Climate Risk Analytics

Steve Sain

The Earth's changing climate affects every aspect of life on our planet. Anthropogenic, or human-induced, climate change has already caused and will continue to cause significant and widespread damage, beyond what might be expected from natural climate variability. While many of the impacts of climate change are expected to worsen in the coming decades, this is not some far-off mid- or end-of-century problem.

As the most recent report of Working Group II of the Intergovernmental Panel on Climate Change (IPCC) notes (*https://www.ipcc.ch/report/sixth-assessment-report-workinggroup-ii*), there are already global-scale changes to the structure of different ecosystems, including shifts in species range and changes in phenology. This also includes human systems, including the availability of water resources, food production, and health, and extends to the built environment and economic sectors, often disproportionately affecting the most vulnerable populations.

The World Health Organization (*https://www.who.int/news-room/fact-sheets/detail/climate-change-and-health*) notes increased deaths and poverty, and billions of dollars in direct costs to health, from climate change. The National Oceanic and Atmospheric Administration (*https://www.ncei.noaa.gov/access/billions/*) has been tracking billion-dollar weather and climate disasters in the United States and their impacts since 1980. The decade of the 2010s saw the highest cost and loss of life from these events, with the average cost over the last five years an all-time record, nearly triple the inflation-adjusted average annual cost.

A recent White House Council of Economic Advisors blog post (*https://www.whitehouse.gov/cea/written-materials/2022/09/01/the-rising-costs-of-extreme-weather-events/*) also discusses the rising costs of extreme weather events, focusing on the impact on economic growth. The blog cites a growing number of econometric studies attempting to quantify the economic impacts of climate change. The results of these

Figure 1. Locations associated with supply chain of a clothing retailer. Shading reflects flood heights associated with 200-year return levels in 2050 based on intermediate future greenhouse gas scenario. Gray shading indicates minimal water heights, yellow indicates water heights up to 3 m, red indicates water heights over 3 m.

studies suggest not only that costs are increasing and impacts are widespread, but also that recovering from disasters is far more difficult than previously thought.

Climate Risk Analytics

With the increasing recognition of the impacts and growing costs of climate change, there is a simultaneous growing demand for actionable assessments of climate risk from business, industry, and governments. Risk arises from the interaction between physical hazard, exposure, and vulnerability. Physical hazards can refer to any weather- or climate-related event that causes direct loss, including health-related losses and damage to or direct loss of property. Physical risks can be contrasted with transition risks, which refer to the potential for losses that might arise from a transition to a carbon-free economy.

Climate risk analytics attempt to quantify the risks due to hazards such as flood, heat, wind, fire, and other weather events that are evolving and becoming more frequent or intense with a changing climate. Climate risk analytics may be thought of as an extension of the long history of research on the impacts of climate and climate change (see the research summarized by Working Group II of the IPCC's Fifth and Sixth Assessments. *https://www.ipcc.ch/reports/*).

The assessment of climate risk is at the heart of climate risk analytics, and these assessments can be used in a variety of use cases. One area involves risk assessment and management, and focuses on quantifying the impacts of extreme weather events on people, infrastructure, property, and assets, along with how those impacts are going to change in a future climate. An application centers on the key components of various networks such as those associated with supply chains, power generation and transmission, or even assets that define a financial portfolio.

For example, Figure 1 displays locations across the world that support the supply chain for a major retailer; represents suppliers of raw materials, manufacturing, and distribution; and highlights flood risks in Asia and the Pacific region. (Flood heights are based on Jupiter Intelligence's Climate Score Global product.) With this kind of information in hand, companies can create resiliency initiatives and planning strategies that can ensure redundancy and business continuity, and ultimately mitigate the risk that could endanger their operations and cause material impacts on their financial performance or even the broader economy.

To see more concretely how such analysis might be used, a blog post from the consulting company Guidehouse discusses a framework for assessing climate risk and identifying potential mitigation strategies using a hypothetical utility whose power distribution network is exposed to tropical storms (*https://guidehouse.com/insights/energy/2022/enhanced-climate-resilience-planning-for-the-power-sector*). The post discusses how potential losses to transmission infrastructure, substations, and distribution poles could be a risk from exposure to flood, wind, and heat; how that risk manifests; and how it increases in a changing climate using an intermediate future greenhouse gas scenario.

Based on this assessment, the post includes concrete examples of measures that could be taken for each of these components and would minimize future risks and ultimately reduce the likelihood of outages and costs associated with those outages.

Another example is presented in Section 4.9 of the World Energy Outlook 2022 report of the International Energy Agency (IEA; *https://www.iea.org/topics/world-energy-outlook*). The report discusses how the increasing frequency and intensity of extreme weather events present major risks to energy infrastructure and supply. Highlighting four specific energy facilities in India, Viet Nam, northern Europe, and the southern United States, the report notes how changing flood risks can increase potential losses and disruption of services and highlights the advantages of investing in flood defenses.

The report also notes reduced efficiency of power plants and strain on power systems caused by increasing temperatures and intense cold; impacts of changing wind speeds in regions with wind operations; and impacts of tropical cyclones on power transmission and distribution infrastructure, among others.

Another focus of climate risk analytics is on regulatory response and disclosure. There is a growing demand for companies to identify and disclose to investors and regulators how risks from a changing climate could endanger their operations or cause material impacts to their financial performance and the broader economy.

Many of these disclosures are following the guidance of the Task Force on Climate-related Financial Disclosures (TCFD; *https://www.fsb-tcfd.org/*). For example, AstraZeneca, a global pharmaceutical company, provides information about a number of their facilities around the world for heat, flood, wind, precipitation, and wildfire for a baseline period (1986–2000), 2030, and 2050, based on an intermediate future greenhouse gas scenario, as well as a broader assessment of physical and transition risks (*https://www.astrazeneca.com/Sustainability/resources.html*).

JLL, a global commercial real estate services company, performed a similar analysis and reported the results in their 2020 Global Sustainability Report (*https://www.us.jll.com/en/about-jll/our-sustainability-leadership/archive-global-sustainability-report*).

Downscaling

A traditional approach for quantifying and assessing climate risks features a top-down strategy. A global climate model (GCM) output provides the starting point. While these climate model experiments are critical to the study of the Earth's climate and how it could change with different assumptions about how greenhouse gases will evolve in the coming decades, the model output is typically on spatial scales that make it difficult to use that output directly in impacts studies (see Figure 2).

Climate model output is often downscaled to provide climate data on spatial scales more appropriate for regional or local analyses. These higher-resolution downscaled data products are then incorporated into impacts modeling or other analyses that can be used to identify potentially adverse impacts or geographical areas of concern, as highlighted in the previous section.

A bottom-up approach is more appropriate for a formal risk analysis, and begins by identifying vulnerabilities in a local system based on experience and historical weather conditions. Downscaled data products can then be interrogated to determine the likelihood of climate and weather conditions that would expose those vulnerabilities.

Of course, these are simplistic characterizations of the different paths an analysis may take, and a more comprehensive approach where aspects of both top-down and bottom-up approaches would be preferred. However such assessments are conducted, it is clear that downscaling GCMs is a critical component, and downscaling is more important to bottom-up approaches.

There are two broad classes of downscaling. Dynamic downscaling incorporates regional climate models (RCMs), such as those associated with the Coordinated Regional Climate Downscaling Experiment (CORDEX; *https://cordex.org/*). These RCMs are generally run over limited spatial domains using output from GCMs as boundary conditions. RCMs are valid climate models based on physical principles, but they can be computationally expensive to run and, as with GCMs, have their own biases that often have to be corrected.

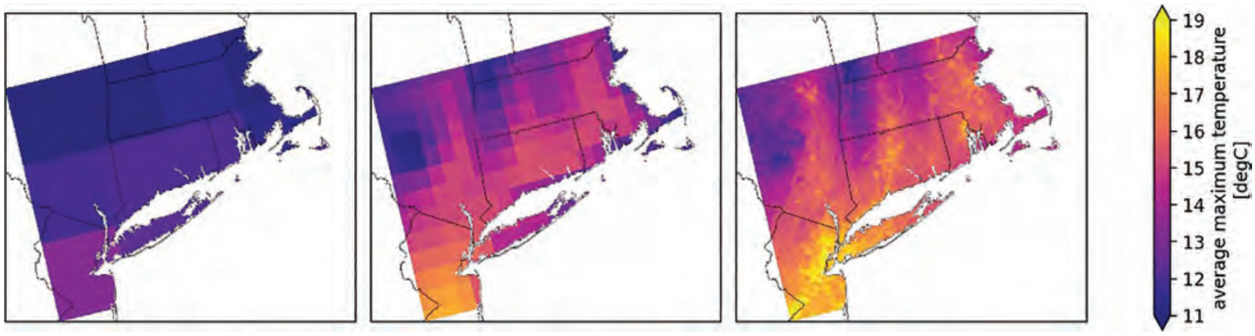

Figure 2. Downscaling example highlighting portions of New Jersey, New York, Connecticut, Rhode Island, and Massachusetts. Left frame displays average maximum daily temperature from a GCM at approximately 1° resolution. Middle frame uses re-analysis data at 30 km resolution; right frame includes downscaled version at 1 km.

A second approach is generally referred to as statistical downscaling, which also focuses on limited spatial domains but exploits empirical relationships to provide higher-resolution climate information. Statistical downscaling can be cheaper to implement, but requires high-quality, high-resolution training data sets and requires assumptions about how the empirical relationships must remain the same in future climates.

Statistical downscaling can be thought of as a supervised learning or regression problem, with the goal of estimating an empirical function $\mathbf{y} \approx g(\mathbf{x})$, where $\mathbf{x}$ represents coarse-scale weather or climate features and $\mathbf{y}$ the finer-scale features. This can be done purely via observational data sets, including reconstructions and reanalysis data sets, or via connecting climate model output with observational data sets.

A broad range of statistical regression methods has been used, including traditional linear regression, generalized linear regression, and generalized additive models, and—more recently— machine learning methods such as random forests and neural networks.

Many downscaling algorithms, particularly when working with climate model output, also incorporate some sort of bias correction of the model output; for example, adjusting time series via a delta method or quantile mapping.

The focus of downscaling is often on creating daily, monthly, or even seasonal time series. It is also possible to downscale specific climate characteristics. For example, the right frame of Figure 2 shows an example resulting from creating a high-resolution (1 km) historical climatology of maximum daily temperature. For comparison, the middle frame shows the same field based on the ERA5 (*https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5*) gridded re-analysis product at 30 km, and the left frame highlights a GCM at 1° resolution.

For this example, the point-based weather station observations represent the local historical weather distributions and climate extremes more accurately. Hence, the image in the right frame of Figure 2 uses a machine learning-based spatial interpolation to connect the parameters of distributions fit to the weather station observations (i.e., the fine scale $\mathbf{y}$) to distribution fits for ERA5 (i.e., the coarse scale $\mathbf{x}$). Other predictors, such as features derived from a high-resolution digital elevation map or from remotely sensed data (land surface temperature, land cover, etc.), were also incorporated.

The left frame shows the output of a GCM over the same area. In this case, the GCM has a cold bias and does not show the temperature response due to topography and urban heat island. However, it should be noted that changes in future temperatures projected by the GCM could be applied in this case using the delta method approach, thus providing a downscaled version of the coarse-resolution GCM projections.

Opportunities for Statistical Science

A wide array of opportunities exists at the intersection of statistical science, data science, and climate science. The study of the Earth's climate, how that climate is changing, and the impacts of those changes are data-driven areas of research with data coming from many diverse sources. Direct observations from weather stations, ocean and river gauges, and radiosondes, as well as remotely sensed observations from satellites, for example, enable specific approaches.

These observations are often restricted in space and time, and there is a demand for data on complete and regular space-time grids, which has led to the creation of climate reconstructions and re-analysis products that use numerical weather models to assimilate observations into a gridded product.

Climate model experiments, such as those included in the most recent CMIP6, are another source of useful data about past, present, and potential future climates.

While statistical science can make contributions in many areas in addition to the downscaling problem discussed previously, two are worth being highlighted.

The first is extremes. Much of climate change research has historically dealt with changes in mean temperature, but there is much interest in the tails of distributions, especially for perils such as precipitation and wind. There is also growing interest in more complex problems, such as compound hazards in which multiple perils, such as precipitation, wind, and storm surge, can simultaneously lead to damage and loss, or the impacts of extreme weather and changes in extreme weather in spatially distributed networks.

The ensembles of climate model output in different model experiments are crucial resources in the assessments and the impacts of climate change, but the construction of these ensembles, particularly multi-model ensembles, does not fit nicely into the statistical concept of a simple random sample. The incorporation of these ensembles into an assessment of climate risk is challenging. Bayesian approaches that incorporate hierarchical models have been used. An alternative framework involves weighting.

These approaches often attempt to balance the ability of the models to reproduce observed climate (higher weights) with dependence or similarity between the models (lower weights). This continues to be an open area of research.

Machine learning is also seeing a great deal of attention well beyond the downscaling problem. There is a long history of the use of machine learning in climate and weather research, and Climate Informatics (*http://www.climateinformatics.org*) has improved the visibility of machine learning and collaboration between machine learning and climate science. Some exciting areas of research involve deep learning, physics-guided machine learning, and interpretable machine learning.

Final Remarks

The statistical sciences have had, and will continue to have, a critical impact on climate analytics as a growing field. Climate change is already affecting broad swaths of society and is projected to continue and, in many cases, increase in the future. Climate risk analytics is emerging as a powerful tool to help companies, governments, and other entities understand their climate risk and build resiliency in response. It involves significant collaboration between data, climate, and other domain scientists, as well as software and data engineers, to develop the computational pipelines needed to build large-scale, global downscaled data sets that are crucial to assessments of climate risk.

This is very much a data-driven effort, and statistical thinking is a crucial component in addressing the uncertainty inherent in the many science- and business-related questions that arise. C

Further Readings

Berliner, M., Hoar, L., Nychka, D., and Hoar, T. 2000. *Studies in the Atmospheric Sciences*. Lecture Notes in Statistics. New York: Springer.

Coles, S. 2001. *An introduction to statistical modeling of extreme values*. Springer Series in Statistics. London: Springer-Verlag.

Cooley, D., Hunter, B. D., and Smith, R. L. 2017. *Univariate and Multivariate Extremes for the Environmental Sciences*. Boca Raton, FL: CRC Press.

Haupt, S. E., Gagne, D. J., Hsieh, W. W., Krasnopolsky, V., McGovern, A., Marzban, C., Moninger, W., Lakshmanan, V., Tissot, P., and Williams, J. K. 2021. The history and practice of AI in the environmental sciences. Available at *doi.org/10.1175/BAMS-D-20-0234.1*.

Huser, R., and Wadsworth, J. L. 2022. Advances in statistical modeling of spatial extremes. *WIREs Computational Statistics* 14 (1), e1537.

Maraun, D., and Widmann, M. 2018. *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge, England: Cambridge University Press.

Monteleoni, C., Schmidt, G., Alexander, F., Niculescu-Mizil, A., Steinhaeuser, K., Tippett, M., Banerjee, A., Blumenthal, M., Ganguly, A. Smerdon, J., et al. 2013. Climate informatics. In T. Yu, N. Chawla, and S. Simoff (Eds.), *Computational Intelligent Data Analysis for Sustainable Development*, Chapter 4, 81–126. Boca Raton, FL: CRC Press.

About the Author

Steve Sain is a senior principal data scientist and senior director, Geospatial and Data Sciences, at Jupiter Intelligence, where he heads the Data Sciences group and serves in Jupiter's science and technology organization. Before joining Jupiter, he held faculty positions in statistics and applied mathematics, and worked in industry as a data scientist where he has led data science research and development programs, as well as data science teams. He has long worked at the intersection of climate research and applied statistics, including a focus on spatial methods for large data sets, extremes, uncertainty quantification, and climate risk analytics. Sain is an affiliate faculty member in the University of Colorado's Department of Applied Mathematics, a fellow of the American Statistical Association, and past recipient of the Distinguished Achievement Award from the American Statistical Association's Section on Statistics and the Environment.

Kotamarthi, R., Hayhoe, K., Mearns, L. O., Wuebbles, D., Jacobs, J., and Jurado J. 2021. *Downscaling Techniques for High-Resolution Climate Projections: From Global Change to Local Impacts*. Cambridge, England: Cambridge University Press.

Nychka, D. 2000. Challenges in understanding the atmosphere. *Journal of the American Statistical Association* 95(451), 972–975.

Reisinger, A., Garschagen, M., Mach, K. J., Pathak, M., Poloczanska, E., van Aalst, M., Ruane, A. C., Mintenbeck, K., Pedace, R., Rojas Corradi, M., et al. 2020. The concept of risk in the IPCC sixth assessment report: A summary of cross-working group discussions: Guidance for IPCC authors.

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., et al., 2022. Tackling climate change with machine learning. *ACM Computing Surveys Review* 55(2), 1–96.

von Storch, H., and Zwiers, F. W. 1999. *Statistical Analysis in Climate Research*. Cambridge, England: Cambridge University Press.