Rapid Adaptive Climate Change Model: Application of a Probabilistic Centred Approach to the Minas Passage Bay of Fundy datasets

Ian Kenny, *Member, IEEE*, Dhouha Kbaier, *Member, IEEE*

***Abstract*—We are interested in modelling smaller climate datasets in order to generate more accurate forecasts. Our approach would be to develop a fractal-based system around the mathematical attractor for a particular dataset, building on the ideas developed by the authors of this paper, to "measure" the distance between the model and the subsequently observed dataset. The goal of the research would be to develop a model where climate data could be iterated comparatively between models and a meaningful comparison made from which forecasts can be drawn. This is expected to be particularly useful given the increasing rapidity of climate change. This paper presents the initial findings of this research focused around discoveries related to the number of Intrinsic Mode Functions (IMFs) required for the decomposition of the signal, obtained by the Empirical Mode Decomposition (EMD), independent from the number of observations.**

***Index Terms*— climate change, modelling, Empirical Mode Decomposition (EMD), adaptive forecasting, mathematical attractor, forecasting model, hydrospheric data, atmospheric data**

I. INTRODUCTION

ACCURATE and reliable climate predictions across a range of temporal and spatial scales is of increasing importance to policy makers. Forecasts of extreme weather a few days ahead, seasonal variations, inter-annual climate variability and decadal projections of the long-term changing climate can drive preparedness and disaster mitigation. Some examples are forecasts of US Hurricane magnitude and landfall (days), predictions of the onset of the Indian Monsoon (seasonal) [1], estimates of El Nino's influence on European climate (year-to-year) [2] and projections of global mean temperate for the end of century [3] In the last decade, we have experienced some of the warmest years on record, driving the emergence of new and increased extreme weather [4], [5], highlighting the importance of predicting patterns of future climate and extremes for local, regional and global decision making.

In addition to investing considerable resources into weather forecasts, society also invests in complex global climate models (GCMs) - deterministic numerical models that solve equations of motion for the atmosphere and oceans to provide quantitative forecasts and projections. GCMs are widely considered the state-of-the-art dynamical method in seasonal forecasting. [6], [7] however, argue that obtaining accurate short-term forecasts of extreme conditions is challenging. Another drawback of GCMs is that they tend to be biased and have a significant spread across ensembles [8], making them unsuitable for decision-support tools. On interannual and decadal scales, modelling the climate, in particular its variability, remains a challenge and projections can differ from model to model [9].

The aim of our research is to explore the modelling of smaller datasets to produce useful forecasts and subsequently produce effective recommendations about climate and other timeseries datasets where the dataset is lacking richness or quality, for example where the climate is changing more rapidly than represented in the historical data. Building on the ideas in [10], [11] and [12] using a fractal-based system, after [13], [14] to derive the mathematical attractor for a specific dataset and use that to estimate the rate of change between the current model and the subsequently observed data. Allowing for probabilistic predictions to be made regarding the evolution of the system. Based on research by [15] that using an iterative approach, one could develop a series of models which are not based on big data models but would in the case of climate data enable predictions to be made over the 5-to-10-year period. More importantly to keep this cycle iterating so that the model can be more responsive and adaptive to changes in the climate.

II. PROPOSED MODEL

Figure 1 shows a schematic of the proposed model development process. Our innovative Rapidly Adaptive Climate Change (RACC) model is expected to be particularly useful given the increasing rapidity of climate change. As an initial step towards this goal, we intend to build a model which integrates the atmospheric data with the hydrospheric data with the intent of allowing the model to derive its own relationship between the heat cycle within the hydrosphere, including the cryosphere, and the effect on the global mean atmospheric temperature.

Our RACC approach of building models at different layers preserves the abstraction needed to keep individual datasets distinct whilst at the same time relating them in a way that could be brought together to produce forecasts. To find the underlying model we will use Empirical Mode Decomposition (EMD)

.

[16], this is in order to derive the Intrinsic Mode Functions (IMFs). However, instead of relying on the IMFs as the end result, for example [17], these will be used to derive an Iterated Function System (IFS) the purpose for which is to overcome the limitations of EMD namely, it is difficult to identify the periodic oscillators from which IMFs are derived and the number of IMFs cannot be predicted and rely on empirical observations of the timeseries dataset [18]. The aim is to produce a model which incorporates all sources of climate science data to produce an integrated comparative model.

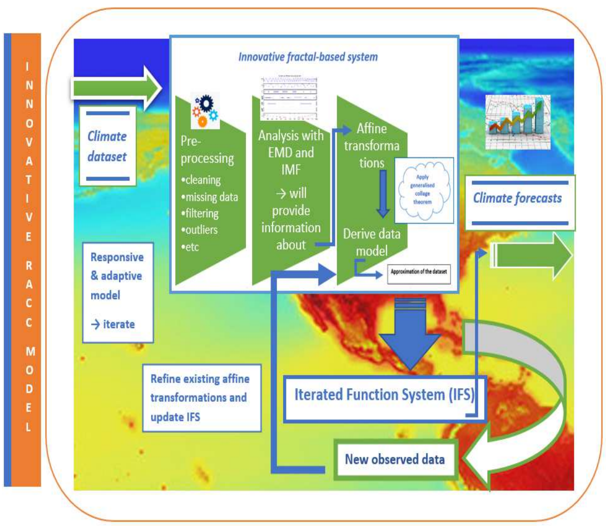

Fig. 1. Schematic of Rapid Adaptive Climate Model.

III. OUR RESEARCH APPROACH

An approach to modelling localised climate change is proposed, using an iterative model. We propose that applying the principle to climate science datasets allows for forecasts to be developed in such a way that a reproducible forecasting model can be applied in a way that is adaptive to the changing dataset observed in the real world; without significant computational overhead to recompute the interactions between large-scale datasets. Mathematically, this is the process of inverse problem construction. Further, our approach enables the relationships within the data to be discovered without making assumptions about issues such as lag applied within each dataset and across the datasets. This, for example, would potentially allow the relationship between the heat capacity of the hydrosphere and atmosphere to be approximated within the transformations which were discovered. By using the set of affine transformations and the probabilistic associations, we can then rebuild the original dataset in abstract manner to determine the characteristics of the part of the climate system under study. This has the advantage of being able to discover the inverse of the dataset which represents the observation seen in the real world. Once derived, these sets of functions that describe the model, can be used to make predictions about future change on a probabilistic basis. The predictions therefore derived will be based on a probabilistic forecast model similar to that seen in [15]. However, by making the model work in this way we can iteratively repeat the process as new observational data becomes available. The proposal is that a system of transitional matrices is developed to probabilistically move the system from the current state to the next state. Each matrix, will contain a representation of the weights required to move the system into the next state given a probabilistic value and applied to the current state. This means that any change in the lag of influence between the respective spheres, for example, the effect of melting ice caps on the hydrosphere's capacity to store and release heat, will be adjusted for, over time.

In summary, the problem we are trying to solve is to produce a method for iteratively using an adaptive climate change model which integrates datasets representing the different spheres. The current approaches to this problem typically summarise the parameters as a single matrix, [19], or are non-iterative, [20]. The former approach of summating the complex relationships between datasets, and indeed elements of the climate system, has the shortcoming that should any of the relationships change, for example, due to climate change, this has to be recalculated. Any model derived in this way has a limited life, and needs to be reproduced from the original principles rather than evolving as the climate system changes, or a climate event occurs for example an El Niño. The latter approach whilst an improvement on the summarising matrix, does not evolve. Given the complexity of the relationships between various aspects of the climate system, there is a need for an integrated approach which is data-centric and allows for changes to occur within the relationships which are subsequently modelled.

IV. ROLE OF EMPIRICAL MODE DECOMPOSITION (EMD)

We took a stepwise approach to our innovation. We would like to explore whether there is a relationship within the EMD technique which enables us to model relationships based on the number of variables not the number of observations. If this could be achieved it would enable a probabilistic model of the data to be developed subsequently. In so doing we would separate the derivation of the model from the data more effectively than is currently achievable within EMD.

A key aspect of this approach is the ability to decompose any complicated data set into a finite and small number of intrinsic mode functions (IMFs), representing different scales of the original time series and physically meaningful modes. An IMF is defined as a function having the same number of extrema and zero crossings. It also has symmetric upper and lower envelopes defined by the local maxima and minima, respectively. EMD has been tested and validated exhaustively, though only empirically. Therefore, EMD has been applied to a wide-ranging data c.f. [21] [22], to analyse the variability. The aim here is to identify the internal components within the signal in the different datasets and to study their correlations. Further study will then intend to use these as the starting point for an iterative model.

This paper presents the initial findings with regard to EMD. We were able to show that the previous empirical findings that the number of functions (IMFs) required by EMD can be reduced from $\log_2 N$, where N is the number of observations, to a function of the number of observations:

$$nf = nv + \text{ round}\left(\frac{nv}{3}\right)$$

Eq. 1. Proposed new empirical equation for the number of IMFs required to decompose a signal.

where nf is the number of functions, nv is the number of variables. Variables refers to the number of distinct measurements entered into the model, alternatively these can be considered as inputs. In our case, we have four variables, air temperature, humidity, rainfall and barometric pressure.

This result was derived from the assumption that the relationship between the variables within the signal and the signal itself was some function of the number of variables over time. Essentially the assumption for every three "spatial" variables there must exist a function that represents the time relationship between the variables. It is important to note that this is not a direct analogy, however our assumption is that there is an interaction between the three spatial variables and time which needs to be represented in its own function. In effect it captures the interaction between the variables which is not part of the Intrinsic Mode Function of current focus.

In summary, the problem we are trying to solve is to produce a method of iteratively using an adaptive climate change model which integrates datasets representing the different spheres. We do not see this as replacing the existing climate models but rather as a useful tool to augment the existing mathematical models which allow for the prediction of climate. The aim would be to produce an adaptive short-term, in terms of the current climate timescales, forecasting model which is rapid and informative enough to contribute to public policy-making.

V. SIMULATION RESULTS

In this section we present the data and statistical results from the experiments running EMD with the new matrix for deriving the number of IMFs required.

Figure 2 shows the data for the first six months of 2015 for each of the variables, air temperature, barometric pressure and humidity. We have considered 4 variables including rainfall.

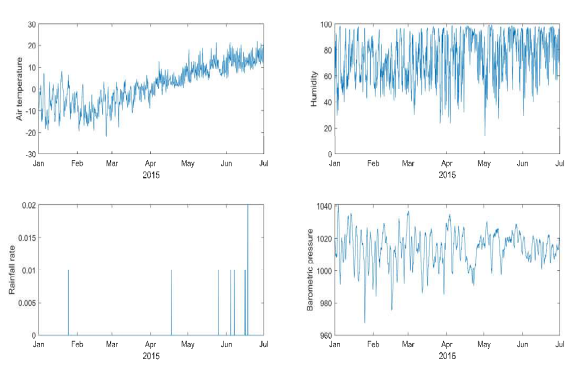

Fig. 2. Data from first six months of 2015 for each of the four variables.

As can be seen with reference to the rainfall data in Figure 2, the rainfall throughout the period is extremely minimal and will be omitted from the results presented below because of its lack of statistical significance.

The following figures demonstrate the results of applying EMD to the data whilst restricting the number of IMFs in accordance with the formula given above in Equation 1. The results presented represent the equivalencies in the comparable statistics.

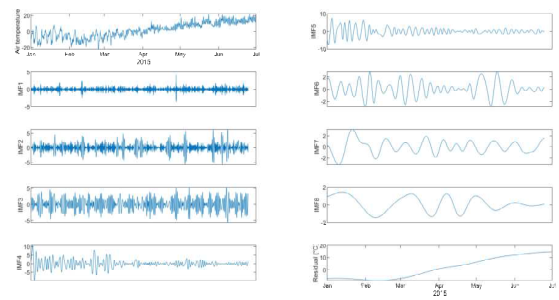

Fig. 3. Air temperature IMFs for the first six months of 2015

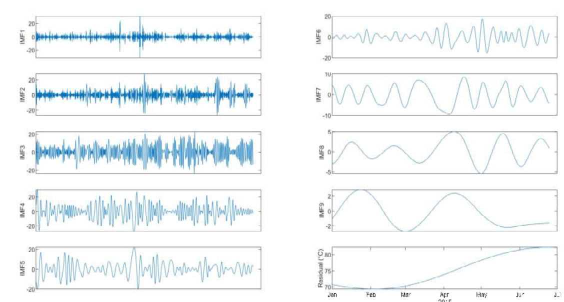

Fig. 4. Barometric pressure IMFs for the first six months of 2015.

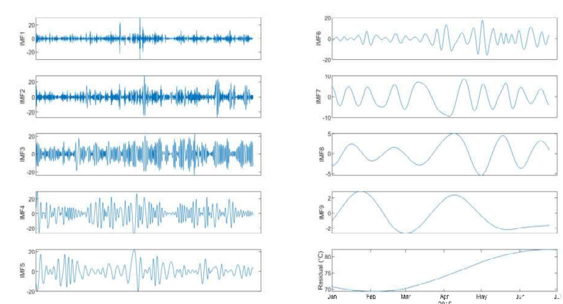

Fig. 5. Humidity IMFs for the first six months of 2015.

Respectively for temperature, Figure 3, barometric pressure, Figure 4 and humidity Figure 5. The above figures present the results of EMD with the new empirical measure for the number of IMFs required. According to these results that the first six IMFs identify the significant modes within the signal.

We argue that the remaining characteristic in the signal that is not captured by an IMF, excluding the residual, is the relationship between the other components of the signal over time. Further decomposition of the signal has the effect of modelling these interactions as if they were individual characteristics, which we would like to avoid in developing the model further. This is because we want to preserve some level of cost interaction within the modes.

Although the residual is very much larger the data represented by each IMF together captures the relevant characteristics of the underlying signal. This is important because it allows further development of a probabilistic model which is the subject of a further paper.

The following figures illustrate the statistical significance of each variable demonstrating that each mode is not identifying noise.

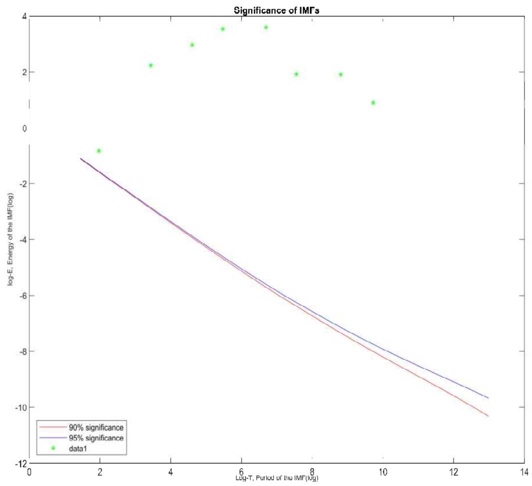

Fig. 6. Significance plot for air temperature, for the first six months of 2015.

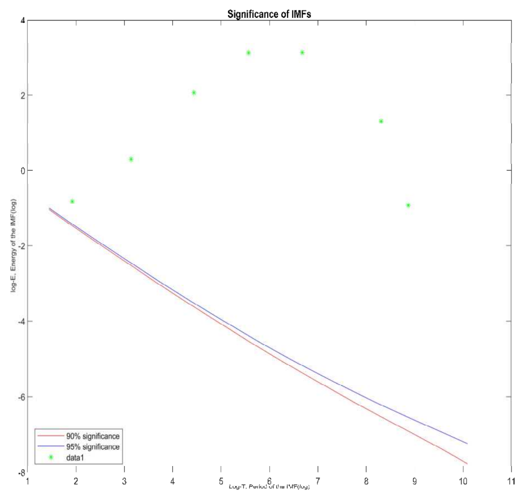

Fig. 7. Significance plot for humidity, for the first six months of 2015.

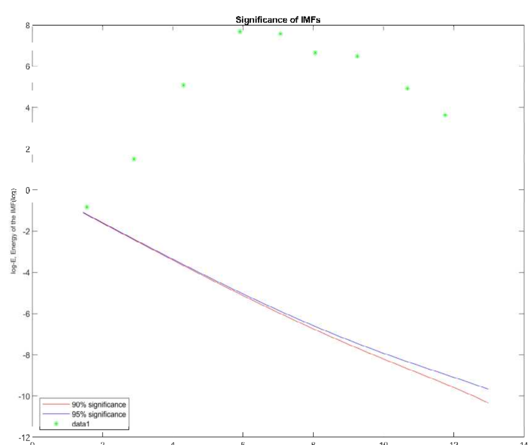

Fig. 8. Significance plot for barometric pressure, for the first six months of 2015.

VI. Conclusion

This paper presented the initial findings of research into the development of a probabilistic base model for localised climate modelling. Such a model would inform public decision-making and resource allocation regarding the effects of climate change in the locality. According to the findings presented in this paper, it is possible to reduce the number of IMF required for a decomposition, and further, to relate it to the number of variables rather than the number of observations. Our approach provides for a more robust approach to deriving the starting point for a subsequent iterative model. Currently, we are expanding this model into an iterated function system, which enables the relationships between each variable to be retained over time and a probabilistic framework for predicting localised climate models is proposed.

Acknowledgment

The authors gratefully acknowledge financial support from the Open University to present the paper at MTS/IEEE Oceans’22. Also, we would like to thank Prof. Schmitt (CNRS Log France), Prof. Budd (University of Bath), and Dr Okada (senior researcher at the Faculty of Wellbeing, Education and Language Studies, visiting Associate Professor in Brazil and Portugal) for the insightful discussions and useful suggestions.

References

[1] A. Chevuturi, A. G. Turner, S. J. Woolnough, G. M. Martin and C. MacLachlan, "Indian summer monsoon onset forecast skill in the UK Met Office initialized coupled seasonal forecasting system (GloSea5-GC2)," *Climate Dynamics,* vol. 52, no. 11, pp. 6599-6617, 2019.

[2] S. Brönnimann, "Impact of El Niño–southern oscillation on European climate," *Reviews of Geophysics ,* vol. 45, no. 3, 2007.

[3] IPCC, 2021, IPCC, 2021: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, V. [Masson-Delmotte, A. P. Zhai, S. Pirani, C. Connors, S. Péan, N. Berger, Y. Caud, L. Chen, M. Goldfarb, M. Gomis, K. Huang, E. Leitzell, J. Lonnoy, T. Matthews, T. Maycock, O. Waterfield, R. Yelekçi and B. Z. (. Yu, Eds., Cambridge University Press, 2021.

[4] P. Stott, "How climate change affects extreme weather events," *Science,* vol. 352, no. 6293, pp. 1517-1518, 2016.

[5] A. Ossó, R. H. E. Allan, L. Shaffrey and D. Maraun, "Emerging new climate extremes over Europe," *Climate Dynamics,* vol. 58, no. 1, pp. 487-501, 2022.

[6] R. Knutti and J. Sedláček, "Robustness and uncertainties in the new CMIP5 climate model projections," *Nature Climate Change,* vol. 4, no. 3, pp. 369-373, 2013.

[7] S. J. Vavrus, M. Notaro and D. J. Lorenz, "Interpreting climate model projections of extreme weather events," *Weather and Climate Extremes,* vol. 10, pp. 10-28, 2015.

[8] J. Schewe, S. N. Gosling, C. Reyer, F. Zhao, P. Ciais, J. Elliott, L. Francois, V. Huber, H. K. Lotze, S. I. Seneviratne, M. T. H. van Vliet, R. Vautard, Y. Wada, L. Breuer, M. Büchner, D. A. Carozza, J. Chang, M. Coll, D. Deryng, A. de Wit, T. D. Eddy, C. Folberth, K. Frieler, A. D. Friend, D. Gerten, L. Gudmundsson, N. Hanasaki, A. Ito, N. Khabarov, H. Kim, P. Lawrence, C. Morfopoulos, C. Müller, H. Müller Schmied, R. Orth, S. Ostberg, Y. Pokhrel, T. A. M. Pugh, G. Sakurai, Y. Satoh, E. Schmid, T. Stacke, J. Steenbeek, J. Steinkamp, Q. Tang, H. Tian, D. P. Tittensor, J. Volkholz, X. Wang and L. Warszawski, "State-of-the-art global models underestimate impacts from climate extremes," *Nature Communications,* vol. 10, no. 1, p. 1005, 2019.

[9] M. Déqué, S. Somot, E. Sanchez-Gomez, C. M. Goodess, D. Jacob, G. Lenderink and O. B. Christensen, "The spread amongst ENSEMBLES regional scenarios: regional climate models, driving general circulation models and interannual variability," *Climate Dynamics,* vol. 38, no. 5, pp. 951-964, 2012.

[10] I. Kenny, "An Evaluation of Performance Enhancements to Particle Swarm Optimisation on Real-World Data.," in *PhD Thesis*, 2016.

[11] I. Kenny, "Hydrographical Flow Modelling of the River Severn Using Particle Swarm Optimization," *The Computer Journal,* vol. 63, no. 11, pp. 1713-1736, 2020.

[12] D. K. B. Ismail, P. Lazure and I. Puillat, "Advanced Spectral Analysis and Cross Correlation Based on the Empirical Mode Decomposition: Application to the Environmental Time Series," *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS,* vol. 12, no. 9, pp. 1968-1971, 2015.

[13] A. Garralda-Guillem, H. Kunze, D. L. Torre and M. R. Galán, "Using the generalized collage theorem for estimating unknown parameters in perturbed mixed variational equations," *Communications in Nonlinear Science and Numerical Simulation,* vol. 91, p. 105433, 2020.

[14] H. Kunze, D. L. Torre, K. Levere and M. R. Galán, "Inverse Problems via the (Generalized Collage Theorem) for Vector-Valued Lax-Milgram-Based Variational Problems," *Mathematical Problems in Engineering,* 2015.

[15] F. Sévellec and S. S. Drijfhout, "A novel probabilistic forecast system predicting anomalously warm 2018-2022 reinforcing the long-term global warming trend," *Nature Communications,* vol. 9, August 2018.

[16] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung and H. H. Liu, "The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis," *Proceedings of the Royal Society. A, Mathematical, Physical, and Engineering Sciences,* vol. 454, no. 1971, pp. 903-995, 1998.

[17] J. A. Le, H. M. El-Askary, M. Allali, E. Sayed, H. Sweliem, T. C. Piechota and D. C. Struppa, "Characterizing El Niño-Southern Oscillation Effects on the Blue Nile Yield and the Nile River Basin Precipitation Using Empirical Mode Decomposition," *Earth Systems and Environment,* vol. 4, no. 4, pp. 699-711, 2020.

[18] D. Kbaier, "Advanced statistical analysis of environmental data in the Gascoyne Inlet," San Diego, 2021.

[19] T. MIYAKAWA, S. YOKOI, N. HIROTA, S. HIRAHARA, T. INOUE and Y. N. TAKAYABU, "An Iterative Selection Method: A Summer Eastern Asian Metric for CMIP3 Climate Models as an Example," *Journal of the Meteorological Society of Japan,* vol. 91, no. 3, pp. 305-21., 2013.

[20] V. M. Gryanik, C. Lüpkes, D. Sidorenko and A. Grachev., "A Universal Approach for the Non-Iterative Parametrization of Near-Surface Turbulent Fluxes in Climate and Weather Prediction Models," *Gryanik, V. M, C. Lüpkes, D. Sidorenko, and A. Grachev. "A Universal Approach for the Non-Iterative Parametrization of Near-Surface Turbulent FluJournal of Advances in Modeling Earth Systems,* vol. 13, no. 8, p. N/a, 2021.

[21] N. E. Huang and N. O. Attoh-Okine, The Hilbert-Huang Transform in Engineering, 1 ed., Boca Raton: CRC Press, 2005.

[22] N. Huang and S. Shen, "Hilbert-Huang Transform and Its Applications [M]," *Interdiscip. Math. Sci.,* vol. 5, 2005.