How to Calculate and Communicate Extreme Event Attribution

Sophia Heinrich

February 27, 2026

Table of contents

Assignment Flow

This project is set up to scaffold this user guide and set you up for success.

Start at `topic.qmd` and identify a topic for your user guide.

Once you've proposed a topic and received my approval/feedback, you can convert the contents of `topic.qmd` into an introductory paragraph in your report. Please leave the `topic.qmd` file as an appendix, to document the different stages of this project.

Once we've agreed on a topic for your user guide, you will proceed to `needs.qmd` and `task-analysis.qmd`. Both will be submitted at the same check-in, and both will be included in your final report as appendices. You should be able to re-purpose most of the information in `task-analysis.qmd` to orient the user to different components of the task you've chosen.

Your next step is to actually write the content that satisfies the work you outlined in `task-analysis.qmd`.

Once you're finished with that, ideally, you'll have plenty of time to edit and streamline your report. Feel free to trade reports with a friend and try to complete your friend's task - this will help you both identify areas where the guide isn't as clear. Try to complete the task with a different data set - that often helps find trouble spots.

When you submit your final report, you may remove `index.qmd` from the `_quarto.yml` file, which will remove this chapter (which is quite unnecessary for your user) from the report. You should also take the time to make sure your name is listed as both author and copyright holder in the `_quarto.yml` file. Feel free to make a custom cover for your manual if you would like to do so. You can also tweak the CSS/theme for the book, so long as you're conscious of accessibility concerns.

> Building the report
>
> If you are using RStudio to complete this report, you can hit (Ctrl/Cmd)-Shift-B to build the whole report. You can also type `quarto render .` on the command line in the project folder to accomplish the same thing.
> If you have questions about how to customize or debug your book, it may be helpful to start at the quarto book documentation (Posit PBC 2024).

1 Introduction

This user guide provides step by step instructions on how to calculate and communicate Fraction of Attributable Risk (FAR) and Risk Ratio (RR). These metrics are fundamental parts of Extreme Event Attribution (EEA). EEA aims to evaluate how human-induced changes in Earth's state, like increased atmospheric greenhouse gases, affect the likelihood of extreme climate events. Along with calculations, this guide will show how these metrics can be used to visualize the shift in event probability between the factual and counterfactual world. Additionally, the tutorial will explain the significance of FAR and RR in climate science and how these metrics and visuals may be used to communicate the causality of human behaviors on the climate to decision-making stakeholders.

1.1 Learning Objectives

By the end of this guide, users should be able to

- Calculate Fraction of Attributable Risk (FAR) and Risk Ratio (RR)
- Interpret FAR and RR though causal frameworks
- Visualize and communicate results

1.2 Key Terms

- Extreme Event Attribution (EEA)
- Factual vs Counterfactual Worlds
- Fraction of Attributable Risk (FAR)
- Risk Ratio (RR)
- TODO more

2 Preparation & Prerequisites

2.1 Knowledge Assumptions

List the required knowledge in basic probability ($p = P(X > u)$), distribution functions, and recent climate risk findings.

2.2 Data Requirements

Explain the necessity of having data for the same random variable (X) across two scenarios.

Guidance on Distribution: Mention why Extreme Value Theory (EVT) is preferred for heavy-tailed "rare" events over normal distributions.

3 Defining the Scope

3.1 Threshold Selection (u)

How to set a benchmark for "rare" while balancing confidence intervals.

3.2 Spatiotemporal Scaling

Guidance on defining the event in terms of time and geography (e.g., single day vs. three-month continental scale).

4 Probability Estimation

Instruction on choosing a method for estimating p_1 (factual) and p_0 (counterfactual).

Briefly warn users about sample size requirements to ensure statistical validity.

5 Calculating FAR & RR

Risk Ratio (RR): Formula and step-by-step calculation instructions.

Fraction of Attributable Risk (FAR): Formula and step-by-step calculation instructions.

6 Theoretical Causal Framework

Probability of Necessary Causation (PN): Interpreting if the event would have happened without the causal factor.

Probability of Sufficient Causation (PS): Interpreting if the factor alone was enough to trigger the event.

7 Visualization & Communication

Visualizing Probability Shifts: Tutorial on using univariate density plots to show how a causal factor shifts the distribution.

Stakeholder Communication: Best practices for translating these metrics into real-world decision-making narratives.

8 Troubleshooting & Evaluation

Self-Check List: A set of "Follow-up Questions" for the user (e.g., Is the threshold high enough? Is the distribution model honest?).

Product Evaluation: Criteria to ensure the final conclusion correctly distinguishes between necessary and sufficient causes.

References

Posit PBC. 2024. "Creating a Book. Quarto." July 1, 2024. https://quarto.org/docs/books/.

A Topic

For my user guide, I'm going to provide a tutorial on calculating and communicating Fraction of Attributable Risk (FAR) and Risk Ratio (RR). These metrics are fundamental parts of Extreme Event Attribution (EEA). EEA aims to evalutate how anthropogenic forcings, like increased atmospheric greenhouse gases, affect the likelihood of specific extreme climate events. Along with calculations, the tutorial will include how these metrics can be used to visualize the shift in event probability between the factual and counterfactual world. Additionally, the tutorial will explain the significance of FAR and RR in climate science and how these metrics and visuals may be used to communicate the causality of human behaviors on the climate to decision-making stakeholders.

> **Comments/suggestions**
>
> Ok. Make sure you connect them appropriately to the climate risk issue, and you may want to also look at different ways to visualize climate risk - Lace Padilla has some good stuff on that.

B Needs Assessment

B.0.0.1 Guide Content

This user guide will cover how the metrics of Fraction of Attributable Risk (FAR) and Risk Ratio (RR) can be used as tools for determining the contribution of a causal factor to extreme events. More specifically, the guide will begin by giving background on the factual vs counterfactual setup where it defines two scenarios, one where a specific causal factor is present (p1) and one where it is absent (p0). It will then define the metrics of FAR and RR by giving step by step instructions for calculation RR and FAR. We will then explain how to interpret these metrics along with how to apply them to the theoretical framework of probability of necessary causation and the probability of sufficient causation. In a final application, the guide will show how to visualize the shifts in event probability between the factual and counterfactual world through the use of univariate density plots. These visuals will demonstrate how a causal factor shifts the distribution and changes the likelihood of extreme outcomes. Throughout this content, each topic/step will be explained generally and will also follow an application example for Extreme Event Attribution (as it relates to climate science) to show how these metrics translate to real-world decision-making.

B.0.0.2 User Assumptions

The primary users of this user guide are statisticians, data scientists, and risk analysts across various fields (e.g., climate science, epidemiology, or engineering) who need to quantify the influence of a specific factor on the occurrence of a rare event. The user guide assumes that readers have the following prerequisites:

- Understanding of basic probabilities: $p = P(X > u)$, where X is a random variable and u is a threshold value. Additionally, an understanding of random variables and probability distribution functions is needed.
- An interest in the logic/reasoning of “necessary” vs. “sufficient” causation.
- Basic knowledge of climate risk issues and themes of recent scientific findings in the last 20 years related to human effects on climate change.

B.0.0.3 User Needs

Ultimately, the user wants to provide a numerical answer to the question "How much did factor A increase the risk of event B?" where the numerical answer lies in the calculation and interpretation of FAR and RR. Someone trying to do this will first need help with understanding and following mathematical equations to calculate their metrics. After that, they will need guidance on interpreting FAR and RR through the lens of complex theoretical frameworks and how these interpretations can be communicated to stakeholders. When introducing new complex ideas, it will be important to make conscious design choices that will keep the user engaged and motivated to continue learning. Additionally, it will be important to caution users of requirements or restrictions of the data that these metrics are being applied to (e.g. sample size and distribution).

There are a couple resources that can provide a more thorough background on some of the concepts covered. The first resource is the book *Causality: Models, Reasoning and Inference* by Judea Pearl which teaches the foundations of the underlying causality theory in the user guide. Additionally, the annual review article that inspired this user guide, "Statistical Methods for Extreme Event Attribution in Climate Science" by Naveau et al. also serves as an in-depth reference for applying FAR and RR to complex, dynamic systems.

C Task Analysis

C.0.0.1 Task Prerequisites

Given that this is a user guide for calculating statistical metrics, FAR and RR, and interpreting them, the only prerequisites are knowledge ones as we will not be applying these concepts to any code. The knowledge prerequisites are:

- Understanding of basic probabilities: $p = P(X > u)$, where X is a random variable and u is a threshold value. Additionally, an understanding of probability distribution functions needed.
- An interest in the logic/reasoning of "necessary" vs. "sufficient" causation.
- Basic knowledge of climate risk issues and themes of recent scientific findings in the last 20 years related to human effects on climate change.

C.0.0.2 Data Requirements

- The data must represent the same random variable (X) measured under two different scenarios: the factual and counterfacutal.
- The data is generally assumed to be real numbers, but the range matters significantly based on the event being analyzed. For example, if analyzing an extreme precipitation event, data should always be positive whereas temperature data can be positive or negative.
- Distribution assumptions: the data can be normally distributed but it does not have to be. In fact, when studying the attribution of causal factors for rare events, Extreme Value Theory (EVT) distributions are preferred because they are better for data with heavy tails.

C.0.0.3 Outline of Task Steps

1. Define the Event and Threshold (u)
2. Define the Factual and Counterfactual Scenarios
    1. Choose a method for probability estimation

3. Calculate FAR and RR
4. Apply calculated metrics to Theoretical Causal Framework

    a. Probability of Necessary Causation
    b. Probability of Sufficient Causation

5. Visualize and Communicate Shift in Probability Density Functions

    1. Provide example plots
    2. Explain interpretation and communication for stakeholders

C.0.0.4 On-the-fly Decisions

- Threshold Decision (u): user must decide where to set the threshold for the event of interest. The threshold sets the benchmark/definition for what makes an event “rare”.
- Distribution Model Choice: user must evaluate which type of distribution is best for their data. This will often vary based on the event type (e.g. rainfall)
- Spatiotemporal Scaling: user must decide how to define the event in terms of time and space/geography. For example, for a climate event, should the event be defined by a single day at one station, or a three-month over an entire continent.

C.0.0.5 Task Follow-up Questions

- Is the chosen threshold high anough to represent a truly extreme event without making confidence intervals too wide?
    - Adjustment: if the sample size above u is too small, the user might need to use Extreme Value Theory.
- Is the statistical model appropriate? (e.g. Did I assume a normal distribution for a variable that actually has heavy tails?)
    - Adjustment: choose a statistical model that more honestly represents the data

C.0.0.6 Product Evaluation Criteria

- Can the user complete the calculations?
- Do the calculated FAR and RR align with the visual shifts in the density plots?
- Was an honest distribution model used to represent the data so risk could be accurately estimated?

- Is the chosen threshold (u) high enought to be considered “extreme” while maintaining useful confidence intervals?
- Can the user accurately interpret metrics and communicate visualizations to stakeholders? Can stakeholders understand the communication and make decisions?
- Does the final conclusion correctly distinguish between the cause being necessary (the event wouldn’t have happened without it) versus sufficient (the cause alone triggers the event)?

C.1 Additional Guidance

Your check-in should answer these basic questions (and similar concerns that apply more directly to your topic).

Once you’ve completed the check-in, you can use this section to jump-start an introduction/set-up/getting started section in your user guide. This document should remain as an appendix to your main report.