Analyzing Cloud Seeding Using Data Science

Karteek Sankar

Introduction

Cloud seeding is a weather modification method created to increase rain or snowfall by adding particles such as silver iodide to appropriate clouds (Donohue & Lamb, 2025; University of Wyoming, 2020). Although this technique has been utilized across the United States for multiple years, determining the exact amount of additional precipitation caused directly by seeding remains a challenge (Donohue & Lamb, 2025). This project applies data science and AI to help organize historical cloud seeding records and analyze patterns regarding where, when, and why weather modification is used. The hypothesis is that the data driven methods can reveal underlying trends in U.S. cloud-seeding activity based on geographic location, temporal distribution, and the specific seeding agents that are employed.

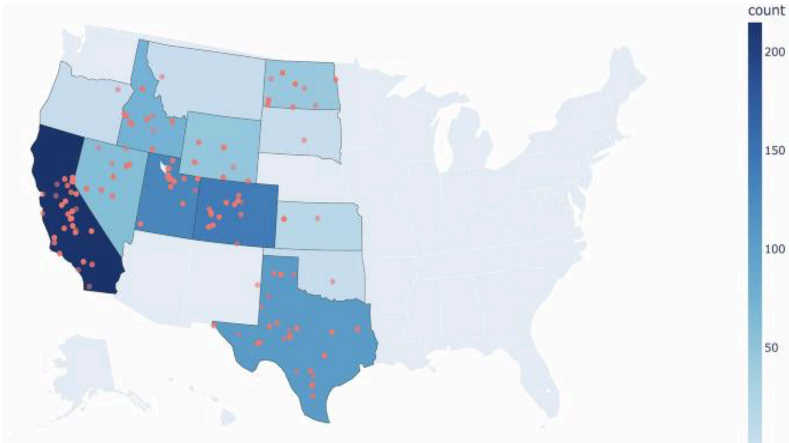
Methodology

The study began by collecting historical National Oceanic and Atmospheric Administration (NOAA) cloud seeding reports spanning the years 2000 to 2025 (Donohue & Lamb, 2025). A custom Python data pipeline was then developed to convert unstructured information from PDF reports into structured data. To handle scanned and handwritten records optical character recognition (OCR) and a large language model (LLM) were utilized for automated extraction. This process generated a dataset that contains important operational variables including the year, season, seeding agent, and stated purpose. Finally, data analysis and visual tools were applied to identify geographic, seasonal, and operational patterns across the records. In total 832 unique cloud-seeding projects were successfully extracted from the NOAA reports, with LLM-based extraction achieving an accuracy of 98.38% based on a manual review of 200 randomly sampled records (Donohue & Lamb, 2025).

Results

Cloud-seeding activity was concentrated heavily in western U.S. states like California, Colorado, and Utah which reflects regional dependence on winter snowpack for water resources (Donohue & Lamb, 2025). Moreover, Texas had substantial activity related to rain enhancement in the summertime (Donohue & Lamb, 2025). Overall, these geographic patterns demonstrate how operational programs target specific climatic zones with acute water management needs.

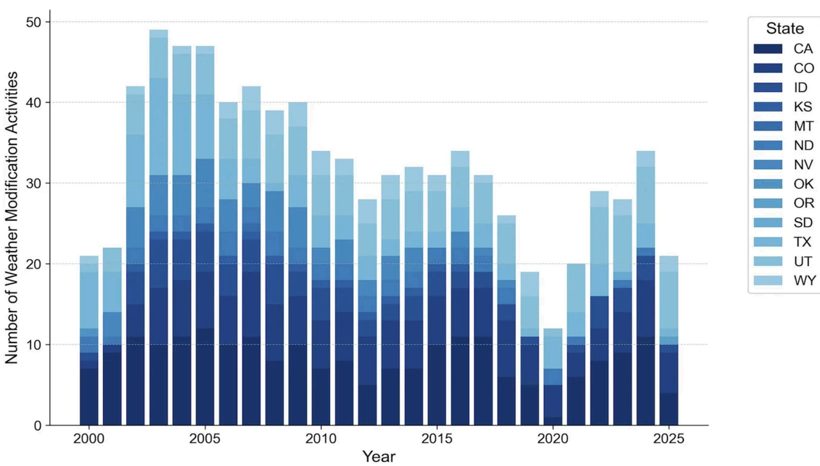
Figure 1



Cloud-seeding activity concentration by each U.S. state (2000–2025)

Reported cloud-seeding activity increased around 2003–2005 which reflects a period of expanded operational investments and reporting (Donohue & Lamb, 2025). Activity decreased through the 2010s before increasing again after 2021 (Donohue & Lamb, 2025). This fluctuation demonstrates how long-term weather modification efforts respond to shifting state funding, drought conditions, and the administrative priorities.

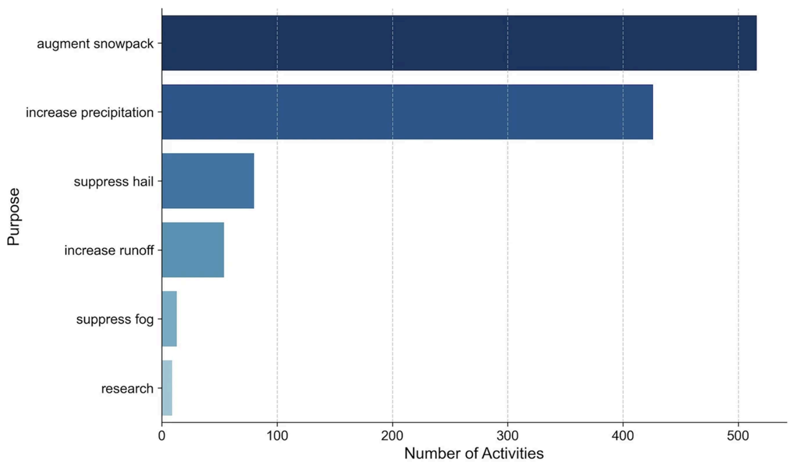
Figure 2



Annual distribution of reported weather modification activities across the U.S. states

Increasing snowpack was the most common stated purpose of cloud seeding across the extracted records which is followed by increasing precipitation and general rain enhancement (Donohue & Lamb, 2025). Secondary objectives like suppressing hail, increasing runoff, and supporting research accounted for smaller fractions of total reported activities (Donohue & Lamb, 2025). These varying objectives help highlight the primary role of winter orographic seeding in sustaining regional watershed reserves.

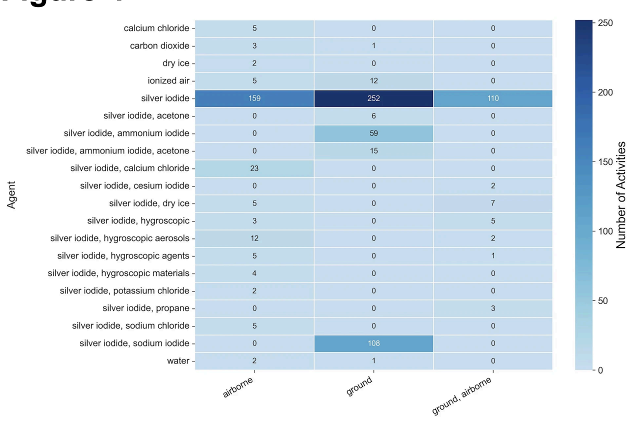
Figure 3



Stated purposes of cloud-seeding projects based on total number of activities

Silver iodide was overwhelmingly the most reported seeding agent used in operations due to its effectiveness as an ice nucleating structure (Donohue & Lamb, 2025). Furthermore, ground-based deployment methods were significantly more common than airborne techniques across the dataset (Donohue & Lamb, 2025). These operational preferences reflect logistical and cost considerations inherent in long-term atmospheric modification programs.

Figure 4



Cross-tabulation of seeding agents

The cross-tabulation of operational records reveals patterns regarding the combination of materials and delivery mechanisms used in field operations (Donohue & Lamb, 2025). Specifically silver iodide utilized through ground-based apparatuses forms the highest concentration of reported project methods by a huge margin (Donohue & Lamb, 2025). Other combinations like airborne seeding agents appear far less frequently across the dataset (Donohue & Lamb, 2025).

Discussion

The analyzed data indicates that cloud-seeding operations are mainly concentrated in western states where winter snowpack and regional water supplies are critical (Donohue & Lamb, 2025). However, this dataset alone doesn't establish a direct relationship proving that cloud seeding increases precipitation. Since cloud enhancement relies heavily on the presence of suitable atmospheric conditions, separating the effects of seeding interventions from natural weather variability remains difficult (University of Wyoming, 2020; Statistical evaluation of the results of cloud seeding, 1980). While the application of AI and data science provides an efficient workflow for organizing historical records the accuracy of extracted information must be verified against the original source reports (Donohue & Lamb, 2025).

Conclusion

Data science can turn difficult-to-use cloud seeding reports into structured datasets that can be analyzed for the patterns (Donohue & Lamb, 2025). The results show how cloud seeding is mainly used in the western U.S. and how silver iodide is especially common (Donohue & Lamb, 2025). Overall, cloud seeding could be useful under the right conditions, but controlled research and improved data are necessary to determine how effective it truly is (University of Wyoming, 2020).

Acknowledgments

This research would not have been possible without the contribution of the resources/mentors from the Science Mentorship Institute Program.

References

Donohue, J. J., & Lamb, K. D. (2025). Structured dataset of reported cloud seeding activities in the United States (2000–2025) using an LLM. *Scientific Data*, *12*, 1996. https://doi.org/10.1038/s41597-025-06273-1

Donohue, J. J., & Lamb, K. D. (2025). Structured dataset of reported cloud seeding activities in the United States (2000–2025) using a large language model. *arXiv*. https://doi.org/10.48550/arXiv.2505.01555

Statistical evaluation of the results of cloud seeding. (1980). *International Geophysics*, *24*, 135–161. https://doi.org/10.1016/S0074-6142(08)60125-4

U.S. Bureau of Reclamation. (2000). *The feasibility of operational cloud seeding programs*. U.S. Department of the Interior. https://platteriverprogram.org/sites/default/files/2025-10/usbr-2000feasibility-cloud-seeding.pdf

University of Wyoming. (2020, February 24). *UW researchers contribute to follow-up study of first quantifiable observation of cloud seeding*. https://www.uwyo.edu/news/2020/02/uw-researchers-contribute-to-follow-up-study-of-first-quantifiable-observation-of-cloud-seeding.html