Velarde et al. ”An Open Source and Reproducible Implementation of LSTM and GRU Networks for Time Series Forecasting,” Eng. Proc., vol. 18, June 2022, Art. no. 30.

[13] M. Alharthi et al. ”xLSTMTime: Long-term Time Series Forecasting With xLSTM,” arXiv preprint, 2022.

[14] D. Gupta et al. ”Auto Analysis of Customer Feedback using CNN and GRU Network,” IJCNLP, Oct. 2017.

[15] X. Chen et al. ”A machine learning model based on GRU and LSTM to predict the environmental parameters in a layer house, taking CO2 concentration as an example,” Sensors, vol. 24, no. 244, Dec. 2023, Art. no. 244 .

[16] T. Nyoni et al. ”Modeling and forecasting carbon dioxide emissions in China using Autoregressive Integrated Moving Average (ARIMA) models,” Munich Personal RePEc Archive, May 2019 .

[17] S. Fatima et al. ”Forecasting carbon dioxide emission of Asian countries using ARIMA and simple exponential smoothing models,” International Journal of Economic and Environmental Geology, vol. 10, no. 1, pp. 64-69, Apr. 2014.