HardiPy: An Open-Source Program for the Development of an Updated Plant Hardiness Zone Map

Zach Pilgrim, 2020

GISC 3200K - Programming For Geospatial Science & Technology

Dr. Huidae Cho

1. Abstract

An open-source program that is able to process the PRISM weather dataset to recreate the current USDA Plant Hardiness Zone Map and create an updated version using the same parameters but with updated data. This is a program that will continue to be developed, to add in support for new parameters to develop a more accurate plant hardiness zone map that is more representative of actual zones throughout the United States. The weather data is binned into classes using Numpy Arrays and calculating the mean minimum temperatures of the input datasets. Currently the only factor considered for hardiness zones is minimum annual temperatures, the eventual support for other factors, such as chilling hours and the Persephone period are planned.

2. Overview

2.1 Introduction:

In an era that has experienced rapid climatic shifts in a relatively short period of time, the importance of analyzing this change could not be greater. This is not a change that is simply imminent; the impacts of climate change have already been felt. (Matthews, Stephen N., et al. 2018; Mckenney, Daniel W., et al.2014; Teqja, Zydi, et al. 2017; Daly, Christopher, et al. 2012) The current Plant Hardiness Zone Map (PHZM) that is available through the USDA is based on mean annual extreme minimum temperature data from 1976-2005. The need for an accurate and updated PHZM is integral to efficient agriculture (Matthews, Stephen N., et al. 2018; Krakauer, Nir Y 2012; Coder, Kim D 2012). The zones that were defined in the 2012 update are no longer an accurate representation of plant hardiness (Matthews, Stephen N., et al. 2018; Parker, Lauren E., and John T. Abatzoglou 2016; Krakauer, Nir Y 2012). This presents a problem for both large-scale and small-scale agriculture. As hardiness zones are shifting, crops that have traditionally been grown in certain regions will no longer be viable there (Gloning, Philipp, Nicole Estrella, and Annette Menzel 2013; McKenney, Daniel W., et al. 2007). This project aims to address the need for an updated PHZM that can be easily updated with newer data as the climate continues to change. The addition of other factors into the delineation of zones would be beneficial to create more accurate plant hardiness zones. Currently the only factor, that being minimum annual temperature, used to calculate plant hardiness does not account for seasonal plants. It only accounts for plants that are non-seasonal by only taking this factor into account. The main reason for this is because this hardiness zone map was originally intended for trees instead of agriculture. As most agriculture is seasonal and is not grown during the coldest time of the year, this is the dormant time so it does not truly impact most agriculture, but these hardiness zones are used as the standard. If this process was to be recalculated taking into account seasonal variation and calculating minimum temperatures during the different seasons rather than annual variation. As well as other important factors such as chilling hours, daylight hours and average precipitation. These would all assist in providing a more accurate representation of hardiness zones.

This project has been developed using open-source data and material, using NumPy, MatPlotLib, among others, there are additional plans for the additional need for the Geospatial Data Abstraction Library (GDAL) (Van der Walt, S., Colbert, S. C., and Varoquaux, G. 2011; GDAL/OGR Contributors 2020). The libraries that are used allow for reading and manipulating the data while maintaining geospatial fidelity, as well as the required computational ability for analysis. The choice to develop this project in an open-source format is to provide an easily accessible process that is capable of providing an up-to-date map of plant hardiness zones. Which could be used to benefit agricultural decision making. This version of HardiPy is only to recreate the current version of the USDA PHZM, this is not the final intended version of HardiPy, but only the current version.

2.2 Project Statement:

The purpose of this project is to develop an open-source program that is capable of processing the PRISM weather dataset to provide an updated and current PHZM. This program will be built with as few dependencies as possible, this will provide the simplest method for creating this updated map. The ability to quickly and effectively update plant hardiness zones could possibly help to serve in agricultural decision making. The ability to effectively predict hardiness zones is an asset and a necessity in a changing environment. The eventual support for creating a newer more accurate representation of plant hardiness zones will be added as this program is continually developed.

2.3 Completed Objectives:

1) Download complete PRISM weather dataset in ASCII format

2) HardiPy development to convert the ASCII files into arrays and sort into bins

3) Use Numpy Digitize to return the values of indexed bins

4) Calculate mean minimum temperature for each input file

5) Mask No Data Values

6) Display completed results

2.4 Planned/Future Objectives:

1) Complete change to function requiring only input and output directory

2) Output as .asc

3) Develop solution to allow for array to use float as step value

4) Include support for other factors for band delineation that better represent hardiness zones

5) Web scraping support to allow for the entire process to be automated

6) Support for statistical comparison of datasets.

3. Materials & Methods:

3.1 Materials:

HardiPy requires the installation of Numpy, and Matplotlib, and makes use of the OS and glob modules as well. The data that is processed by HardiPy is the publicly available dataset process by Oregon State University. This dataset is available as both monthly and annual datasets with a spatial resolution of 4km. The PRISM dataset also contains data at higher temporal resolution as well as the availability of the data from the original point data. This dataset is what was used in the original USDA PHZM dataset, this is why it was chosen for this project.

3.2 Methods:

HardiPy only requires the desired years of weather data in .asc from the PRISM dataset. This contains weather data for the entire conterminous United States for a single year. There is no limitation to the number of input files that HardiPy can analyze as there are no hardcoded variables within the program. Currently, the program has been tested with up to 38 files, weather data from 1981 to 2019, and processed and output the result in ~35 seconds. This update to HardiPy to process a range of years has enabled for the development of the same result of the original USDA PHZM. Currently, there is not support for the statistical comparison of the newer dataset and the original year range dataset, but both datasets have been analyzed.

Once the desired .asc files have been chosen and downloaded all that is required is to load this file into the same directory as HardiPy. Once the program is initiated, it loads each individual file ending in .asc using the numpy.loadtxt function. Where it skips the header of each file and accounts for the no data values of the dataset. Each file is then binned into an array using the numpy.array function, with a range from -90°C to 50°C with a step of 3. These values were chosen because they are greater than the range of any of the possible datasets in the PRISM data. A step of 3 degrees is used because in the original USDA data a step of 5°F was used. The equivalent of this is 2.7°C, however support for float values is not yet completed for this project. This binning process will also be updated to read the minimum and maximum values of the dataset and automatically create bins that will contain the data with a step of 2.7°C to create the most accurate result possible.

After the creation of the bins, the data is then separated and categorized within them using Numpy Digitize (Numpy; Wasser, L., Holdgraf, C., & Morrissey, M. 2018; Hass, Bridget 2018). This process separates the values of the array into the according bins. For example, a zone with an average minimum temperature of 24°C would fall in the bin containing values 22-25. This method effectively reclassifies the data but maintaining the possibility of ranges shifting through the years for comparison of the greater dataset. Each of the files that are input into the program after being processed are added to a list. This list is then used to calculate the mean minimum temperatures for the entire dataset.

Currently the last major step in the analysis portion that HardiPy performs is masking the No Data values to ensure data accuracy and prevent inappropriate bin creation. This is accomplished using numpy where class '0', the class containing NoData values is removed from the result (Wasser, L., Holdgraf, C., & Morrissey, M. 2018; Lawhead, J. 2013). Lastly the data is displayed on screen and saved as a .png in the containing folder using the Matplotlib plotting function. The eventual support for saving the final mean data as a new .asc file is planned for development for HardiPy.

4. Results & Discussion:

4.1 Results:

The results from HardiPy achieve the intended results that this project set out to accomplish. It is able to process the PRISM weather dataset and create an output of plant hardiness zone maps using updated data. There are still steps that need to be taken to get this program to its initially intended capability. The ability for statistical comparison of the datasets as well as output of a mean average .asc file are yet to be completed. However, the majority of this project has been completed and has accomplished what was intended. The results of the same dataset used in the USDA PHZM (figure 1.) has been processed using HardiPy. This result when compared to the updated dataset (figure 2.) do not seem to be as notably different as initially expected. However, there are still differences between the two results, most notably in the southern regions and through portions of the middle latitudes. In the southern regions' bands are pushing northward as temperatures rise, and in the middle latitudes bands appear to widen. Results from smaller datasets are more clear in the change that has happened in a short period of time, but this can also be a misrepresentative or misleading result as weather patterns can wax and wane in short periods of time. Climate patterns are more consistent in their change, and a better representation of total changes over time. Comparisons of 10 year time periods from 1981 – 1991 and 2001 – 2010 (figures 3 & 4) show short term change.

4.2 Discussion:

The ultimate goal of this project was to develop a completely open-source program from scratch and learn in the process how to accomplish this. The focus was less on the exact analysis, but more upon the development of the program that is capable of accomplishing this. Prior to this project this is not a process that I was familiar with and this was a much more difficult undertaking than initially anticipated. This has been both a rewarding and difficult project that enabled me to learn far more about the process of developing a program that is capable of accomplishing this type of analysis. To develop this program required a greater understanding of how to apply numpy and how arrays actually worked so that I would be able to convert the data into the appropriate format and analyze and process the data as I required it. This process took a lot of reading and learning about all of the steps needed to accomplish this. The results from this program are not as precise as would have been preferred, they accomplish the goal but are not as good as they could be. With more time put into the development of this program I hope to improve the quality of the results that it outputs. However, the learning process in developing this program has been invaluable. The amount of time that was required to learn how to accomplish a basic program of this type took quite a long time to understand. As more time has been put into developing this program and learning to implement the fundamentals it has been very helpful to understand how to better form the approach for my next foray into python programming.

5. Conclusion:

5.1 Conclusion:

This project was able to accomplish the some of the goals that it set out to do, it is able to process data using an open-source program and output the results. This was a large undertaking, and while the final product is not perfect, or even entirely complete, I am proud of the results that I was able to create. This was a much larger process than I anticipated, but the results have been rewarding. Everything that I learned through this process has been very helpful in my fundamental knowledge of python. This is a project that I plan to continue working on, not just to develop a more effective method for developing plant hardiness zones, but also to further my own understanding of how the best process for developing an open-source program is done.

6. References:

• Hass, Bridget, Classify a Raster Using Threshold Values in Python - 2018. (n.d.). Retrieved September 16, 2020, from https://www.neonscience.org/classify-raster-thresholds-2018-py

• Coder, Kim D. "Hardiness zone changes: Across the United States and Georgia." (2012).

• Daly, Christopher, et al. "Development of a new USDA plant hardiness zone map for the United States." *Journal of Applied Meteorology and Climatology* 51.2 (2012): 242-264

• GDAL/OGR Contributors (2020). GDAL/OGR Geospatial Data Abstraction software Library. Open Source Geospatial Foundation.

• Gloning, Philipp, Nicole Estrella, and Annette Menzel. "The impacts of climate change on the winter hardiness zones of woody plants in Europe." *Theoretical and applied climatology* 113.3-4 (2013): 683-695.

• Krakauer, Nir Y. "Estimating climate trends: application to United States plant hardiness zones." *Advances in Meteorology* 2012 (2012).

• Matthews, Stephen N., et al. "Assessing potential climate change pressures across the conterminous United States: mapping plant hardiness zones, heat zones, growing degree days, and cumulative drought severity throughout this century." *RMAP-NRS-9. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station. 31 p.* 9 (2018): 1-31.

• McKenney, Daniel W., et al. "Beyond traditional hardiness zones: using climate envelopes to map plant range limits." *BioScience* 57.11 (2007): 929-937.

• Mckenney, Daniel W., et al. "Change and evolution in the plant hardiness zones of Canada." *BioScience* 64.4 (2014): 341-350.

• Parker, Lauren E., and John T. Abatzoglou. "Projected changes in cold hardiness zones and suitable overwinter ranges of perennial crops over the United States." *Environmental Research Letters* 11.3 (2016): 034001.

• Taylor, S. (2020, September 9). Sdtaylor/pyPRISMClimate. Retrieved September 16, 2020, from https://github.com/sdtaylor/pyPRISMClimate

• Teqja, Zydi, et al. "A study of the impacts of climate change scenarios on the plant hardiness zones of Albania." *Journal of Applied Meteorology and Climatology* 56.3 (2017): 615-631.

• United States Department of Agriculture. (n.d.). Retrieved September 16, 2020, from https://planthardiness.ars.usda.gov/PHZMWeb/

• Van der Walt, S., Colbert, S. C., and Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. Computing in Science & Engineering, 13(2):22–30.4

• Lawhead, J. (2013, December 4). Python and Elevation Data: ASCII Grid Files. Retrieved October 28, 2020, from http://geospatialpython.com/2013/12/python-and-elevation-data-ascii-grid.html

• Numpy.digitize¶. (n.d.). Retrieved October 28, 2020, from https://het.as.utexas.edu/HET/Software/Numpy/reference/generated/numpy.digitize.html

• Wasser, L., Holdgraf, C., & Morrissey, M. (2018, February 05). Classify and Plot Raster Data in Python. Retrieved October 28, 2020, from https://www.earthdatascience.org/courses/use-data-open-source-python/intro-raster-data-python/raster-data-processing/classify-plot-raster-data-in-python/

5.2 Figures:

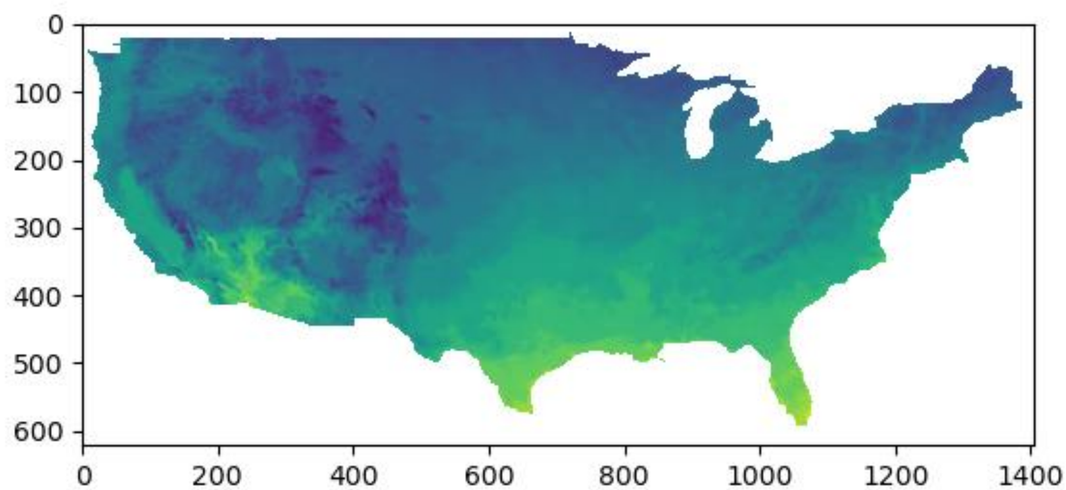

Figure 1. HardiPy Plant Hardiness Zone Map 1981 – 2010

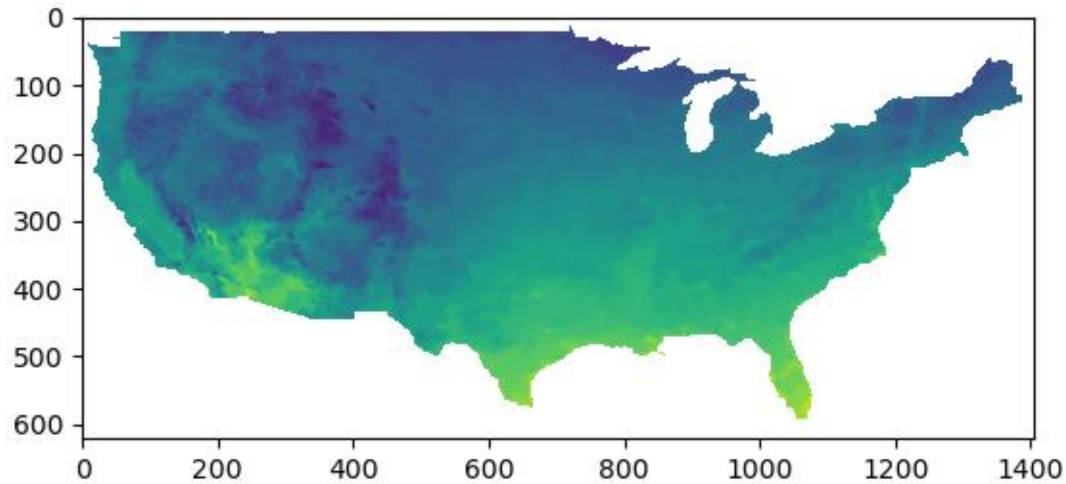

Figure 2. HardiPy Plant Hardiness Zone Map 2009 - 2019

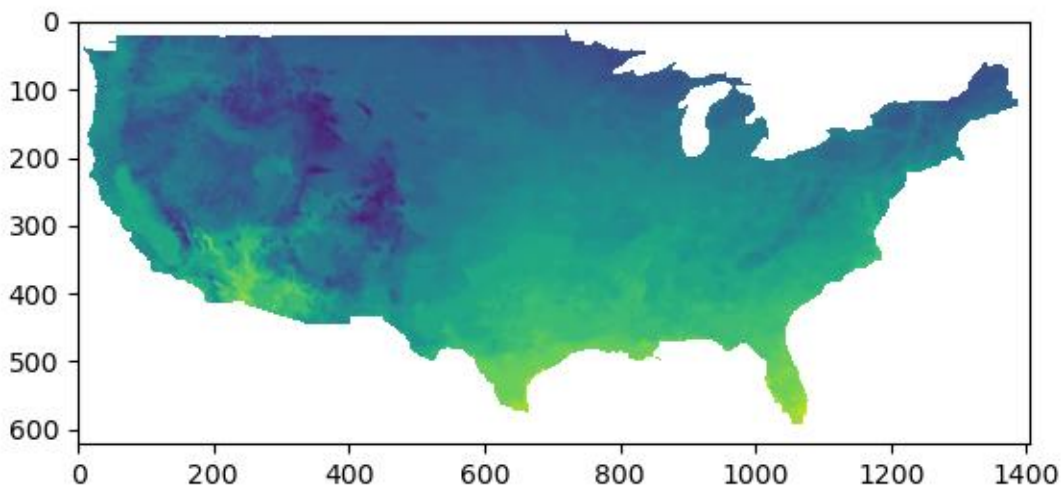

Figure 3. HardiPy Plant Hardiness Zone Map 1981 – 1991

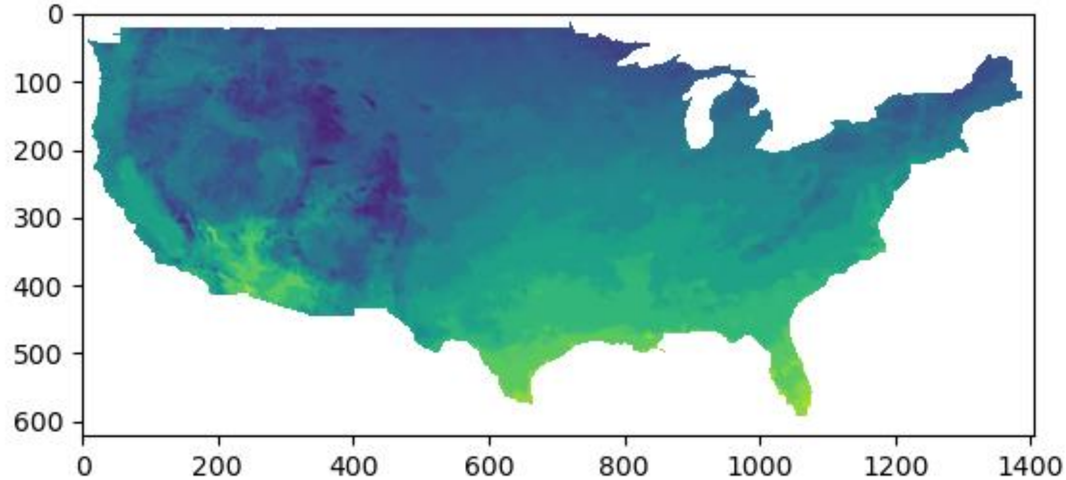

Figure 4. HardiPy Plant Hardiness Zone Map 2001 – 2010