Creating a Climatological Snowfall Map for the National Weather Service Buffalo County Warning Area Using an Ordinary Least-Squares Regression of PRISM Data with Residual Correction Scheme

Jefferson Wood*
NOAA/National Weather Service
Buffalo, New York

Abstract

An updated 30-year snowfall climatology map of the National Weather Service Weather Forecast Office Buffalo, NY county warning area is constructed from National Centers for Environmental Information 1981-2010 Normals point data using a Geographic Information System-driven objective analysis technique, replacing the previous map that was constructed using human-expert techniques. The new technique, henceforth referred to as Ordinary Least-Square Regression with Residual Correction (OLRwRC), involves a regression analysis of 30-year normal snowfall point data across the CWA against values derived from gridded cold-season precipitation data obtained from the PRISM Climate Group. The resulting regression equation is applied to the gridded Parameter-elevation Regression on Independent Slopes (PRISM) precipitation data to create an initial snowfall map, which is then corrected using an interpolated grid of the residuals from the regression. The resulting map, while not without flaws, is a significant step forward in the representation of snowfall climatology in western and north-central New York, revealing previously unresolved details and highlighting the significant variability in snowfall that is caused by the area’s unique geography.
1. Introduction

The National Weather Service Weather Forecast Office (WFO) in Buffalo, New York is located on the eastern flank of the Laurentian Great Lakes, a region renowned for its lake-effect snow. The climatological distribution of snowfall across the WFO County Warning Area (CWA) is a source of great interest to the public, the media, and other stakeholders across the region. An up-to-date highly detailed snowfall climatology of the area, deliverable in a digital format, is highly desired; however, to this point, no such map exists. This is likely a result of the inherent complexities in the mapping of snowfall brought about by the complex terrain of western and north-central New York in addition to the relative lack of snowfall observing sites when compared to standard temperature and precipitation observing networks, an issue addressed in Durre et al. (2013), as well as Romanov et al. (2000).

Historically, climate maps tend to be constructed using either the human-expert or statistical methods (Daly et al. 2002). In fact, a snowfall climatology map for the WFO Buffalo CWA, created approximately 20 years ago using the former methodology, still exists (Fig. 1). This map captures the basic trends in snowfall across the CWA quite well, and has held up well over the years. However, when 1981-2010 snowfall normals from the National Centers for Environmental Information (NCEI) are plotted on top of this map, as seen in Fig. 1, both the strengths and weaknesses of the map become more apparent. Clearly, there are details in snowfall patterns across the CWA that this map does not fully capture, and assumptions were made that are not supported by the latest climatological data. Therefore, the goal of this project was to produce an updated snowfall climatology.

**Figure 1.** The previous Buffalo CWA mean annual snowfall map, as originally created by Steve McLaughlin, WFO Buffalo Senior Forecaster (retired). Color-fill and contours depict snowfall in inches. Points depict snowfall in inches using 1981-2010 NCEI/Environment Canada (EC) climate normals.
map of the CWA that utilized the latest snowfall climate data from NCEI, the 1981-2010 U.S. Climate Normals, and leveraged the power of a Geographic Information System (GIS) to provide a more detailed insight into the patterns of snowfall across the area. An additional advantage of using GIS is the ability to easily output the data into a digital format that can be flexibly utilized by the NWS and its partners, as well as the general public.

2. Study Area

*Daly (2006)* states that “spatial climate patterns are most affected by terrain and water bodies, primarily through the direct effects of elevation, terrain-induced climate transitions, cold air drainage and inversions, and coastal effects.” These factors are particularly amplified on scales of less than 10km (*Daly 2006*). The geography of western and north-central New York provides an ideal conflation of the factors described in *Daly (2006)* to produce significant variability in snowfall across the WFO Buffalo County Warning Area (CWA). The CWA ranges in elevation from a minimum elevation of 72 m along the St. Lawrence River in Jefferson County to a maximum elevation of 776 m in southern Allegany County, near the Pennsylvania border (*Fig. 2*). In general, the terrain rises from the relatively flat lake plains around lakes Erie and Ontario into the steeply cut hills of the Allegheny Plateau in the Southern Tier and Finger Lakes regions. Relatively sharp changes in relief are observed downwind of the Great Lakes, most notably across the Tug Hill Plateau.

![Figure 2](image-url) A digital elevation map (DEM) of the NWS Buffalo County Warning Area (CWA) Color-fill depicts elevation in meters (m). Notable geographic features are labeled as follows: A. Tug Hill Plateau; B. Chautauqua Ridge; C. St. Lawrence River; D. Black River Valley; E. Mohawk Valley.
east of Lake Ontario and the Chautauqua Ridge east of Lake Erie. By virtue of their downwind location from the lakes, these prominences are particularly susceptible to orographic enhancement of precipitation. Likewise, the eastern sides of these features can be subject to localized relative precipitation minima; the Black River Valley in Lewis County is a prime example.

In addition to the variations in precipitation induced by topography, an additional complicating factor with regards to the pattern of snowfall across the CWA is the predominance of lake-effect precipitation during the cool season. As noted by Scott and Huff (1996), lake-effect precipitation accounts for a significant increase in the amount of precipitation observed around the lakes, resulting in an approximate 25% increase in winter precipitation downwind of Lake Erie and 30% downwind of Lake Ontario. The compounding effects of orographic enhancement and lake-effect produce pronounced maxima across the Tug Hill Plateau and the Chautauqua Ridge, as noted in Veals and Steenburgh (2015), Niziol et al. (1995), and Schmidlin (1993).

3. Methodology

Academic literature is rich with studies describing and testing various methods of interpolating and plotting climate data; Hartkamp et al. (1999), Daly (2006), and Hofstra et al. (2008) all include comprehensive summaries of the methods of interpolation of climate data, including discussions of their various strengths and weaknesses. In mapping the snowfall climatology of the WFO Buffalo CWA, a series of direct interpolation schemes of the observed snowfall values were tried, including inverse distance weighting (IDW), kriging, and natural neighbor. Of these methods, kriging was discarded due to issues with the significant anisotropy that exists within the snowfall data due to the lake-effect influence of two lakes, making production of an appropriate semi-variogram unfeasible. IDW produced more realistic results than direct kriging of the 30-year snowfall data; however, the interpolation scheme produced multiple bullseyes (Fig. 3a) that are not physically justified given the mean low-level flow and topography existing across the forecast area.

Of the direct interpolation schemes

![Figure 3](image-url)

**Figure 3.** Initial snowfall maps created using (a) IDW and (b) natural neighbor methods (color-fill, inches). Note the multiple bullseyes produced by the IDW method. These are reduced by using the natural neighbor method, however this method tends to “smear” out topographically-induced details.
attempted, the natural neighbor method produced the most realistic results (see Fig. 3b). However, even the natural neighbor method, while capturing the general trends of snowfall quite well, tended to miss the details imposed by topography and lake-effect patterns. This is not an entirely unexpected result; as Daly (2006) points out, methods such as natural neighbor and IDW are more adept at illustrating spatial climate patterns across areas “without major terrain features and at least 100 km from climatically important coastlines.” A more sophisticated scheme was necessary to capture not only the relationship of elevation and snowfall, but aspect, slope, and position relative to the lakes.

The Parameter-elevation Regressions on Independent Slopes Model (PRISM 2016), developed at Oregon State University, is an ideal candidate for modeling the spatial distribution of precipitation across the Buffalo CWA. Utilizing the technique of Climatologically-Aided Interpolation (CAI), PRISM incorporates a number of climatologically important factors, including terrain height, slope orientation (referred to in PRISM literature as “topographic facets”), and coastal proximity (Daly et al. 2002; Daly 2002). The PRISM dataset is peer-reviewed and represents the best in mapping of spatial climate patterns across the United States (Daly et al. 2008). In addition, the PRISM dataset is recognized as the official climatological dataset of the United States Department of Agriculture (UCAR 2017). Unfortunately, there is no PRISM dataset for snowfall climatology; however, as will be demonstrated in this study, PRISM climatological precipitation distribution data can be utilized as a proxy for snowfall distribution by performing a regression analysis of the PRISM precipitation data against observed snowfall climatology data and then adjusting the results of the regression by using an interpolated map of the residuals from the regression to produce a final snowfall map. Finally, using a precipitation-based dataset versus a snowfall-based dataset takes advantage of the significantly denser precipitation observing network, allowing for greater resolution of details in precipitation, and thereby, snowfall, patterns, than would be possible using

![Figure 4. NCEI 1981-2010 climatological stations, color-coded by precipitation type measured. The red and purple points are stations used in this study.](image-url)
snowfall data alone. For example, within the United States portion of the study area, there are 33% more stations that measure liquid precipitation versus stations that measure snowfall alone, as illustrated in Fig. 4. It can also be observed from Fig. 4 that there are significant gaps in the distribution of snowfall observing sites that are filled by precipitation sites.

4. Data Used

Observed mean annual snowfall data from the NOAA 1981-2010 Climate Normals dataset, provided by the National Climate Data Center (NCEI), for locations in the Buffalo CWA and within a 100 km buffer around the CWA were used. A detailed description of the dataset can be found in Durre et al. (2013). In addition, 1981-2010 mean annual snowfall for Canadian sites was obtained from the Environment Canada (EC) Canadian Climate Normals dataset and included in the 100 km buffer around the CWA. The Canadian data was used for the purposes of expanding the boundaries for the basic interpolation schemes; however, for the final regression analysis, only sites that fell within the domain of the PRISM dataset were used (Fig. 4). PRISM 30-Year 1981-2010 gridded monthly precipitation data, of 800 m resolution, downloaded from the PRISM Climate Group at Oregon State University, were used for the regression analysis.

Figure 5. PRISM-derived mean cold season precipitation (color-fill, inches) (Nov-Mar) – 1981-2010. Data courtesy of PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 1 Jun 2016.
5. Procedure

The first step in creating the snowfall climatology map involved quality control of both the NCEI and Environment Canada snowfall point data. The snowfall data was plotted using Environmental Systems Research Institute’s (ESRI) ArcGIS software, a geographic information system for working with maps and geographic information (ESRI 2017), and examined for inconsistencies against known snowfall climatology patterns. Obvious outliers whose means deviated significantly from neighboring sites and previously known climatology of the area were identified and discarded. In the end this only resulted in removal of 3 sites out of 159 total sites in the study area.

After quality control was complete, the PRISM data were assembled and prepared. Monthly 800 m PRISM 1981-2010 normals for the cool season months of November through March were downloaded. The vast majority of snowfall in the Buffalo CWA occurs during these months; thus, it can be assumed that the precipitation patterns observed in these months would likely correlate closely to snowfall patterns. Although snowfall is also observed in the cool season months of October and April, rain tends to be the predominant precipitation mode, and therefore those months were left out of the analysis. The individual Nov-Mar monthly datasets were then added together to create a seasonal precipitation mean map raster, shown in Fig. 5.

Once the PRISM-derived mean seasonal precipitation map raster has been created, the next step is to generate an ordinary least-squares regression (OLR) equation to relate the observed snowfall values at the snowfall climatological stations to the underlying values of the PRISM raster surface. This equation is then used to generate a new interpolated raster of mean seasonal snowfall across the forecast area. This yields an OLR equation $MAS = 10.680P - 77.623$, where $MAS =$ Mean Annual Snow, and $P =$ PRISM raster value at the observed snowfall site. The model yields a fairly robust adjusted R-squared value of 0.59, with the scatterplot shown in Fig. 6.

While in Fig. 6 the residuals appear to be fairly randomly distributed, Fig. 7, the initial output map, with the residuals superimposed, displays some interesting spatial trends in the distribution of the residuals. PRISM winter seasonal precipitation does a satisfactory job overall at predicting the snowfall totals across the Niagara Frontier and Southern Tier; however, it can be seen that the farther south one goes into Pennsylvania, the model tends to over-predict snowfall, as demonstrated by the negative trends to the residuals. This can also be seen across the far eastern portions of the domain in the Mohawk Valley. Likewise, as one moves from the Rochester metro, and particularly to the eastern Lake Ontario region, the higher residuals indicate that the model under-predicts snowfall. There are a couple of possible physical explanations for these trends. Regarding the lower snowfall amounts in Pennsylvania and the Mohawk Valley, the fact that mean snowfall tends to trend lower the farther south and east one goes, even though total precipitation changes little, means that the assumption that the large majority of winter precipitation falls as snow does not hold true in that area. On the other end of the spectrum, the higher residuals across the Eastern Lake Ontario region could be a function of the predominance of lake-effect snows across the area, with their associated higher snow-to-liquid ratios (Baxter et al., 2005).
Figure 6. (a) Scatterplot showing the distribution of Annual Mean Snowfall vs. PRISM gridded winter precipitation values, and (b) the scatterplot of the standard deviation of the residuals from the OLR model output vs the predicted value.
While the initial output from the OLR model does a relatively decent job of predicting the snowfall across the Buffalo CWA, clearly there is room for improvement. In order to accurately portray the climatological mean snowfall across the CWA, adjustments have to be made to the initial OLR model output. These adjustments are accomplished by creating a spatial interpolation of the OLR residuals and adding that back to the “first-guess” spatial map of the predicted snowfall values, a method demonstrated in Kyriakidis et al. (2001), as well as Ninyerola et al. (2000) and (2007) and abbreviated henceforth as OLRwRC. Three interpolation methods were considered for the spatial interpolation of the residuals: kriging, inverse distance weighting (IDW), and the natural neighbor method. The complexity of precipitation patterns across the study area, driven by the region’s unusual geography, makes it nearly impossible to create the accurate estimate of the spatial autocorrelation present in the residuals required to utilize kriging. Thus, a natural neighbor approach was used to map the residuals. However, the one limitation to the natural neighbor method is that it is constrained to the boundary of the points being interpolated. This resulted in an area of missing data across the northwestern Niagara County (Fig. 8). To resolve this gap, an IDW interpolation of the residuals was also created, and data from this map was used to fill in the gaps in the natural neighbor residual map. The end result is referred to as “Residuals - Meld” in Fig. 9. This final residual map is then added to the original OLR output map to create the final CWA snowfall climatology map, shown in Fig. 10.
**Figure 8.** Inset of map depicting OLR residuals (color-fill, inches), interpolated using a natural neighbor method. This method is constrained by the boundary of the data points, hence the lack of data (white fill) over Niagara County, indicated by the arrow.

**Figure 9.** Map of the residuals (color-fill, inches) from the OLR scheme, interpolated using a natural neighbor method, with IDW interpolated residuals filling in the data void indicated in Fig. 8.
6. Results

The final map is shown in Fig. 10. As can be seen from Fig. 11, the map produced using the OLRwRC scheme more accurately reflects the known effects of topography and position relative to the Great Lakes on distribution of snowfall across the WFO Buffalo CWA, as compared to a more basic interpolation scheme, such as natural neighbor. Of particular note is the success in capturing the snowfall minimum in the upper Genesee Valley (Fig. 12, inset A) and the corresponding relative maxima in the hills of the Allegheny Plateau. The regression scheme also does an excellent job of capturing the maximum of snowfall across the Tug Hill Plateau, and particularly the sharp drop-off in snowfall amounts in the lee of the Tug, across the Black River Valley (Fig. 12, inset B). It additionally shows the diminishing gradient as one moves north into the Saint Lawrence River Valley, to the north of the central east-west axis of Lake Ontario.

While the snowfall map created using the OLRwRC scheme captures the overall pattern of climatological snowfall across the WFO Buffalo CWA successfully, there are likely areas where the model still falls short. There is a minimum noted along the Monroe/Orleans county border that is most likely an artifact of the interpolation scheme, and anecdotal evidence suggests that the snowfall amounts across the towns to the east of Rochester, particularly in Wayne County, are likely underdone. However,
due to the lack of climatological observations across the New York State Thruway corridor between Rochester and Syracuse, this cannot be substantively verified. Another area where the map may not be entirely representative is in the far western Southern Tier (inset C). The inland effects of the local snowfall minimum of 52.5” observed at Silver Creek, located along the Lake Erie shoreline, are likely over-exaggerated. The Silver Creek minimum is likely a function of relatively warmer temperatures stemming from its location on Lake Erie, making it more likely to receive rain instead of snow during marginal winter precipitation events. However, these warmer temperatures do not extend nearly as far inland or as high in elevation as the snowfall map indicates.

A third anomalous minimum is noted south of Westfield. This minimum happens to be co-located with a minimum observed in the PRISM dataset, which is curious given its location along the crest of the Chautauqua Ridge. While the cause of this minimum is not known with certainty, it is hypothesized that it is the result of interpolation between the relatively lower precipitation and snowfall amounts along the Lake Erie shoreline.

**Figure 11.** Comparison of (A) the final mean snowfall climatology map, created using the OLRwRC method, with maps generated using the (B) IDW method, and (C) the natural neighbor method. The PRISM-derived mean cold-season precipitation is shown in (D), for reference. All data is color-filled and expressed in inches.
shoreline and corresponding low terrain in the vicinity of Chautauqua Lake, and a relative dearth of observing sites in the vicinity, with possible interpolation between the two lower elevation (and theoretically lower precipitation) areas cutting across the Chautauqua Ridge, where precipitation should be enhanced instead.

7. Conclusion

It has been approximately 20 years since a climatological snowfall map of the NWS Buffalo CWA was constructed. This original map was constructed utilizing what Daly et al. (2002) refers to as the “human-expert” method, a method that leverages knowledge of local meteorological patterns and terrain features in conjunction with statistical data to manually construct a representation of the snowfall climatology. While this map has held up well over the years, significant improvements in technology and methods for performing spatial analysis on datasets, namely advances in Geospatial Information Systems (GIS), have allowed for the construction of a snowfall climatology map that better captures the subtleties in snowfall.

Figure 12. As in Fig. 10, but with inset detail. Note the detail provided by the OLRwRC scheme in the sheltered areas of (inset A) the upper Genesee Valley, as well as the sharp drop-off in the lee of (inset B) the Tug Hill Plateau. The map is not perfect though – note the lack of snow along (inset C) the Chautauqua Ridge.
distribution across the Buffalo CWA driven by topography, proximity and orientation to the Great Lakes, and synoptic-scale patterns.

A series of direct-interpolation schemes to model the snowfall distribution across the Buffalo CWA based upon NCEI 1981-2010 Climate Normals data were utilized, with varying levels of success. However, in order to construct a more accurate representation of the snowfall distribution across the CWA, it was decided to create an ordinary least-squares regression model that compared the point snowfall climate data to a PRISM-derived gridded cold-season precipitation dataset, as the PRISM dataset is considered the preeminent gridded precipitation climatology dataset currently in use. The residuals from the regression were then interpolated and added back to the regression map to produce the final map. While not without faults, this new map is the most accurate depiction of climatological snowfall distribution across the Buffalo CWA currently available. In addition, the digital format of the map allows for a plethora of future applications.

Acknowledgements

The author would like to thank David Zaff, Science and Operations Officer, NWS Buffalo, NY for his encouragement and feedback throughout this research project. In addition, Steven McLaughlin, Senior Forecaster (retired), NWS Buffalo, NY provided helpful comments regarding the history of his map and feedback on this project. Finally, the author would like to thank Brian Miretzky, NWS Eastern Region Scientific Services Division, as well as the anonymous reviewers for their invaluable suggestions and edits.

References


Daly, C., 2002: Variable influence of terrain on precipitation patterns: delineation and use of effective terrain height in PRISM (white paper).


PRISM Climate Group, Oregon State University, [http://prism.oregonstate.edu](http://prism.oregonstate.edu), created 1 Jun 2016.


Appendix: The Step-by-Step Guide to Producing a Snowfall Climatology Map Using the OLRwRC Method in ArcGIS

Part 1 – Preparing the Data

Snowfall Data
1. NCEI data downloaded into spreadsheet – optionally convert into .csv format.
2. Mean snowfall point data then plotted in ArcMap. All points within 100km of CWA boundary are kept. This is determined using the Buffer tool in ArcMap.
3. Once data is plotted, points are inspected for consistency. Significant outliers with no reasonable cause are discarded.

PRISM Data
4. PRISM 800m resolution 30-year mean precipitation data by month are downloaded for the cool season months (Nov-Mar).
5. The individual Nov-Mar PRISM datasets are added together using the Map Algebra tool to create the seasonal precipitation average.

Part 2 – Executing the Regression

1. Extract the values from the PRISM raster to the CWA Snowfall data point file, using the Extract Values to Points tool.
2. Use the Ordinary Least Squares Regression tool in Spatial Analyst to run the OLS between the Annual Snowfall (dependent variable) and the extracted PRISM value (explanatory variable). This will generate a regression equation.
3. The next step is to apply the formula derived from the OLR procedure in Step 2 to the PRISM data grid using the Map Algebra (Raster Calculator) tool in ArcMap. This gives you the initial CWA snowfall climatology map.
4. To accomplish the residual correction procedure, create an interpolated map of the residuals. Natural Neighbor is the preferred method, as long as the domain of the point data exceeds that of the CWA. In the case of the Buffalo CWA, this is complicated by the proximity of the Great Lakes and Canada, which causes a portion of the CWA to have no data. If this occurs, a second residual interpolation utilizing Inverse Distance Weighting needs to be created. Combine the two maps using the Mosaic to New Raster tool – make sure that the Natural neighbor map is selected last on the list of rasters, and the “Mosaic Operator” method is also set to Last.
5. The final residual raster is then added back into the OLR map using the Map Algebra tool to produce the final map.