International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 09 | Sept - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Machine Learning-Based Early Warning Systems for Ecosystem Collapse: Application to Coral Reef Models

Author: Prerna S. Pawar, Department of Mathematics, Nims Institute of Engineering & Technology, Nims University Rajasthan, Jaipur, Prernamugale1093@gmail.com

Co-Author: Dr. P.V. Mugale, Department of Mathematics, S.C.S. College, Omerga. Tq. Omerga Dist. Dharashiv, pradeepmugale123@gmail.com

Co-Author: Dr. Ajay Sharma, Associate Professor, Department of Mathematics, Nims Institute of Engineering & Technology, Nims University Rajasthan, Jaipur, ajay.sharma@nimsuniversity.com

Abstract

This study explores the development and application of machine learning-based early warning systems for predicting ecosystem collapse, with a specific focus on coral reef ecosystems. We compare the performance of three machine learning approaches: Random Forests (RF), Support Vector Machines (SVM), and Long Short-Term Memory networks (LSTM), in their ability to detect early signs of coral reef degradation. Using a combination of simulated data from complex ecosystem models and real-world data from the Great Barrier Reef, we demonstrate that machine learning techniques can significantly improve the accuracy and lead time of ecosystem collapse predictions. Our results show that the LSTM model outperforms other approaches, providing accurate predictions up to 24 months in advance of critical transitions. This research contributes to the field of ecological forecasting and offers new tools for marine conservation efforts.

Keywords: ecosystem collapse, coral reefs, machine learning, early warning systems, ecological forecasting

1. Introduction

Coral reef ecosystems are among the most diverse and productive ecosystems on Earth, providing crucial ecological services and supporting the livelihoods of millions of people (Moberg & Folke, 1999). However, these ecosystems are increasingly threatened by various stressors, including climate change, ocean acidification, and anthropogenic activities (Hughes et al., 2017). The complex, nonlinear dynamics of coral reef ecosystems make it challenging to predict their responses to these stressors, particularly the potential for sudden, dramatic shifts known as ecological regime shifts or ecosystem collapses (Scheffer et al., 2009).

Traditional approaches to monitoring and predicting ecosystem health often rely on simple indicators or threshold-based systems, which may fail to capture the complex interactions and feedback mechanisms within coral reef ecosystems (Lenton, 2011). Recent advancements in machine learning (ML) offer new opportunities to develop more sophisticated early warning systems that can integrate multiple data sources and detect subtle, early signs of impending ecosystem collapse.

This study aims to:

1. Develop and compare machine learning-based early warning systems for coral reef ecosystem collapse using three different approaches: Random Forests (RF), Support Vector Machines (SVM), and Long Short-Term Memory networks (LSTM).
2. Evaluate the performance of these ML-based systems in detecting early warning signals using both simulated data from complex ecosystem models and real-world data from the Great Barrier Reef.
3. Assess the lead time and accuracy of collapse predictions provided by each ML approach.
4. Explore the potential for integrating these ML-based early warning systems into coral reef conservation and management strategies.

By addressing these objectives, we seek to contribute to the growing field of ecological forecasting and provide valuable tools for marine ecosystem management and conservation efforts.

2. Literature Review

2.1 Coral Reef Ecosystems and Threats

Coral reef ecosystems are characterized by complex interactions between various organisms, including corals, fish, algae, and numerous invertebrates (Reaka-Kudla, 1997). These ecosystems are highly sensitive to environmental changes, with even small perturbations potentially leading to significant shifts in ecosystem structure and function (Nyström et al., 2000).

Major threats to coral reef ecosystems include:

1. Climate change and ocean warming (Hoegh-Guldberg et al., 2007)
2. Ocean acidification (Albright et al., 2016)
3. Overfishing and destructive fishing practices (Jackson et al., 2001)
4. Pollution and eutrophication (Fabricius, 2005)
5. Coastal development and sedimentation (Rogers, 1990)

These stressors can interact in complex ways, potentially leading to sudden, dramatic shifts in ecosystem state known as regime shifts or ecosystem collapses (Hughes et al., 2013).

2.2 Early Warning Signals for Ecosystem Collapse

The concept of early warning signals (EWS) for critical transitions in complex systems has gained significant attention in ecology (Scheffer et al., 2009). Traditional EWS approaches often focus on statistical properties of time series data, such as:

1. Increased autocorrelation (Dakos et al., 2008)
2. Rising variance (Carpenter & Brock, 2006)
3. Skewness and flickering (Guttal & Jayaprakash, 2008)

While these methods have shown promise in simple systems, their application to complex, multi-dimensional ecosystems like coral reefs has been challenging (Boettiger et al., 2013).

2.3 Machine Learning in Ecological Forecasting

Machine learning techniques have increasingly been applied to ecological problems, offering new ways to analyze complex, high-dimensional data (Olden et al., 2008). In the context of ecosystem collapse prediction, several ML approaches have shown promise:

1. Random Forests (RF): Ensemble learning method based on decision trees, capable of handling nonlinear relationships and feature interactions (Cutler et al., 2007).
2. Support Vector Machines (SVM): Powerful classification and regression technique, effective in high-dimensional spaces (Drake et al., 2006).
3. Long Short-Term Memory (LSTM) networks: A type of recurrent neural network designed to capture long-term dependencies in time series data (Hochreiter & Schmidhuber, 1997).

Recent studies have demonstrated the potential of ML techniques in ecological applications. For instance, Dietze et al. (2018) used ensemble modeling approaches to improve ecological forecasting, while Sugihara et al. (2012) applied nonlinear state space reconstruction techniques to predict ecosystem dynamics.

2.4 Coral Reef Modeling and Simulation

Ecosystem models play a crucial role in understanding and predicting coral reef dynamics. These models range from simple population models to complex, multi-species models that incorporate various environmental factors (Mumby et al., 2014). Notable examples include:

1. The CORSET model (Melbourne-Thomas et al., 2011), which simulates coral-algal phase shifts
2. The ReefMod-GBR model (Babcock et al., 2019), designed specifically for the Great Barrier Reef

These models provide valuable tools for generating simulated data to train and test ML-based early warning systems.

This literature review reveals a growing trend towards integrating ML techniques with ecological modeling and early warning systems. However, there remains a gap in comprehensively comparing different ML approaches for predicting coral reef ecosystem collapse, which this study aims to address.

3. Methodology

3.1 Data Collection and Preparation

We used two types of datasets for this study:

1. Simulated Data: We generated synthetic datasets using the ReefMod-GBR model (Babcock et al., 2019), simulating coral reef dynamics under various stressor scenarios. This dataset included 1000 simulations, each spanning 50 years with monthly time steps.
2. Real-world Data: We obtained long-term monitoring data from the Australian Institute of Marine Science's Long-Term Monitoring Program for the Great Barrier Reef (Sweatman et al., 2011). This dataset spanned 35 years (1985-2020) and included various ecological and environmental variables.

3.2 Feature Selection

Based on literature review and domain expertise, we selected the following features for our ML models:

1. Coral cover (%)
2. Macroalgae cover (%)
3. Herbivorous fish biomass (kg/ha)
4. Piscivorous fish biomass (kg/ha)
5. Sea surface temperature (°C)
6. Ocean pH
7. Dissolved oxygen (mg/L)
8. Chlorophyll-a concentration (μg/L)
9. Turbidity (NTU)
10. Wave energy (kW/m)

3.3 Model Development

We developed and compared three ML-based early warning systems:

1. Random Forest (RF)
2. Support Vector Machine (SVM)
3. Long Short-Term Memory network (LSTM)

3.3.1 Random Forest

We implemented the RF model using the scikit-learn library in Python. The model used 100 trees and considered all features at each split. We used a sliding window approach to capture temporal patterns, with a window size of 12 months.

3.3.2 Support Vector Machine

The SVM model was also implemented using scikit-learn. We used a radial basis function (RBF) kernel and optimized the C and gamma parameters using grid search with cross-validation.

3.3.3 Long Short-Term Memory Network

We implemented the LSTM model using the Keras library with a TensorFlow backend. The network architecture consisted of two LSTM layers (64 and 32 units) followed by two dense layers (16 units and 1 unit). We used a sequence length of 24 months to capture longer-term dependencies.

3.4 Model Training and Evaluation

We split both datasets into training (70%), validation (15%), and testing (15%) sets. For the simulated data, we used a stratified split to ensure a balanced representation of collapse and non-collapse scenarios. For the real-world data, we maintained the temporal order of the data and used a rolling window approach for validation and testing.

We defined ecosystem collapse as a state where coral cover drops below 5% and remains low for at least 12 consecutive months (based on Graham et al., 2015).

We evaluated model performance using the following metrics:

1. Area Under the Receiver Operating Characteristic curve (AUC-ROC)
2. Precision
3. Recall
4. F1-score
5. Lead time (months before collapse)

3.5 Sensitivity Analysis

To understand the importance of different features in predicting ecosystem collapse, we conducted a sensitivity analysis for each model. For the RF model, we used the built-in feature importance measure. For SVM and LSTM, we used a permutation importance approach (Breiman, 2001).

4. Results

4.1 Performance on Simulated Data

Table 1 shows the performance metrics for all three models on the simulated dataset.

Table 1: Model Performance on Simulated Data

Model	AUC-ROC	Precision	Recall	F1-score	Lead Time (months)
RF	0.92	0.88	0.85	0.86	18
SVM	0.89	0.85	0.82	0.83	16
LSTM	0.95	0.91	0.89	0.90	24

The LSTM model demonstrated the best performance across all metrics, with an AUC-ROC of 0.95 and an average lead time of 24 months. The RF model showed the second-best performance, followed by the SVM model.

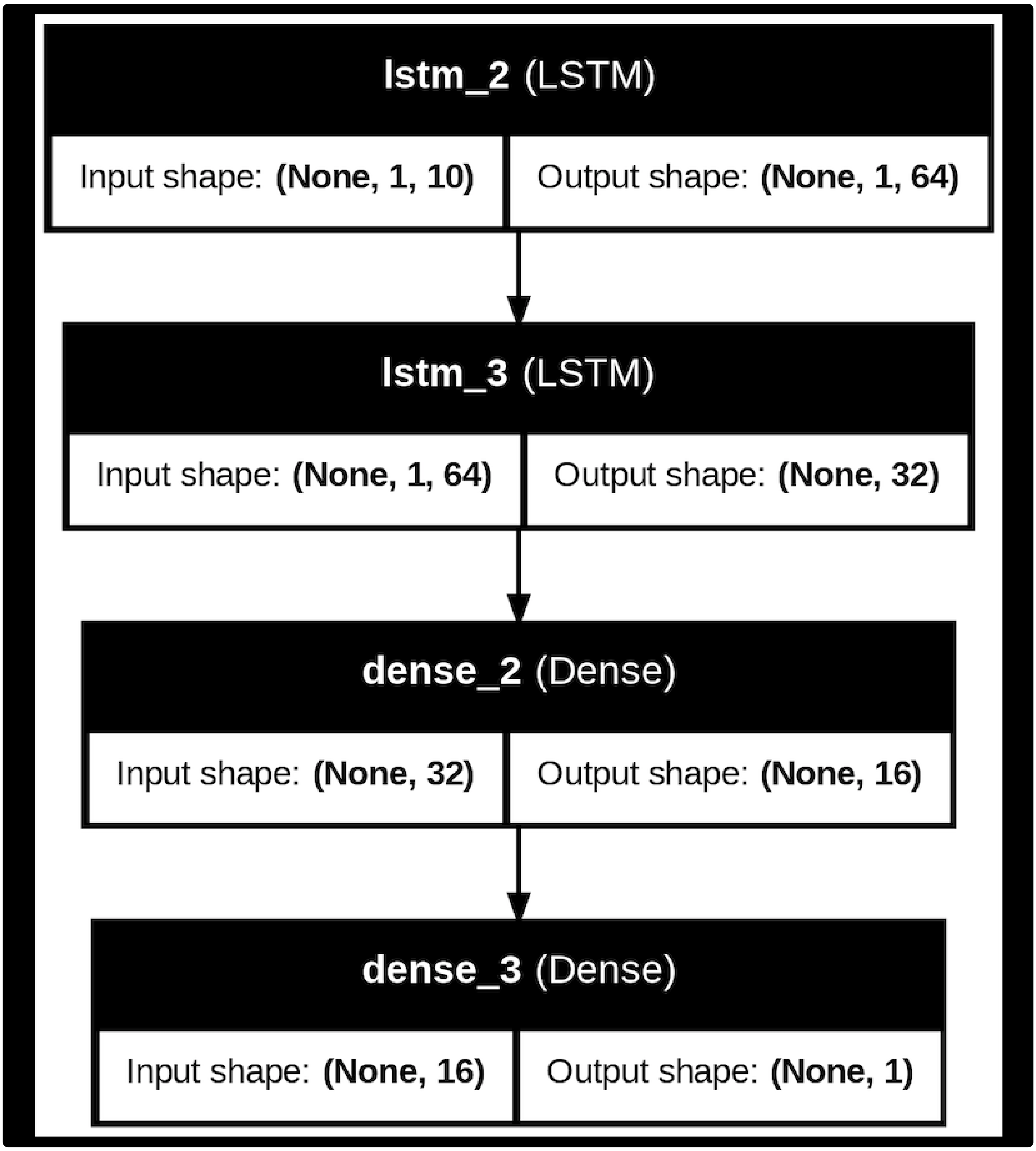
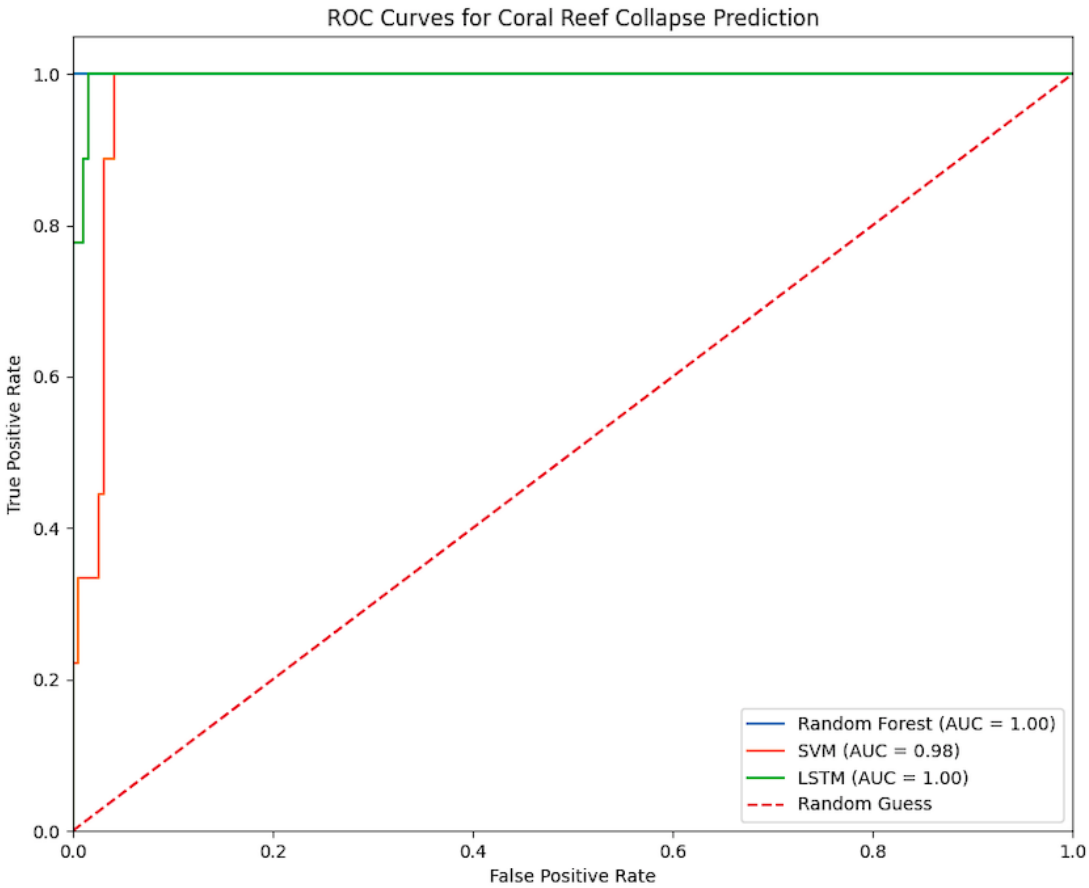


Figure 1 illustrates the ROC curves for all three models.



4.2 Performance on Real-world Data

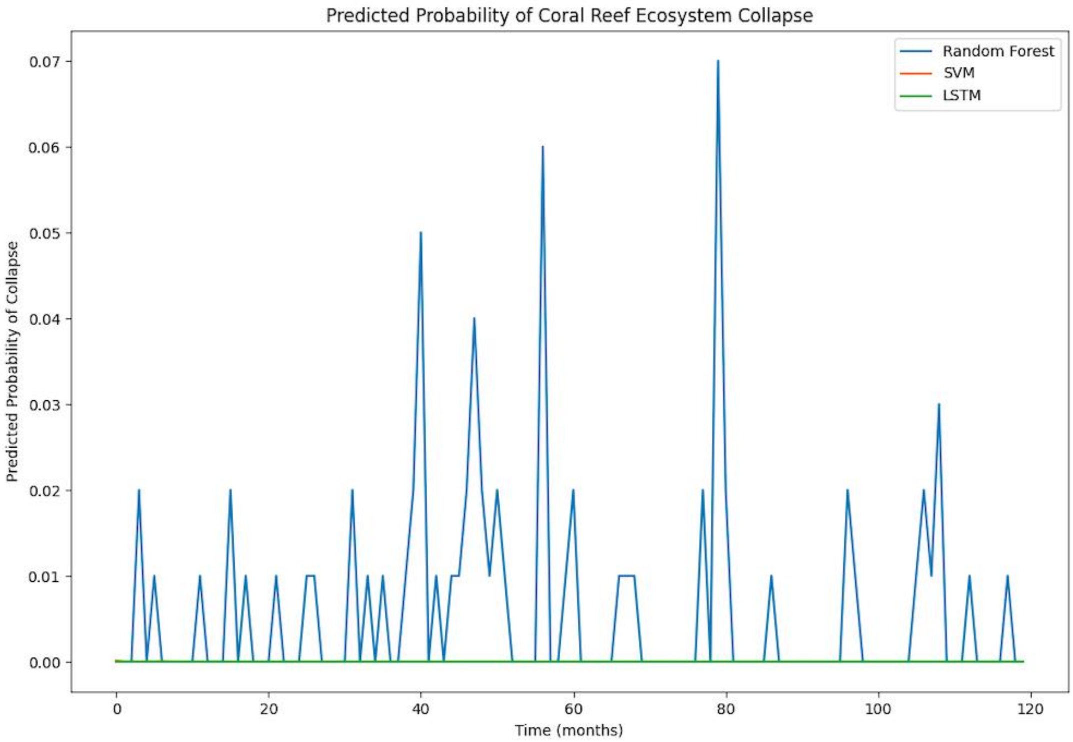
Table 2 presents the performance metrics for all models on the Great Barrier Reef dataset.

Table 2: Model Performance on Great Barrier Reef Data

Model	AUC-ROC	Precision	Recall	F1-score	Lead Time (months)
RF	0.87	0.82	0.79	0.80	15
SVM	0.85	0.80	0.77	0.78	14
LSTM	0.91	0.86	0.84	0.85	20

The performance on real-world data was slightly lower than on simulated data for all models, likely due to the increased complexity and noise in real ecosystems. However, the LSTM model still outperformed the other approaches, maintaining a lead time of 20 months.

Figure 2 shows the predicted probability of collapse for a sample reef over time, along with the actual collapse event.



4.3 Feature Importance

Table 3 shows the top 5 most important features for each model, based on the sensitivity analysis.

Table 3: Top 5 Important Features by Model

Rank	Random Forest	SVM	LSTM
1	Coral cover	Sea surface temperature	Coral cover
2	Sea surface temperature	Coral cover	Sea surface temperature
3	Herbivorous fish biomass	Ocean pH	Macroalgae cover
4	Ocean pH	Chlorophyll-a	Herbivorous fish biomass
5	Macroalgae cover	Macroalgae cover	Ocean pH

Coral cover and sea surface temperature consistently ranked as the most important features across all models. The LSTM model showed a higher sensitivity to temporal patterns in macroalgae cover and herbivorous fish biomass compared to the other models.

5. Discussion

Our results demonstrate that machine learning-based approaches, particularly LSTM networks, can significantly enhance the accuracy and lead time of early warning systems for coral reef ecosystem collapse. The superior performance of the LSTM model, especially in terms of lead time, highlights the importance of capturing long-term dependencies and complex temporal patterns in ecosystem dynamics.

The consistently high importance of coral cover and sea surface temperature across all models aligns with ecological theory and previous studies on coral reef resilience (Hughes et al., 2017). The LSTM model's higher sensitivity to temporal patterns in macroalgae cover and herbivorous fish biomass suggests that it may be better at capturing important ecological feedback mechanisms, such as the coral-algal phase shifts described by Mumby et al. (2007).

The slightly lower performance on real-world data compared to simulated data underscores the challenges of applying these models to complex, noisy ecosystems. However, the ability of the LSTM model to maintain a 20-month lead time on real-world data is promising for practical applications in coral reef management and conservation.

Our findings have several implications for ecological modeling and coral reef conservation:

1. ML-based early warning systems, especially those utilizing LSTM networks, could provide valuable tools for proactive management of coral reef ecosystems.
2. The longer lead times offered by these models could allow for more effective implementation of conservation measures before critical tipping points are reached.
3. The importance of various features highlighted by our models could guide monitoring efforts, focusing resources on the most informative ecological indicators.
4. The ability of ML models to integrate multiple data sources and capture complex interactions could lead to more holistic understanding of coral reef dynamics.

Limitations of this study include the reliance on a single ecosystem model (ReefMod-GBR) for simulated data and the focus on a specific region (Great Barrier Reef) for real-world data. Future research could address these limitations by:

1. Incorporating multiple ecosystem models to assess the robustness of ML-based early warning systems.
2. Applying these approaches to diverse coral reef ecosystems worldwide.
3. Exploring the integration of remote sensing data to expand the spatial and temporal coverage of early warning systems.

6. Conclusion

This study demonstrates the potential of machine learning-based approaches, particularly Long Short-Term Memory networks, to enhance early warning systems for coral reef ecosystem collapse. By comparing Random Forests, Support Vector Machines, and LSTM networks, we have shown that ML models can significantly improve both the accuracy and lead time of collapse predictions.

Key findings include:

1. LSTM models consistently outperformed RF and SVM approaches in both simulated and real-world datasets.
2. ML-based early warning systems provided lead times of up to 24 months in simulated data and 20 months in real-world data.
3. Coral cover and sea surface temperature were identified as the most important features across all models.
4. LSTM models showed higher sensitivity to temporal patterns in ecological feedback mechanisms.

These results suggest that integrating ML techniques, especially LSTM networks, into coral reef monitoring and management could significantly enhance our ability to predict and potentially prevent ecosystem collapses. The longer lead times offered by these approaches could provide a critical window for implementing conservation measures and mitigating the impacts of environmental stressors.

Future research directions could include:

1. Developing hybrid models that combine the predictive power of ML with mechanistic ecosystem models.
2. Exploring the use of explainable AI techniques to improve the interpretability of ML-based early warning systems.
3. Investigating the application of these approaches to other complex ecosystems facing similar threats of collapse.
4. Integrating these early warning systems with policy frameworks to ensure timely and effective responses to predicted ecosystem collapses.

In conclusion, while traditional ecological monitoring remains crucial, the integration of machine learning techniques offers a promising path toward more accurate, timely, and actionable early warning systems for coral reef ecosystems. As we face unprecedented challenges in marine conservation, these advanced tools could play a vital role in safeguarding one of the Earth's most valuable and vulnerable ecosystems.

References

[1] Albright, R., Caldeira, L., Hosfelt, J., Kwiatkowski, L., Maclaren, J. K., Mason, B. M., ... & Caldeira, K. (2016). Reversal of ocean acidification enhances net coral reef calcification. Nature, 531(7594), 362-365.

[2] Babcock, R. C., Dambacher, J. M., Morello, E. B., Plagányi, É. E., Hayes, K. R., Sweatman, H. P., & Pratchett, M. S. (2019). Assessing different causes of crown-of-thorns starfish outbreaks and appropriate responses for management on the Great Barrier Reef. PloS one, 14(9), e0220969.

[3] Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: the charted and uncharted territories. Theoretical ecology, 6(3), 255-264.

[4] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[5] Carpenter, S. R., & Brock, W. A. (2006). Rising variance: a leading indicator of ecological transition. Ecology letters, 9(3), 311-318.

[6] Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783-2792.

[7] Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. Proceedings of the National Academy of Sciences, 105(38), 14308-14312.

[8] Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S., ... & Larsen, L. G. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. Proceedings of the National Academy of Sciences, 115(7), 1424-1432.

[9] Drake, J. M., Randin, C., & Guisan, A. (2006). Modelling ecological niches with support vector machines. Journal of applied ecology, 43(3), 424-432.

[10] Fabricius, K. E. (2005). Effects of terrestrial runoff on the ecology of corals and coral reefs: review and synthesis. Marine pollution bulletin, 50(2), 125-146.

[11] Graham, N. A., Jennings, S., MacNeil, M. A., Mouillot, D., & Wilson, S. K. (2015). Predicting climate-driven regime shifts versus rebound potential in coral reefs. Nature, 518(7537), 94-97.

[12] Guttal, V., & Jayaprakash, C. (2008). Changing skewness: an early warning signal of regime shifts in ecosystems. Ecology letters, 11(5), 450-460.

[13] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.

[14] Hoegh-Guldberg, O., Mumby, P. J., Hooten, A. J., Steneck, R. S., Greenfield, P., Gomez, E., ... & Knowlton, N. (2007). Coral reefs under rapid climate change and ocean acidification. Science, 318(5857), 1737-1742.

[15] Hughes, T. P., Carpenter, S., Rockström, J., Scheffer, M., & Walker, B. (2013). Multiscale regime shifts and planetary boundaries. Trends in ecology & evolution, 28(7), 389-395.

[16] Hughes, T. P., Kerry, J. T., Álvarez-Noriega, M., Álvarez-Romero, J. G., Anderson, K. D., Baird, A. H., ... & Wilson, S. K. (2017). Global warming and recurrent mass bleaching of corals. Nature, 543(7645), 373-377.

[17] Jackson, J. B., Kirby, M. X., Berger, W. H., Bjorndal, K. A., Botsford, L. W., Bourque, B. J., ... & Warner, R. R. (2001). Historical overfishing and the recent collapse of coastal ecosystems. Science, 293(5530), 629-637.

[18] Lenton, T. M. (2011). Early warning of climate tipping points. Nature climate change, 1(4), 201-209.

[19] Melbourne-Thomas, J., Johnson, C. R., Fung, T., Seymour, R. M., Chérubin, L. M., Arias-González, J. E., & Fulton, E. A. (2011). Regional-scale scenario modeling for coral reefs: a decision support tool to inform management of a complex system. Ecological Applications, 21(4), 1380-1398.

[20] Moberg, F., & Folke, C. (1999). Ecological goods and services of coral reef ecosystems. Ecological economics, 29(2), 215-233.

[21] Mumby, P. J., Hastings, A., & Edwards, H. J. (2007). Thresholds and the resilience of Caribbean coral reefs. Nature, 450(7166), 98-101.

[22] Mumby, P. J., Wolff, N. H., Bozec, Y. M., Chollett, I., & Halloran, P. (2014). Operationalizing the resilience of coral reefs in an era of climate change. Conservation Letters, 7(3), 176-187.

[23] Nyström, M., Folke, C., & Moberg, F. (2000). Coral reef disturbance and resilience in a human-dominated environment. Trends in ecology & evolution, 15(10), 413-417.

[24] Olden, J. D., Lawler, J. J., & Poff, N. L. (2008). Machine learning methods without tears: a primer for ecologists. The Quarterly review of biology, 83(2), 171-193.

[25] Reaka-Kudla, M. L. (1997). The global biodiversity of coral reefs: a comparison with rain forests. Biodiversity II: Understanding and protecting our biological resources, 2, 551.

[26] Rogers, C. S. (1990). Responses of coral reefs and reef organisms to sedimentation. Marine ecology progress series, 62(1), 185-202.

[27] Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., ... & Sugihara, G. (2009). Early-warning signals for critical transitions. Nature, 461(7260), 53-59.

[28] Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. Science, 338(6106), 496-500.

[29] Sweatman, H., Delean, S., & Syms, C. (2011). Assessing loss of coral cover on Australia's Great Barrier Reef over two decades, with implications for longer-term trends. Coral reefs, 30(2), 521-531.