

1 DRAFT: Precipitable-Water Model Analysis Tool: An  
2 open-source suite for estimating precipitable water with  
3 low-cost instrumentation.

4 **Spencer Riley**<sup>1¶</sup>, **Vicki Kelsey**<sup>\*2</sup>, and **Kenneth Minschwaner**<sup>1</sup>

5 <sup>1</sup> Physics Department, New Mexico Institute of Mining and Technology <sup>2</sup> Atmospheric and  
6 Environmental Sciences Program, South Dakota School of Mines and Technology ¶ Corresponding  
7 author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#) ↗
- [Repository](#) ↗
- [Archive](#) ↗

Editor: [Open Journals](#) ↗

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain  
copyright and release the work  
under a Creative Commons  
Attribution 4.0 International  
License ([CC BY 4.0](#)).

8 **Summary**

9 Precipitable water is the vertically integrated amount of water vapor in a column of air from the  
10 surface to the top of the atmosphere ([Salby, 1996](#)). The Precipitable-water Model Analysis Tool  
11 (PMAT) is an open-source software suite designed to study the correlation between atmospheric  
12 brightness temperature and precipitable water ([Riley & Kelsey, 2021](#)).

13 **Statement of Need**

14 Precipitable water data can be used in weather forecasting to determine the amount of  
15 potential rainfall available and to study the dynamical evolution of convective storms. The  
16 lack of reliable and sufficient precipitable water data can have a significant impact on the  
17 quality of weather forecasts Smith & Toumi ([2008](#)). PMAT addresses the need for an easily  
18 integrated analysis workflow to quantitatively characterize the relationship between regional  
19 PWAT and localized atmospheric brightness temperature observations for areas that lack high  
20 to moderate-resolution data collection infrastructure. The software suite has already been  
21 configured and deployed for use in the South-Central New Mexico climate zone using a series of  
22 handheld infrared thermometers as the source of atmospheric brightness temperature data. We  
23 plan to expand the scope of the research with community science endeavors in South Dakota;  
24 while also expanding the suite's analysis capabilities with machine learning and additional  
25 statistical products for predictive modeling.

26 **Software Architecture and Design**

27 PMAT contains three primary modules: deployment, pre-processing, and analysis. The three  
28 modules can collect and organize data, conduct primary and secondary analysis functions, then  
29 visualize the results.

30 There are two primary prerequisites for the full deployment of PMAT:

- 31 (1) A data file consisting of measurements that include ground and sky temperatures, time  
32 and date stamps, and visual condition of the sky (either clear-sky or overcast)
- 33 (2) A configuration file that details sensor information, analysis parameters, and information  
34 required to retrieve data

\*co-first author

The deployment mechanism allows end-users to implement the software suite on local and cloud-based systems, which facilitates access to the generated data products and visualizations. We have configured a template repository that illustrates the required directory structure.

The pre-processing module involves aggregating atmospheric brightness temperature data collected in the field with data from radiosondes and local ground stations. The core of this subprocess is a Python script that will read in the raw data and collect atmospheric data from the University of Wyoming Upper-Air database via siphon and the University of Utah MesoWest database. PMAT supports the usage of external data sources, such that local data files that contain precipitable water and relative humidity data be processed. This process also includes a series of three primary filters to organize and refine the data. The first filter separates the labeled overcast and clear-sky data. Next, the data which contains the label Do Not Analyze (DNA) is filtered out. This DNA label is added manually to the raw data file and can be used to remove outliers. The last filter compared the standard deviation of the individual days with the average standard deviation of the dataset. The threshold of this filter originates from the end-user's configuration file.

The primary role of the Analysis module is to present the correlation between PWAT and zenith sky temperature in terms of regression statistics. In the PMAT analysis methodology, we have implemented an iterative process for computing the exponential regression statistics and coefficients. The PWAT and temperature data undergo a linearization in the form of the exponential equation:

$$\text{PWAT} = Ae^{BT_b}. \quad (1)$$

After the iteration process is complete, the regression statistics are averaged together to further minimize random bias. The results of the individual step runs and the final averages are presented as a part of the data products that are generated. In addition, we have secondary analysis functions which include instrument comparisons and sinusoidal regression for time-series data.

## Future Developments

As we continue to develop and maintain the PMAT software, we intend to add additional modules and optional analytical functions. One of the major developments is a machine learning module that utilizes TensorFlow's support vector machine framework to classify the dataset into a binary weather condition scheme.

## Acknowledgements

## References

- Riley, S., & Kelsey, V. (2021). Precipitable-water model analysis tool. In *GitHub repository*. GitHub. <https://github.com/physicsgoddess1972/Precipitable-Water-Model>
- Salby, M. L. (1996). *Fundamentals of atmospheric physics* (1st ed.). Elsevier. ISBN: 0126151601
- Smith, S., & Toumi, R. (2008). Measuring Cloud Cover and Brightness Temperature with a Ground-Based Thermal Infrared Camera. *Jamc*, 47(2), 683–693. <https://doi.org/10.1175/2007JAMC1615.1>
- Yang, L., & Smith, J. (2018). Sensitivity of extreme rainfall to atmospheric moisture content in the arid/semiarid southwestern united states: Implications for probable maximum precipitation estimates. *Jgr: Atmospheres*, 123(3), 1638–1656. <https://doi.org/10.1002/2017JD027850>