Approaches to Boulder's Extreme Wind Events with Chaotic Spatiotemporal Data

Wisang Sugiarta
Dept. of Computer Science
University of Colorado - Boulder

1 Introduction

The study of dynamical systems has been a central theme in scientific research since the late nineteenth century, originating with Henri Poincaré's investigations into celestial mechanics [1]. Dynamical systems encompass a wide range of physical phenomena, from simple pendulums to complex, high-dimensional models of the atmosphere. Poincaré and subsequent researchers sought to understand how the state of such systems evolves over time, typically described by differential equations (DEs). These equations provide a mathematical foundation for modeling system behavior, enabling simulations that offer insights into future states. Early studies soon revealed that many systems, despite exhibiting clear underlying patterns, could diverge dramatically in response to small perturbations in initial conditions. This sensitivity to initial conditions became a hallmark of chaotic systems and was later popularized by meteorologist Edward Lorenz, who introduced the concept of the "butterfly effect" in his seminal work [2].

The modern understanding of weather as a chaotic system originates from Lorenz's seminal work [3]. While physicists and atmospheric scientists have traditionally focused on deriving explicit differential equation (DE) models to describe such systems, the explosion of high-resolution, easily accessible data in recent years has led to the development of a broad range of data-driven modeling techniques.

Modeling chaotic systems, however, presents challenges, as their sensitive dependence on initial conditions imposes severe limitations on predictability. Traditional approaches, outlined in [4, 5, 6], explicitly address these challenges by adapting tools from nonlinear time-series analysis. More recently, exciting advances such as reservoir computing have demonstrated remarkable success in modeling chaotic spatiotemporal systems [7, 8].

Extreme wind events in Boulder, Colorado, provide a compelling local expression of spatiotemporal chaos, arising from nonlinear interactions among synoptic-scale circulations, terrain turbulence, and local boundary layer dynamics. These events not only pose significant forecasting challenges but also have important practical implications: improving predictive capabilities could aid in mitigating risks associated with wildfires, power outages, and public safety threats. Motivated by this context, this project investigates whether different modeling frameworks—specifically, classical chaotic time-series analysis methods and modern reservoir computing architectures—can effectively forecast these extreme events and recover structure within the underlying dynamics. The high-level goal is to assess the relative strengths and limitations of these approaches in capturing the essential features of chaotic atmospheric phenomena.

2 Methodology

The goal of the study is to optimize algorithms to best predict extreme winds. From [5], Local Linear Factor and Method of Analogues will be tested for experimentation with extreme winds. The RC approach outlined in [7] will also be used. This section will provide descriptions of the methodology used.

2.1 Data Description

We chose to pursue a data-driven approach, where a reliable, robust dataset is important. We decided to select the National Center for Atmospheric Research (NCAR)[9] Boulder weather site which may be found at

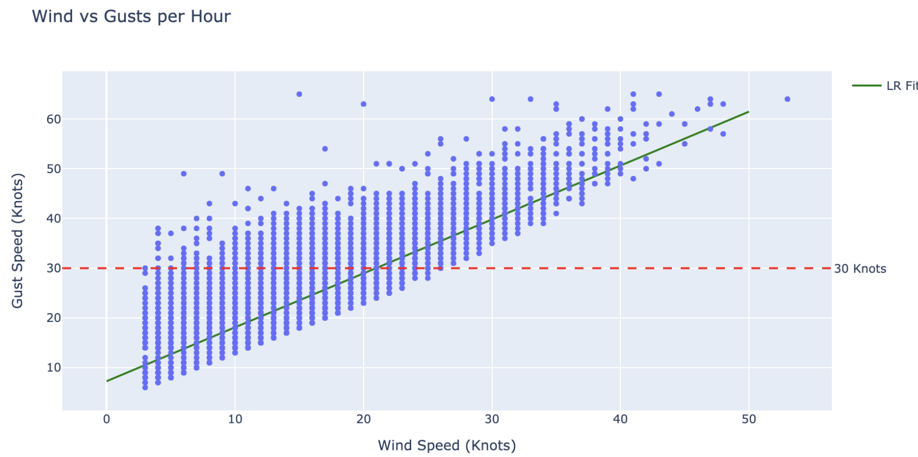


Figure 1: Linear regression fit of gust factor and wind speed.

`https://archive.eol.ucar.edu/cgi-bin/weather.cgi?site=ml&units=english`

It is a reliable hourly dataset that dates back to 2000. It is, however, a simple dataset with basic features including wind speed and gust data. The motivation of this study is to create an optimized model and thus previous research [8, 6] shows that supplemental atmospheric data can augment predictive performance significantly.

We use synoptic-scale pressure-level data over western North America to supplement the dataset. We decided to use the ERA5 [9] single pressure level atmospheric data which may be found at

`https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels?tab=overview`

The complete list of features is provided in the appendix. Although the ERA5 reanalysis product offers atmospheric variables at a spatial resolution of 0.25° × 0.25°, this study restricts input to full-degree grid points to reduce computational overhead. Despite this coarser spatial sampling, the inclusion of ERA5 data enables a regionally informed representation of Boulder's wind dynamics, augmented by synoptic-scale atmospheric features.

A notable limitation of the NCAR surface weather dataset is the presence of missing gust observations, with approximately 28% of gust values unavailable. To address this, we introduce a derived feature termed the *gust factor*, defined via a linear regression model:

$$\text{Gust Factor} = \alpha \cdot \text{WS} + \beta$$

where WS denotes wind speed, and α, β are regression coefficients estimated from the subset of the data for which both wind speed and gust measurements are available. This model achieves an R^2 score of 0.79, indicating a strong linear relationship. For time steps with missing gust values, the gust factor is used to impute estimates, while observed gust values are retained wherever present.

As the dataset created is already high dimensional ($D \geq 1000$), there will be no time-delayed embedding.

2.2 Local Linear Forecasting

Local linear forecasting is a classical approach for predicting the evolution of chaotic systems based on reconstructed state-space dynamics. Rather than fitting a global model, this method exploits the assumption that the system's behavior can be well-approximated by a linear model within small neighborhoods of the attractor.

To implement local linear forecasting, for a given query point, its nearest neighbors are identified using Euclidean distance. A linear regression is then locally fit to predict the future evolution.

Formally, let $\mathbf{x}(t)$ denote the state. For each prediction, the k nearest neighbors of $\mathbf{x}(t)$ are identified, and a local linear model is constructed to minimize the prediction error:

$$\hat{y}(t + \Delta t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}$$

where $\mathbf{A}$ and $\mathbf{b}$ are learned from the local neighborhood. This approach allows for capturing the locally linear structure of the otherwise globally nonlinear attractor.

2.3 Method of Analogues

The method of analogues, originally proposed by Lorenz, provides an intuitive yet powerful framework for forecasting chaotic systems by leveraging historical similarity. Instead of constructing an explicit model, predictions are made by identifying past states of the system that are close to the current state and assuming that their subsequent evolution can inform future behavior.

Given a current state $\mathbf{x}(t)$, the method searches for past states $\{\mathbf{x}(t_i)\}$ that are within a specified distance threshold or among the $k = 3$ nearest neighbors. The future state is then predicted based on the observed evolution of the most similar past instance:

$$\hat{y}(t + \Delta t) = y(t_i + \Delta t)$$

where t_i is the time index of the closest analogue.

Nearest neighbors are identified using Euclidean distance in the data, and both single-analogue and weighted-average predictions are tested. This approach provides a nonparametric baseline for comparison against more complex machine learning models like RC.

2.4 Reservoir Computing

RC is a machine learning framework particularly well-suited for modeling nonlinear and chaotic dynamical systems. It consists of a fixed, high-dimensional recurrent neural network (the reservoir) that projects input signals into a rich dynamic space, followed by a trainable linear readout that maps the reservoir state to the desired output. This architecture enables efficient training, as only the readout weights are optimized, while the reservoir remains untrained.

In this study, RC is applied to forecast extreme wind events in Boulder, Colorado, using features derived from local meteorological observations. Each input to the reservoir consists of a delay-embedded representation of the time series, capturing both current and historical states of atmospheric variables.

The reservoir dynamics are governed by the update equation:

$$\mathbf{r}(t + 1) = \tanh\left(W_{in}\mathbf{u}(t) + W_r\mathbf{r}(t)\right)$$

where $\mathbf{r}(t)$ is the reservoir state, $\mathbf{u}(t)$ is the input vector, W_{in} and W_r are fixed input and recurrent weight matrices, respectively. The output prediction is computed as:

$$\hat{y}(t) = W_{out}\mathbf{r}(t)$$

where W_{out} is learned via linear regression.

Reservoir computers were selected for their proven success in chaotic systems and efficient training compared to fully trainable recurrent networks [8]. Detailed performance metrics are discussed in 3.

3 Results

This section presents the results of applying the three algorithms to the task of forecasting extreme wind events in Boulder, Colorado. We analyze aspects of model performance across a range of forecast horizons and feature configurations. The predictive task is framed as a binary classification problem: determining whether an extreme wind event—defined as a gust exceeding 35 knots—will occur within a specified future window.

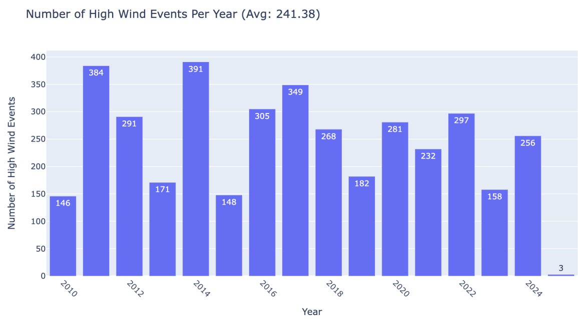
3.1 Temporal Distribution of Wind Events

We first examine the temporal distribution of extreme wind events to provide context for the forecasting task. Figure 2 shows distributions across year, month, temperature, and flow direction. Events occur most frequently during winter months, are associated with warmer conditions, and predominantly occur under easterly flow patterns. These findings provide insight into the atmospheric features that may be most predictive.

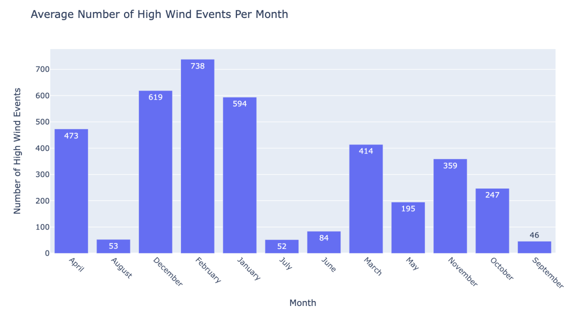
3.2 Predictive Performance: Model Comparison

Table 1 presents the predictive performance of MoA, LLF, and RC at a 12-hour forecast horizon (±30 minutes).

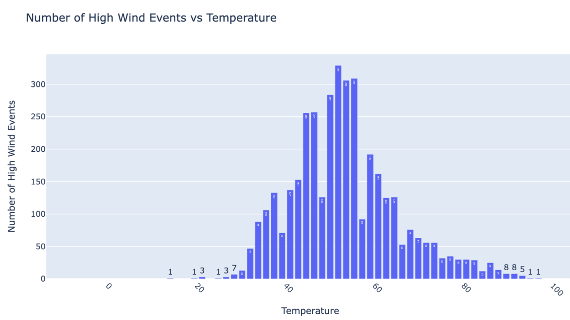
Reservoir Computing achieved strong performance, with a sensitivity of 0.76, positive predictive value (PPV) of 0.83, and a positive-class F1 score of 0.79. The negative-class F1 score of 1.00 indicates perfect identification of non-events.



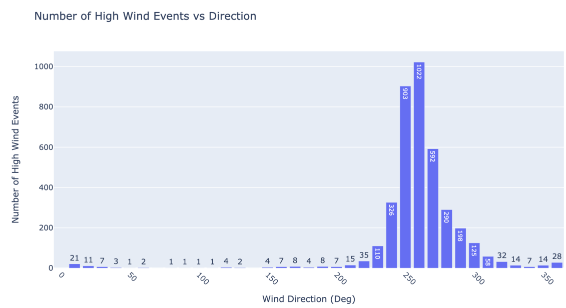
(a) Wind events per year.



(b) Wind events per month.



(c) Wind events per temperature.



(d) Wind events per flow direction.

Figure 2: Distributions of important features in high wind events.

These results suggest that the RC model captures sufficient memory and dynamics to anticipate the onset of high wind conditions with relatively high confidence.

In contrast, both MoA and LLF performed poorly, with positive-class F1 scores below 0.1, indicating that these traditional methods were unable to capture the complexity of the system dynamics at even short forecast horizons.

Metric	Method of Analogues	Local Linear Forecasting	Reservoir Computing
Sensitivity	0.04	0.13	0.76
PPV	0.09	0.02	0.83
Positive F1 Score	0.07	0.03	0.79
Negative F1	0.99	0.99	0.99

Table 1: Comparison of 12-hour forecasting performance for MoA, LLF, and RC.

3.3 Time Horizons with Reservoir Computers

Given the extremely poor performance of MoA and LLF at short horizons, these methods were not extended to longer forecast horizons. The observed limitations highlight the challenges of applying traditional chaotic forecasting methods to high-dimensional, spatiotemporally complex atmospheric systems. Therefore, only RC models were evaluated across multiple forecast ranges.

Table 2 shows RC model performance from short-term to longer-term horizons.

3.4 Forecast Horizon Sensitivity

As expected for chaotic systems, predictive performance deteriorated with increasing forecast horizon. RC models maintained strong predictive skill up to approximately 72 hours, but accuracy declined sharply beyond that range, with positive-class F1 scores falling below 0.20 at one week and beyond. This trend is consistent with the known exponential error growth in chaotic dynamical systems.

Time Horizon	Sensitivity	PPV	Positive F1	Negative F1
Random	0.51	0.14	0.22	0.64
12h	0.76	0.83	0.79	1.00
48h	0.86	0.81	0.83	0.99
72h	0.87	0.89	0.88	0.99
1 week	0.40	0.09	0.15	0.13
1 month	0.40	0.12	0.19	0.67

Table 2: Reservoir Computing performance across forecast horizons.

The decline in PPV at long horizons further suggests that the model becomes overconfident in flagging events, even when predictive reliability drops. These results highlight the natural predictability limits inherent to chaotic weather systems.

3.5 Practical Implications

The ability to accurately forecast extreme winds up to 72 hours in advance has significant implications for wildfire prevention, emergency response, and power infrastructure management. For example, during recent years, Xcel Energy has implemented proactive power shutoffs in Colorado during forecasted high-wind events to mitigate wildfire risks. Reliable short-term prediction could enhance these efforts, reducing the impacts of extreme weather.

While RC models offer computational efficiency and strong short-term performance, their limited interpretability and reduced reliability at longer horizons suggest the need for future hybrid approaches. Incorporating physical constraints or developing physics-informed reservoir computing (PINN-RC) architectures may further improve robustness and operational trust.

These results motivate a deeper examination of the strengths, limitations, and broader implications of the different forecasting approaches, which we discuss in the following section.

4 Discussion

Reservoir computing proved to be an effective algorithm for forecasting extreme wind events in Boulder, particularly for short-term time horizons. However, the model's strengths must be weighed against its major limitation: lack of interpretability.

One of the key trade-offs in this study is between predictive power and model explainability. Reservoir computers excel at learning nonlinear temporal patterns without requiring explicit physical modeling, but they operate as a black-box. Unlike classical methods such as local linear forecasting or phase-space reconstruction, reservoir models do not provide direct access to interpretable dynamical quantities, such as attractor geometry or Lyapunov exponents. Although the RC may capture these structures within its high-dimensional state space, they are not recoverable for analysis. This limitation raises a question: can black-box models be used by meteorologists operationally?

The poor performance of Method of Analogues and Local Linear Forecasting highlights the challenges classical chaotic forecasting methods face in high-dimensional, spatiotemporal complex systems. Both approaches rely on finding similar past states via nearest neighbor algorithms, but in high-dimensional spaces, the task is much harder. Extreme wind events are relatively rare, introducing class imbalance. K-nearest neighbor algorithms are particularly sensitive to such imbalance, often biasing toward predicting the majority class [10]. These factors likely explain the poor F1 scores for MoA and LLF.

While the reservoir model exhibited strong performance for 12–72 hour forecasts, it declined sharply beyond that, consistent with known predictability limits in chaotic systems. Although loosening the prediction window can artificially inflate performance metrics by relaxing precision requirements, it does not address the deeper challenge of accurately capturing system dynamics. As such, future research should look into enhancing interpretability over longer prediction windows.

Feature selection emerged as another key factor. Incorporating thermodynamic and synoptic-scale variables—such as temperature, pressure, and humidity—significantly improved forecasting skill. These variables likely encode precursor conditions for downslope wind events, including cold-air damming, lee-side pressure gradients, and boundary layer instability [11]. This finding reinforces the importance of embedding physically meaningful signals into machine learning models.

Ultimately, this study highlights the promise—and the current limitations—of reservoir computing in chaotic, spatiotemporal forecasting. It provides strong evidence for its use in short-term prediction but also motivates future exploration into hybrid architectures, interpretable dynamical embeddings, and physically informed learning models.

5 Conclusion

This study explored the challenges and opportunities of forecasting extreme wind events in Boulder, Colorado, using both classical and modern approaches to chaotic time-series prediction. Traditional techniques such as MoA and LLF struggled to capture the complex dynamics of the system, achieving poor predictive performance even at short forecast horizons. Their limitations were attributed to high dimensionality and sensitivity to class imbalance, which reduced nearest-neighbor searches.

In contrast, RC demonstrated strong predictive skill, particularly for short-term horizons up to 72 hours. RC models successfully captured nonlinear spatiotemporal patterns. However, this success came with a trade-offs with interpretability and raises issues for deployment.

Nonetheless, several limitations warrant future research. Improving model interpretability remains critical, especially for operational meteorological applications. Future work could explore hybrid architectures that combine reservoir computing with physical constraints, or develop physics-informed reservoir models that better align machine learning predictions with known dynamical structures. Extending predictive skill beyond the 72-hour horizon, while respecting the inherent unpredictability of chaotic systems, also remains an open challenge.

In sum, this study highlights both the promise and the challenges of applying modern machine learning tools to the prediction of chaotic atmospheric physics. It motivates continued development of forecasting models that balance predictive power with physical insight and understanding.

References

[1] H. Poincaré. Les méthodes nouvelles de la mécanique céleste, Jan 1899.

[2] Edward N. Lorenz. Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, Mar 1963.

[3] Bo-Wen Shen, Roger A. Pielke Sr., Xubin Zeng, Jong-Jin Baik, Sara Faghih-Naini, Jialin Cui, and Robert Atlas. Is weather chaotic?: Coexistence of chaos and order within a generalized lorenz model, Jan 2021.

[4] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*. Cambridge University Press, 2005.

[5] Elizabeth Bradley and Holger Kantz. Nonlinear time-series analysis revisited. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(9), April 2015.

[6] Kerry A. Emanuel. An air-sea interaction theory for tropical cyclones. *Journal of the Atmospheric Sciences*, 43(6):585–605, Mar 1986.

[7] Jaideep Pathak, Brian Hunt, Michelle Girvan, Zhixin Lu, and Edward Ott. Model-free prediction of large spatiotemporally chaotic systems from data: A reservoir computing approach. *Physical Review Letters*, 120(2), Jan 2018.

[8] Daniel J. Gauthier, Erik Bollt, Aaron Griffith, and Wendson A. S. Barbosa. Next generation reservoir computing. *Nature Communications*, 12(1), September 2021.

[9] Nsf ncar mesa laboratory. `https://archive.eol.ucar.edu/cgi-bin/weather.cgi?site=ml&units=english`. Accessed: 2025-04-20.

[10] Haibo He and Edwardo A Garcia. Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9):1263–1284, 2009.

[11] Andrew E. Mercer, Michael B. Richman, Howard B. Bluestein, and John M. Brown. Statistical modeling of downslope windstorms in boulder, colorado. *Weather and Forecasting*, 23(6):1176–1194, Dec 2008.