

Data Science Approaches for Identifying Climate Variability Patterns From Historical Environmental Records

Dr. Lucas Anderson

Associate Professor, Ottawa Institute of Sustainable Data Science, Canada

Abstract

Climate variability represents an important component of environmental systems because temperature, precipitation, and related atmospheric conditions fluctuate across different temporal and geographical scales. Historical environmental records provide opportunities to examine these variations and identify patterns that may influence ecological conditions, resource availability, and environmental planning. However, the analysis of long-term climatic information is complicated by incomplete observations, measurement inconsistencies, seasonal fluctuations, geographical differences, and nonlinear relationships among environmental variables. Data science approaches provide computational and statistical methods for organizing historical records, examining temporal relationships, and identifying meaningful patterns within complex environmental information. The present paper develops a hypothetical methodological framework for investigating climate variability through historical temperature and precipitation records. The proposed research design incorporates environmental data preprocessing, climatological anomaly calculation, time-series analysis, statistical trend assessment, and predictive modeling. Regularized regression, Random Forest, and Gradient Boosting are considered alongside a seasonal baseline method.

Two illustrative analytical tables describe hypothetical temperature and precipitation anomalies across different historical periods and comparative prediction errors among selected data science techniques. Mean absolute error and root mean squared error are used to demonstrate the evaluation of temperature and precipitation predictions. All numerical values are manually constructed and do not represent measured environmental observations, executed computational experiments, or verified climate trends. The methodological discussion examines data completeness, temporal dependence, measurement uncertainty, model complexity, and the interpretation of climatic patterns. Particular attention is given to distinguishing long-term climate variability from short-term fluctuations and separating predictive performance from causal climate attribution. The paper concludes that data science offers a useful methodological foundation for analyzing historical environmental information when supported by dependable observations, appropriate statistical procedures, and reproducible evaluation.

Keywords: Climate Variability, Data Science, Historical Environmental Records, Climate Data Analytics, Temperature Trends, Precipitation Variability, Time-Series Analysis, Environmental Modeling, Machine Learning, Climate Pattern Identification.

1. Introduction

Climate variability describes changes in atmospheric and environmental conditions occurring across different periods and geographical regions. Temperature, precipitation, humidity, and other climatic variables exhibit natural fluctuations associated with interacting physical processes. Historical environmental records provide valuable information for examining these fluctuations and understanding how climatic conditions vary over time. Meteorological observations collected through monitoring stations and environmental measurement systems support investigations of seasonal behavior, interannual differences, and longer-term trends. The increasing availability of historical records has created opportunities to apply computational methods to climate research and environmental information analysis.

Traditional climate analysis frequently relies on descriptive statistics, temporal averaging, trend estimation, and comparisons among historical observation periods. These procedures provide an important foundation for identifying changes in environmental conditions. However, climatic datasets often contain complicated relationships that extend beyond simple linear variation. Temperature and precipitation may respond differently to seasonal conditions, regional geography, and broader atmospheric processes. Moreover, climatic observations collected at different stations may exhibit varying levels of completeness and measurement consistency. Such characteristics require analytical approaches capable of examining temporal patterns while accounting for uncertainty and differences in observational quality.

Data science introduces computational techniques for processing, organizing, and interpreting extensive environmental datasets. Statistical learning, time-series modeling, data visualization, and predictive analytics provide complementary methods for examining climatic information. These approaches can support the identification of recurring seasonal patterns, unusual observations, and relationships among environmental variables. Machine learning algorithms may also represent nonlinear relationships that are difficult to describe through conventional statistical models. Nevertheless, the effectiveness of these methods depends on appropriate data preparation, representative observations, and evaluation procedures that account for the temporal structure of climate records.

Historical temperature and precipitation records are particularly relevant to climate variability analysis because they describe important characteristics of atmospheric conditions. Temperature observations can reveal differences in average thermal conditions across periods, while precipitation records provide information about moisture availability and rainfall variability. However, changes in these variables do not necessarily follow identical temporal patterns. A period characterized by higher average temperature may exhibit increasing, decreasing, or irregular precipitation. Consequently, climatic variables should be examined individually and jointly, with careful attention to their physical interpretation and the limitations of available environmental observations.

The present paper examines data science approaches for identifying climate variability patterns through a hypothetical methodological investigation. It considers historical environmental records represented by monthly temperature and precipitation measurements across six fictional monitoring stations. The proposed framework combines climate anomaly calculation, descriptive trend analysis, regularized regression, Random Forest, and Gradient Boosting. Two illustrative analytical tables demonstrate how temporal differences and predictive model performance could be reported. No actual meteorological observations were collected or analyzed, and the numerical examples are included exclusively to demonstrate a potential research methodology.

2. Objectives of the Study

The principal objective of this study is to examine the methodological application of data science techniques for identifying climate variability patterns within historical environmental records. The investigation focuses on temperature changes, precipitation fluctuations, predictive modeling, and analytical reliability. Three specific objectives establish the conceptual direction of the proposed research framework.

2.1 Examination of Historical Climatic Patterns

The first objective is to investigate how historical environmental observations can be organized to identify recurring and changing climatic patterns. Monthly temperature and precipitation records provide information about seasonal conditions, interannual variability, and differences among historical periods. The proposed study considers how descriptive statistics, standardized anomaly calculations, and temporal comparisons could be used to examine these characteristics. Particular attention is given to distinguishing ordinary seasonal fluctuations from patterns that may persist across extended observation periods.

The investigation further aims to examine how temporal aggregation influences the interpretation of environmental information. Monthly observations may contain short-term fluctuations that become less apparent when expressed through annual or multiyear averages. However, excessive aggregation can also conceal important variation. Consequently, the proposed framework considers several temporal scales and emphasizes consistent reference periods. This objective seeks to establish a structured analytical procedure for examining climate variability without assuming that every observed difference represents a statistically significant long-term trend.

2.2 Comparative Assessment of Predictive Data Science Models

The second objective focuses on comparing data science approaches for predicting temperature and precipitation values from historical environmental observations. A seasonal baseline, regularized regression, Random Forest, and Gradient Boosting are selected to represent contrasting prediction strategies. Their proposed evaluation considers the magnitude of prediction errors under consistent data preparation and testing conditions. The investigation seeks to establish whether nonlinear methods could offer methodological advantages when environmental variables exhibit complex relationships across seasons and observation periods.

A further objective involves examining the relationship between model complexity and predictive reliability. Statistical regression provides comparatively transparent parameter relationships, whereas ensemble learning methods can represent nonlinear interactions among environmental features. However, increasingly complex algorithms may require more observations and extensive parameter optimization. The proposed study therefore emphasizes fair computational comparison, appropriate temporal validation, and the independent evaluation of prediction errors. This objective encourages model selection based on defensible performance evidence rather than assumptions regarding the superiority of sophisticated computational techniques.

2.3 Identification of Environmental Data Analysis Constraints

The third objective is to identify methodological difficulties that may influence the analysis of historical environmental records. Missing observations, inconsistent measurement practices, changes in monitoring equipment, and geographical differences can complicate climatic interpretation. Temporal

autocorrelation and seasonal structure may also influence statistical testing and predictive model assessment. The proposed study considers how quality control, data standardization, and appropriate validation procedures could address selected difficulties while preserving meaningful information about environmental variability.

The investigation also aims to establish analytical safeguards for interpreting climate-related findings. Predictive associations among environmental variables do not independently establish the physical causes of observed changes. Similarly, favorable prediction errors do not demonstrate that a model can reliably represent future climatic conditions outside its evaluation period. This objective therefore emphasizes uncertainty assessment, transparent documentation, and careful differentiation between descriptive patterns, predictive estimates, and causal interpretations. These considerations support responsible data science applications in historical climate research.

3. Review of Literature

3.1 Development of Historical Climate Observation Datasets

Jones and Moberg (2003) examined large-scale surface temperature variations through revisions to historical land-based observations. Mitchell and Jones (2005) developed procedures for constructing monthly climate datasets and high-resolution environmental grids. Brohan et al. (2006) investigated uncertainty in historical temperature estimates and introduced improvements to a global observational dataset. These studies established important methodological foundations for handling long-term climate measurements. Their contributions emphasize the relevance of data quality, station coverage, measurement consistency, and uncertainty estimation when examining climatic variability.

Smith et al. (2008) investigated improvements to merged land–ocean surface temperature analysis, while Hansen et al. (2010) examined long-term global surface temperature changes using observational information. Morice et al. (2012) developed an approach for representing uncertainty within historical temperature estimates. Harris et al. (2014) subsequently described updated high-resolution grids of monthly climate observations. Together, these contributions demonstrate the importance of systematic environmental data construction. They also indicate that historical climatic interpretation depends on how observations are processed, combined, and assessed for uncertainty.

3.2 Statistical Analysis of Temperature and Precipitation Variability

Taylor (2001) developed a statistical diagram for summarizing several aspects of model performance, including correlation and root mean squared differences. Yue et al. (2002) investigated how autocorrelation influences trend detection in hydrological records. Yue and Wang (2004) examined modifications to the Mann–Kendall test for serially correlated observations. These studies established important statistical principles for evaluating environmental time series. Their contributions demonstrate that temporal dependence can influence conclusions drawn from conventional trend analysis and performance assessment.

Hijmans et al. (2005) developed high-resolution interpolated climate surfaces using environmental observations. Alexander et al. (2006) investigated global changes in daily temperature and precipitation extremes. Zhang et al. (2011) reviewed climatic indices used to monitor changes in extreme temperature and precipitation conditions. Wilks (2011) provided a comprehensive treatment of statistical methods used in atmospheric research. Collectively, these contributions emphasize that

climatic variability requires suitable temporal indicators, carefully selected analytical methods, and consideration of differences among environmental measurement scales.

3.3 Predictive Modeling and Environmental Data Science

Breiman (2001) developed Random Forest as an ensemble learning technique for predictive analysis. Friedman (2001) introduced Gradient Boosting through sequential optimization of predictive functions. These approaches established methods for modeling nonlinear relationships and interactions among explanatory variables. Their methodological relevance extends to environmental analysis, where climatic observations may exhibit complicated statistical dependencies. However, the usefulness of these techniques depends on the composition of the dataset, the selected predictive variables, and the quality of evaluation procedures.

Hastie et al. (2009) discussed statistical learning principles associated with regression, classification, model complexity, and generalization. Mudelsee (2014) examined climate time-series analysis through statistical and bootstrap-based methods. The Intergovernmental Panel on Climate Change (2013) provided a broader scientific assessment of observed climate change and associated uncertainties. These publications establish complementary foundations for climate data science. They indicate that reliable analysis requires integration of computational modeling, statistical evaluation, uncertainty assessment, and scientifically appropriate interpretation of environmental observations.

4. Research Methodology

4.1 Research Design and Environmental Analytical Framework

The present paper adopts a hypothetical comparative quantitative research design intended to examine historical climate variability through statistical and machine learning methods. The proposed framework combines descriptive climatic analysis with one-month-ahead prediction of temperature and precipitation. A seasonal baseline, regularized regression, Random Forest, and Gradient Boosting are considered as alternative approaches. These methods would operate on a common historical dataset containing monthly observations from multiple environmental monitoring stations. No actual computational experiments or observational analyses were conducted.

The proposed analytical procedure consists of environmental data preparation, seasonal reference calculation, anomaly estimation, trend assessment, feature construction, predictive modeling, and independent evaluation. Historical measurements would first be examined for completeness and consistency. Temperature and precipitation anomalies would then be calculated using an established reference period. Predictive features would include seasonal indicators and historical measurements available before the forecast date. Independent chronological testing would provide the intended basis for comparing model performance without allowing future information to influence training decisions.

4.2 Hypothetical Historical Environmental Dataset

For illustrative purposes, the study assumes a fictional environmental dataset containing observations from six monitoring stations across the period 1981–2015. Each station is assumed to contribute monthly mean temperature and monthly precipitation totals for 35 consecutive years. The proposed dataset therefore contains 2,520 station-month observation records. The monitoring stations are hypothetical, and no geographical coordinates or verified environmental measurements are claimed.

These specifications have been constructed exclusively to demonstrate a possible climate data science research design.

The proposed preprocessing procedure would involve identifying missing monthly records, examining measurement consistency, and checking for unexpected changes in the observation series. Temperature anomalies would be calculated relative to station-specific monthly means for 1981–2010. Precipitation anomalies would be expressed relative to the corresponding historical precipitation reference. Feature engineering would include lagged measurements, seasonal indicators, and selected rolling statistics. All predictive features would be constructed using information available before the relevant forecast month to prevent temporal data leakage.

4.3 Statistical Techniques and Predictive Evaluation

The descriptive component would examine annual and multiyear temperature and precipitation anomalies through summary statistics and trend estimation. Linear trend analysis and appropriate nonparametric procedures would be considered, with adjustments for serial dependence where necessary. The predictive component would compare a seasonal baseline, Ridge Regression, Random Forest, and Gradient Boosting. Model configuration would be conducted using historical development observations, while the final period would remain reserved for independent testing. These procedures represent a proposed research protocol.

The intended temporal division would use observations from 1981–2010 for model development and observations from 2011–2015 for independent testing. Earlier development years would support training and later development years would provide chronological validation. Mean absolute error and root mean squared error would serve as principal predictive evaluation measures. Temperature and precipitation would be assessed separately using their respective measurement units. Confidence intervals and repeated temporal evaluation would be necessary before establishing statistically supported differences among models.

5. Analysis

5.1 Illustrative Assessment of Historical Temperature and Precipitation Patterns

Historical climate variability can be examined through deviations from established reference conditions. Temperature anomalies describe differences from a selected climatological mean, while precipitation anomalies indicate departures from reference precipitation totals. Table 1 presents hypothetical temperature and precipitation anomalies across four historical periods. The first three periods represent the 1981–2010 reference interval, and the final period illustrates subsequent conditions. The numerical values were manually constructed and do not describe actual environmental changes within any geographical region.

The hypothetical temperature values increase across successive periods, while precipitation anomalies vary between negative and positive values. This constructed pattern demonstrates how thermal and rainfall conditions may be presented separately within a historical analysis. However, period averages alone cannot establish statistical significance or identify the physical mechanisms responsible for observed differences. A genuine investigation would require complete monthly records, uncertainty estimates, and appropriate statistical testing. The illustrative table therefore provides a reporting example rather than evidence of verified climate trends.

Table 1: Hypothetical Historical Climate Anomalies Across Observation Periods

Historical Period	Duration (Years)	Mean Temperature Anomaly (°C)	Mean Annual Precipitation Anomaly (%)
1981–1990	10	−0.20	−4.0
1991–2000	10	−0.01	+1.0
2001–2010	10	+0.21	+3.0
2011–2015	5	+0.44	−6.0

The illustrative pattern provides a basis for discussing historical variability but does not demonstrate climate change in a real geographical location. Any empirical interpretation would require additional information concerning station coverage, measurement uncertainty, and temporal dependence.

5.2 Illustrative Comparison of Predictive Data Science Models

Predictive climate data analysis requires evaluating how closely estimated environmental values correspond to independent observations. Mean absolute error summarizes the average magnitude of prediction differences, while root mean squared error gives greater weight to larger deviations. Table 2 presents hypothetical prediction errors for four data science methods. The fictional evaluation covers the 2011–2015 test period. The numerical values were manually constructed and do not represent outputs generated through actual predictive modeling or statistical computation.

The illustrative comparison assigns the lowest prediction errors to Gradient Boosting, followed by Random Forest, Ridge Regression, and the seasonal baseline. This pattern demonstrates how increasingly flexible models could be compared under common evaluation conditions. However, the assigned differences do not establish that nonlinear algorithms provide superior environmental forecasts. Reliable conclusions would require implemented models, representative historical observations, independent testing, and uncertainty estimates.

Table 2: Hypothetical Predictive Model Error Comparison

Data Science Model	Temperature MAE (°C)	Temperature RMSE (°C)	Precipitation MAE (mm)	Precipitation RMSE (mm)
Seasonal Baseline	0.74	1.02	29.0	41.0
Ridge Regression	0.63	0.86	26.2	37.4
Random Forest	0.55	0.78	23.5	34.1
Gradient Boosting	0.52	0.74	22.8	33.2

The table illustrates how different statistical and machine learning approaches could be evaluated using common performance measures. It does not establish forecasting superiority or demonstrate an actual improvement in climate prediction accuracy.

6. Discussion

6.1 Temporal Climate Patterns and Environmental Interpretation

Historical environmental records provide an important basis for examining differences in climatic conditions across successive observation periods. Temperature and precipitation anomalies can help researchers organize complex measurements and identify departures from selected reference conditions. However, the interpretation of these anomalies depends on the quality of observations, the selected baseline, and the temporal scale being examined. A single period with unusual conditions does not necessarily establish a persistent trend or indicate a specific environmental cause.

The hypothetical temperature and precipitation patterns illustrate how different climatic variables may exhibit contrasting behavior across time. In the constructed example, temperature anomalies increase while precipitation changes fluctuate between positive and negative values. Such differences would require careful examination in a genuine study because thermal conditions and precipitation processes are influenced by different physical mechanisms. The proposed framework therefore emphasizes variable-specific analysis and cautious interpretation rather than assuming that all environmental indicators respond uniformly to changing climatic conditions.

6.2 Machine Learning and Environmental Prediction Reliability

Machine learning provides opportunities to examine nonlinear relationships among historical environmental observations and temporal predictors. Random Forest and Gradient Boosting can represent complex interactions without requiring every relationship to be defined through a single linear equation. However, environmental prediction remains challenging because climatic systems involve interacting physical processes and variability across multiple time scales. A model may identify useful historical associations while remaining sensitive to changes in observation conditions or the statistical distribution of future environmental measurements.

The hypothetical comparison illustrates how prediction errors could differ among statistical and machine learning models. Although Gradient Boosting receives favorable illustrative values, the example provides no evidence of genuine predictive superiority. An actual investigation would need to establish whether any improvement remains consistent across seasons, stations, and independent evaluation periods. Researchers should also compare computational requirements and interpretability. A simpler statistical model may remain preferable when additional complexity does not produce meaningful or reproducible benefits.

6.3 Statistical Uncertainty and Climate Trend Assessment

Statistical uncertainty represents a central consideration in the analysis of historical climate records. Environmental observations frequently exhibit temporal autocorrelation, seasonal patterns, and differences in measurement quality. These characteristics can influence trend estimates and the interpretation of statistical significance. The proposed framework therefore emphasizes appropriate treatment of serial dependence and transparent reporting of uncertainty. Without such safeguards, short-

term fluctuations or observational inconsistencies may be incorrectly interpreted as meaningful long-term climatic changes.

The methodological literature supports careful statistical assessment of environmental time series. Yue et al. (2002) examined the influence of autocorrelation on trend detection, while Yue and Wang (2004) investigated adjustments to conventional trend-testing procedures. Mudelsee (2014) also discussed statistical analysis under temporal dependence. These contributions indicate that meaningful climatic conclusions require more than descriptive comparisons. Future implementation of the proposed framework should incorporate appropriate confidence intervals, sensitivity analysis, and transparent evaluation of statistical assumptions.

7. Challenges

7.1 Missing Observations and Historical Measurement Inconsistency

Historical environmental datasets may contain missing observations because of equipment failure, interrupted monitoring, incomplete record preservation, or changes in data collection procedures. Such gaps can reduce the reliability of temporal comparisons and complicate the estimation of reference climatic conditions. Measurement inconsistencies may also arise when monitoring instruments or station locations change over time. Consequently, apparent differences among historical periods may partly reflect observational limitations rather than genuine environmental variation.

Preprocessing historical climate records requires careful decisions concerning missing values, measurement corrections, and data homogenization. Inappropriate interpolation may conceal important environmental fluctuations, while untreated inconsistencies can introduce artificial changes into the observation series. The proposed methodology therefore emphasizes documented quality control and sensitivity assessment. A genuine investigation should evaluate whether estimated patterns remain stable under alternative preprocessing choices. These procedures are necessary to preserve the scientific credibility of long-term environmental analysis.

7.2 Seasonal Dependence and Nonlinear Climate Relationships

Climate observations contain seasonal fluctuations that can dominate short-term changes in temperature and precipitation. These recurring patterns complicate trend estimation when data are compared without accounting for calendar-related variation. Temporal autocorrelation introduces additional dependence among consecutive observations. Consequently, statistical methods designed for independent samples may produce misleading uncertainty estimates when applied directly to environmental time series. Appropriate seasonal adjustment and temporal validation are therefore necessary for reliable climatic interpretation.

Nonlinear relationships create further analytical difficulties because environmental variables can interact through complex physical mechanisms. Temperature and precipitation may respond differently to atmospheric circulation, local geography, and changing environmental conditions. Although machine learning methods can represent nonlinear associations, their predictions do not necessarily explain the physical causes of observed patterns. The proposed framework therefore distinguishes computational prediction from scientific attribution. Meaningful climate interpretation requires suitable domain knowledge alongside statistically reliable model development.

7.3 Predictive Generalization and Environmental Data Uncertainty

Predictive generalization presents a challenge because statistical relationships identified within historical observations may not remain stable under different climatic conditions. A model trained during one observation period may perform differently when applied to later years or another geographical region. Changing environmental distributions can affect prediction reliability even when a model performs satisfactorily during development. Consequently, temporal validation and independent testing are essential for assessing the usefulness of predictive approaches beyond their original training conditions.

Environmental uncertainty also arises from incomplete geographical coverage, measurement error, and differences among monitoring stations. These limitations may influence both descriptive climate patterns and machine learning predictions. A model cannot automatically eliminate uncertainty already present within the source observations. Therefore, future investigations should examine data quality, prediction intervals, and sensitivity to alternative analytical configurations. The proposed methodology emphasizes transparent uncertainty reporting as an essential component of responsible environmental data science research.

8. Conclusion

Data science approaches provide a structured methodological foundation for investigating climate variability through historical environmental records. The present paper developed a hypothetical framework incorporating temperature anomaly analysis, precipitation variability assessment, statistical trend examination, and predictive modeling. Its principal focus was the organization of environmental information and the comparative evaluation of computational techniques. The proposed approach demonstrates how historical climatic observations could be examined through complementary descriptive and predictive procedures.

The illustrative tables demonstrate possible methods for reporting historical temperature and precipitation anomalies alongside prediction errors from statistical and machine learning models. However, all numerical values were manually constructed and do not represent actual environmental measurements or completed computational experiments. Establishing genuine climate variability patterns would require representative observations, documented quality control, appropriate statistical testing, and independent evaluation. The hypothetical results therefore cannot support claims about verified climatic changes or model superiority.

Overall, the study identifies historical data quality, temporal dependence, climatic uncertainty, and predictive generalization as important considerations in environmental data analysis. Future empirical investigations should examine these factors using authentic climate records and reproducible computational procedures. Data science may strengthen the analysis of climatic information when its outputs are interpreted alongside established environmental knowledge. Nevertheless, meaningful scientific conclusions must be supported by verified observational evidence and defensible statistical interpretation.

References

1. Mann, M. E., Bradley, R. S., & Hughes, M. K. (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392, 779–787. <https://doi.org/10.1038/33859>
2. Mann, M. E., Park, J., & Bradley, R. S. (1995). Global interdecadal and century-scale climate oscillations during the past five centuries. *Nature*, 378, 266–270. <https://doi.org/10.1038/378266a0>

3. Hasselmann, K. (1998). Conventional and Bayesian approach to climate-change detection and attribution. *Quarterly Journal of the Royal Meteorological Society*, 124(552), 2541–2565. <https://doi.org/10.1002/qj.49712455202>
4. Loehle, C. (2004). Climate change: Detection and attribution of trends from long-term geologic data. *Ecological Modelling*, 171(4), 433–450. <https://doi.org/10.1016/j.ecolmodel.2003.08.013>
5. Meuer, J., Timmreck, C., Fang, S.-W., et al. (2024). Fingerprints of past volcanic eruptions can be detected in historical climate records using machine learning. *Communications Earth & Environment*, 5, Article 455. <https://www.nature.com/articles/s43247-024-01617-y>
6. Plésiat, É., Dunn, R. J. H., Donat, M. G., et al. (2024). Artificial intelligence reveals past climate extremes by reconstructing historical records. *Nature Communications*, 15, Article 9191. <https://doi.org/10.1038/s41467-024-53464-2>
7. Labe, Z., et al. (2024). Changes in United States summer temperatures revealed by explainable neural networks. *Earth's Future*. <https://doi.org/10.1029/2023EF003981>
8. Time series trend analysis and forecasting of climate variability using deep learning in Thailand. (2024). *Results in Engineering*, 24, Article 102997. <https://doi.org/10.1016/j.rineng.2024.102997>
9. Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195–204. <https://doi.org/10.1038/s41586-019-0912-1>
10. Rolnick, D., Donti, P. L., Kaack, L. H., Kochanski, K., Lacoste, A., Sankaran, K., Ross, A. S., Milojevic-Dupont, N., Jaques, N., Waldman-Brown, A., Luccioni, A. S., Maharaj, T., Sherwin, E. D., Mukkavilli, S. K., Kording, K. P., Gomes, C., Ng, A. Y., Hassabis, D., Platt, J. C., ... Bengio, Y. (2019). Tackling climate change with machine learning. *arXiv*. <https://arxiv.org/abs/1906.05433>
11. Ham, Y.-G., Kim, J.-H., & Luo, J.-J. (2019). Deep learning for multi-year ENSO forecasts. *Nature*, 573, 568–572. <https://doi.org/10.1038/s41586-019-1559-7>
12. McGovern, A., Ebert-Uphoff, I., Gagne, D. J., II, & Bostrom, A. (2017). Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, 98(10), 2073–2090. <https://doi.org/10.1175/BAMS-D-16-0123.1>
13. Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J. A., Nicholls, N., Penner, J. E., & Stott, P. A. (2007). Understanding and attributing climate change. In S. Solomon et al. (Eds.), *Climate change 2007: The physical science basis* (pp. 663–745). Cambridge University Press.
14. Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (3rd ed.). Academic Press.
15. von Storch, H., & Zwiers, F. W. (1999). *Statistical analysis in climate research*. Cambridge University Press.
16. Mudelsee, M. (2010). *Climate time series analysis: Classical statistical and bootstrap methods*. Springer. <https://doi.org/10.1007/978-90-481-9482-7>
17. Mann, M. E., & Lees, J. M. (1996). Robust estimation of background noise and signal detection in climatic time series. *Climatic Change*, 33, 409–445. <https://doi.org/10.1007/BF00142586>
18. Mann, M. E., Bradley, R. S., & Hughes, M. K. (2004). Corrigendum: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 430, 105. <https://doi.org/10.1038/nature02478>
19. Folland, C. K., Karl, T. R., Christy, J. R., Clarke, R. A., Gruza, G. V., Jouzel, J., Mann, M. E., Oerlemans, J., Salinger, M. J., & Wang, S.-W. (2001). Observed climate variability and change. In J. T. Houghton et al. (Eds.), *Climate change 2001: The scientific basis* (pp. 99–181). Cambridge University Press.
20. IPCC. (2021). *Climate change 2021: The physical science basis*. Cambridge University Press. <https://doi.org/10.1017/9781009157896>