Pre-submission Report

Harriet Mason

Overview of the thesis

Background and motivation

(This section essentially reiterates the introduction to Chapter Two of the thesis)

Uncertainty visualisation is important because understanding uncertainty allows us to improve decision-making and draw correct conclusions from graphics. The field is still relatively new, but two distinct camps have established themselves in the literature.

One views uncertainty visualisation as the visualisation of an error, probability, or distribution. This camp treats uncertainty as a variable of interest, distinct from its estimate. The second camp sees uncertainty as noise. They believe the role of uncertainty in visualisation is to dampen misleading or false conclusions and amplify correct ones. They treat uncertainty as a necessary nuisance that should be communicated alongside its estimate.

This conflict in the literature creates several recurring problems. Uncertainty visualisations are often bespoke and not properly integrated into the grammar of graphics. This leads to visualisation approaches that inconsistently alternate between treating uncertainty as a variable in its own right, or as a necessary nuisance required to properly visualise an estimate.

The conflicting motivations behind the field also make it difficult to evaluate uncertainty visualisations. In graphics research, user studies allow us to understand which graphics convey information most effectively. However, uncertainty visualisation presents a unique challenge. Most research tries to evaluate users' ability to extract information from a plot in order to establish a hierarchy of visualisation techniques—but this is not the goal of uncertainty visualisation.

Uncertainty visualisation requires signal amplification when uncertainty is low, and obfuscation when it is high. This means that the best technique is not necessarily the one that allows for the most accurate statistical extraction. Some papers have attempted to evaluate uncertainty visualisations, but without a clear definition of uncertainty, this has proven difficult.

Research questions

The primary goal of this research is to clearly define uncertainty visualisation and establish how it differs from other statistical graphics. This work is broken into three specific projects:

1. Survey the uncertainty visualisation literature to establish the motivations behind the field and define what makes a successful uncertainty visualisation.
2. Develop software that makes uncertainty visualisation intuitive and accessible, and properly integrates uncertainty into the grammar of graphics framework.
3. Establish what makes an effective uncertainty visualisation through user studies designed specifically for evaluating uncertainty visualisations.

Thesis Structure

The thesis will be structured as follows:

1. Chapter One provides an overview of uncertainty visualisation, including existing software options. It also outlines the remaining thesis chapters and establishes each chapter's scope.
2. Chapter Two presents a comprehensive review of the uncertainty visualisation literature. It asks broader questions about the purpose of uncertainty visualisation and highlights the wealth of conflicting information in the field. In this chapter, we coin the term signal suppression to describe a visualisation designed to prevent false conclusions—where the signal (i.e. conclusions drawn from estimates) is intentionally suppressed by the noise (i.e. variance on those estimates). We further discuss the challenges in creating and evaluating such plots, which motivates the subsequent chapters. This chapter was submitted to Annual Review of Statistics and Its Application, and is now an invited contribution to Volume 14 of the journal.
3. Chapter Three introduces a new R package, ggdibbler. The conflicting definitions of uncertainty visualisation have made it difficult to incorporate into the grammar of graphics, resulting in ad hoc software that produces only a small number of bespoke plots, disconnected from the broader language of statistical graphics. This chapter addresses that gap by discussing the theoretical implications of including uncertainty in the grammar of graphics. This framework is implemented in ggdibbler, a ggplot2 extension designed for signal suppression that integrates seamlessly into existing visualisation workflows. The package has been submitted to CRAN and will be presented at useR! 2025 in Durham, North Carolina in August 2025.

4. Chapter Four presents a user study, motivated by problems faced by the Australian Energy Market Operator (AEMO), that evaluates the effectiveness of the graphics developed in ggdibbler. This section applies experimental methods introduced in Chapter One to identify which graphical tools are useful in signal suppression.
5. Chapter Five concludes the thesis with a discussion of the contributions and limitations of the work, and provides suggestions for future research.

Timeline

Date	Description
July 2025	Submit `ggdibbler` to CRAN. Pre-submission milestone. Implement VSUP elements in package.
August 2025	Present `ggdibbler` at useR! 2025 and receive feedback. Finish the `ggdibbler` paper and submit to the Journal of Statistical Software. Decide on a case study for the experiment in consultation with an AEMO research representative.
September 2025	Develop Shiny app, conduct pilot testing, revise design, and begin data collection for experiment.
October 2025	Write paper for the user experiment.
November 2025	Finalise and submit user experiment paper. Compile all three papers into the thesis and write the introduction and conclusion.
December 2025	Present final versions of ggdibbler and user experiment at ASC. Submit thesis.