Using Machine Learning and Bifurcation Theory to Predict Heat Waves

Author: Miles Kent, **Editor:** Russell Milne

February 14, 2023

1 Overview

Heat waves have the potential to wreak havoc on infrastructure, human life, and the functioning of local ecosystems (Wang et al. 2020). Today's most up-to-date climate models offer a high-resolution picture of the local and global climate, but these models are often very complicated, requiring computationally intensive simulation to generate numerical solutions and predictions (Crucifix & Rougier, 2009). Although these models offer high spatial resolution, it is still notoriously difficult to predict the temporally abrupt transitions in temperature that define a heat wave (Wang et al. 2020). An alternative approach to modeling and predicting these extreme temperature transitions involves using the mathematical theory behind dynamical systems. A dynamical system consists of one or more ordinary differential equations that mathematically represent how particular quantities, for instance temperature or other weather indices, change over time. When parameters in the system change, it is possible for a dynamical system to undergo a bifurcation, in which the stability of its equilibrium points change. This can alter the trajectories of the quantities being modelled by the dynamical system, causing phenomena like large jumps or oscillations. Because of the potential for bifurcations to cause unexpected and disruptive changes, early warning signals have been developed so that upcoming bifurcations can be spotted before stability changes occur, based on patterns in data (Bury et al. 2021). There is an abundance of publicly accessible local climate data from many cities around the world, and hence previously-observed heat waves will be detectable in this data. Since the same transitions in time series data that characterize heat waves also characterize bifurcations in dynamical systems, the approach pioneered by Bury et al. that used machine learning to construct early warning signals for bifurcations in paleoclimate data can also be used with heat waves and similar extreme weather events.

2 Project Motivation

Motivation for this project comes from the physics behind heat waves, which describes them in terms similarly to how a bifurcation can happen in a multidimensional dynamical system, and two recent studies that serve as a proof of concept that critical transitions can be identified in climate data.

2.1 Physics of a Heat Wave

A heat wave occurs when local temperature quickly increases to levels significantly above what is expected and remains there for an extended period. This is thematically similar to what happens during a bifurcation; furthermore, the changes in atmospheric and environmental properties that drive heat waves are also analogous to changes in dynamical system parameters that cause bifurcations. Heat waves are caused when a high-pressure system becomes stagnant over a region, creating an environment with no wind. Due to the high pressure, air is forced downward, which acts as a barrier preventing warm air from rising. This phenomenon suppresses precipitation, which allows the soil to dry because the warm air cannot rise high enough to condensate. The result of these environmental conditions is a spike in temperature, which creates a stagnant warm pocket of hot air at ground level with little temperature fluctuation. This is the heat wave. This elevated local temperature state with little fluctuation is correlated with a distinct combination of conditions: high pressure, low wind speed, no precipitation, and no cloud cover. The shift to elevated temperature levels caused by these conditions closely resembles a bifurcation, meaning that early warning signals that have been developed to provide advance notice of bifurcations could also be used to predict this shift.

2.2 Proof of Concept

Two recent papers showed that early warning signals, including both classical ones (variance, autocorrelation) and ones generated by machine learning techniques, can predict critical transitions in climate data. Bury et al. (2021) showed that machine learning can be used to provide early warning signals of bifurcations in noisy paleoclimate data, by training classifiers on time series of many different dynamical systems undergoing bifurcations and then feeding real-world time series into these classifiers. In addition, Wang et al. (2020) were able to use classical early warning signals to identify some recent heat waves using data from prior to their onset, as well as uncover critical transitions in broader-scale climate data. Both of these papers together serve as a proof of concept that early warning signals, and the general framework of bifurcations in dynamical systems, can be used to predict heat waves within climate data, and that machine learning can be used to do this efficiently and accurately.

3 Research Question and Project Breakdown

The purpose of this project is to determine if early warning signals based on machine learning and a multidimensional dynamical system that incorporates the physics of a heat wave can be used to predict heat waves in a wide variety of climates. To answer this research question, the project will be broken into three parts to ensure it is successfully completed.

3.1 Data collection and organizing

The student will be tasked with collecting daily and hourly climate data from the NOAA ISD database, which will include temperature, barometric pressure, cloud cover, wind speed, and dew point from three hundred cities across a broad range of global locations. The student will then filter and organize the data in R. Specifically, this task will include organizing the data into unbroken time series of 60 observations (for daily data) and 500 observations (for the data at three-hour intervals), which will be used as inputs for machine learning classifiers as detailed below. These specific time series lengths both represent approximately two months of data, long enough to ensure that heat waves are detected ahead of time and provide enough data to the classifiers.

3.2 Model identification and creating training data

To create training data for the machine learning algorithm, time series will be generated in which dynamical systems that are known to exhibit bifurcations are simulated with time-varying values for the parameters that cause the bifurcations. This will induce specific identifiable bifurcations. The student will also generate time series that do not have any bifurcations, to produce a null option for the machine learning classifiers described below. Furthermore, the student will use his physics background and previous studies to identify a low-dimensional dynamical system model that can represent local weather patterns such as those in the NOAA ISD data mentioned above. Specifically, this model will be obtained from the data by using the Sparse Identification of Nonlinear Dynamics algorithm (Brunton et al. 2016), and should feature critical temperature transitions when dew point, cloud cover, barometric pressure, and wind speed reach or are near certain values. The student will select a model in which some of these quantities are expressed as state variables and some are expressed as time-varying parameters. This will ensure that the model will be complex enough to be capable of exhibiting the relevant bifurcations, and have relevant parameters that would cause the bifurcations when varied.

3.3 Prediction algorithm development and testing

The student will be tasked with using a previous study (Bury et al. 2021), and the Python code contained therein, to create classifiers to be used in the machine learning algorithm. In particular, the student will modify the algorithm that Bury et al. 2021 produced early warning signals with, so that it will reflect the multidimensional nature of the interacting weather components that produce heat waves. These classifiers will be trained based on the time series generated in part two. Following this, the student will then test the algorithm on the data obtained in part one of the project to verify the method's accuracy and predictive power, as well as compare it to other heat wave detection methods using techniques such as ROC curves. The student will also test the accuracy of the novel dynamical system model constructed in part two, by determining if it can replicate out-of-sample weather data obtained from the same NOAA ISD dataset.

4 Relevance

This research is relevant to both applied mathematics and climate modeling. This is because it represents a unique, cutting-edge application of bifurcation theory for gaining insights into extreme temperature transitions and climate trends. If successful, this project will yield a novel method for predicting heat waves in any given city where local weather data is available. This method would require much less computational power than, and ideally offer more predictive power than, today's most common techniques for predicting temperature transitions. As a result, the output of this project will open the door for future research that is both innovative and mathematically sound. The end results of this project, if everything works as intended, will also include a low-dimensional dynamical system that can accurately represent climate data. This system would be easy to analyze and interpret, without sacrificing much in the way of real-world predictive power. This would pave the way for future research to understand how this model fully behaves and its relevance for atmospheric sciences.

References

[1] Crucifix, M., Rougier, J. (2009). *On the use of simple dynamical systems for climate predictions.* The European Physical Journal Special Topics, 174(1), 11-31.

[2] Bury, T.M., Sujith R. I., Pavithran I., Scheffer M., Lenton T.M., Anand M., Bauch C.T. (2021). *Deep learning for early warning signals of tipping points.* Proceedings of the National Academy of Sciences of the United States of America. 118(39).

[3] Wang, C., Wang, Z.-H., Sun, L. (2020). *Early-warning signals for critical temperature transitions.* Geophysical Research Letters, 47

[4] Brunton, S. L., Proctor, J. L., Kutz, J. N. (2016). *Discovering governing equations from data by sparse identification of nonlinear dynamical systems.* Proceedings of the national academy of sciences, 113(15), 3932-3937.