Eastern Region Technical Attachment
No. 2021-01
May 2021

Lake Effect Snow Warning Polygon Experiment: Verification

David Church[*]
NOAA/National Weather Service
Buffalo, NY

Abstract

Currently, National Weather Service Offices, including WFO Buffalo, issue Lake Effect Snow Warnings in a zone-based format. However, unlike traditional winter storm systems, lake effect snow typically occurs on a localized scale. Starting in the 2015-2016 winter season, WFO Buffalo gained permission to experiment with creating polygon-based warnings for lake effect snow within the operational zone-based warning framework for events off of Lake Ontario. The office expanded the polygon-based warning experiment to events originating off both Lakes Erie and Ontario in winter 2016-2017, and continued the experiment into winter of 2017-2018. Polygon-based warnings are familiar to other severe weather warning operations within the National Weather Service, and offer potentially attractive benefits to lake effect snow warning applications, such as reduced false alarm area and more specific hazard timing information over zone-based warnings.

With several seasons of events containing parallel zone-based and polygon-based Lake Effect Snow warnings, this study examined the verification of both warning methods to determine if the polygon-based warnings are more skillful than zone-based warnings. Verification for the LES polygons is more complex than standard zone-based warnings, as the polygons allow for temporal variation in the location of the high-impact lake effect snow, and can also be updated more frequently to reflect the most current forecast information available. A spatial verification scheme is presented here that allows for equitable scoring of zone-based and polygon-based warnings using common statistics, such as probability of detection (POD), false alarm rate (FAR), and critical success index (CSI). In addition, given the unique nature of the LES polygons compared to most other NWS warning products, statistics like spatial and population savings, and the reduction of unnecessary warning time were also examined.

Spatial verification showed a statistically significant increase in skill for polygon-based warnings over legacy zone-based warnings. Population statistics showed that the number of people that were correctly removed from the zone-based warning was much greater than the number of people that were incorrectly removed from the zone-based warning. Finally, polygon-based warnings were found to reduce unnecessary warning time by approximately half over zone-based warnings. The verification statistics presented show that there is sufficient forecast skill and significant value in switching to a polygon-based warning scheme from the current zone-based warning scheme.

---

[*] *Current Affiliation – NOAA/NWS Weather Forecast Office Salt Lake City, UT*

1. Introduction

Lake effect snow (LES) is a unique winter event, which can cause localized extreme winter weather with high impact to the public and commerce. Snowfall rates of up to three inches per hour can be common in these snow bands and, in extreme cases, can exceed six inches per hour. Snowfall rate often can have a higher impact than actual snowfall totals. In many cases, the cores of these bands of snow have widths of only ten to twenty miles across (Niziol et al. 1995).

WFO Buffalo (WFO BUF) and other offices serving the Great Lakes issue long-fused Lake Effect Snow Watches, Warnings and Advisories on a zone-by-zone basis, where a zone is often defined as part or all of a county. Any warning related to lake effect is challenged by the localized nature of LES, with significant portions of zones within the warning often not receiving substantial, if any, snowfall. Another issue is the transient nature that some of these events can exhibit, as shifting winds move the high-impact intense snows across a narrow region.

In order to address the above limitations of zone-based warnings, WFO BUF experimented with issuing polygons within Lake Effect Snow Warning products for three winters (2015-2016, 2016-2017, and 2017-2018) to delineate more precise locations and timing for the highest impact from LES. The goal of the experiment was to provide enhanced decision support information focusing on impact areas bordering Lake Ontario (all three winters) and Lake Erie (starting with winter 2016-2017), while reducing the area of false alarm and thereby increasing the effectiveness of the warning. Of particular benefit with the polygon approach is the ability to reduce the overall time a location is under a warning when compared to a zone-based warning system. The enhanced information provided by the polygon LES warning areas is expected to allow for more organized and cost-effective use of public resources to minimize the effects of these high-impact LES events, particularly on transportation and commerce.

Lake Effect Snow Warning polygons pose an interesting verification challenge, as lake effect snow is impactful on time and area scales greater than severe thunderstorms, tornadoes, or flash floods that currently utilize polygon-based warnings, but smaller in time and area scales for widespread synoptic snowfall (NOAA 2017a). The current verification scheme defines a zone-based lake effect snow event as a single snowfall report of 7 inches in 12 hours or 9 inches in 24 hours within a warned forecast zone. This scheme breaks down when evaluating a polygon-based approach, as multiple shorter-in-duration and smaller-in-area polygons may cover the same time span as the zone-based warning.

One particular challenge of point verification is that reliable snowfall reports from trained spotter networks, such as the NWS Cooperative Observer Program (COOP) and the Community Collaborative Rain, Hail & Snow Network (CoCoRaHS), are often reported in 24-hour intervals. While intermediate reports from these sources, as well as from the public and social media, are common, these reports are often irregular in space and time, and are especially scarce overnight. Most polygon based-warnings in the 2016-2018 seasons were valid for between 6 and 18 hours, with a median of 12 hours. Thus, few snow reports may be received during the polygon's valid time, especially for polygons in effect overnight. The 24-hour snowfall reports are also problematic, as bands of heavy lake effect snow are often mobile, making identifying the time span of accumulation within that 24-hour window difficult.

Another problem with the current zone-based verification scheme is that it does not address the main disadvantage to zone-based warnings, which is the large false alarm area when lake effect snow covers only a portion of a zone. Traditional verification of zone-based warnings usually has a high POD and fairly low FAR, but can be easy to achieve by "casting a wide net". One of the main goals of the LES polygon experiment is to evaluate whether this concept can be an effective method to reduce user-perceived FAR, and provide more specific and actionable information to increase the effectiveness of the Lake Effect Snow Warning.

2. Methodology

2.1 Spatial Verification

With the above challenges in mind, WFO BUF (Figure 1) developed a spatial verification scheme, which scores both the polygon-based and zone-based warnings on spatial skill, allowing for side-by-side statistical comparison of the two warning techniques. In order to accomplish this, for every lake effect snow event, an area of impact is defined during the valid time of an LES polygon warning. If an event occurred without a warning in effect, this would result in a POD of zero. If a warning was in effect and no event occurred, this would result in an FAR of one. For all events with a polygon warning, this verifying event area is then compared to both the area of the warning polygon and the corresponding warned zones. The probability of detection (POD) is defined in this spatial context as the percentage of the verifying event area that is covered by a warning (polygon or zone based). The false alarm rate (FAR) is defined as the percentage of warned area that does not fall inside of the verifying event area. The critical success index (CSI) is then the combination of POD and FAR that follows the NWS definition for storm-based warnings (NOAA, 2017b). This concept is illustrated with an example in Figure 2 for a polygon-based warning, and in Figure 3 for the matching zone-based warning. The main benefit to this spatial verification scheme is that correctly reducing the warned area is rewarded by lowering the FAR, while failing to warn all of the impacted area results in a lowering of the POD. Using this verification scheme will give a fair assessment as to whether the polygon-based warnings were effective in correctly reducing the warned area. Comparing the CSI will give a one-to-one comparison between the zone based and polygon-based warning as to which showed more skill.

A complete LES warning event will consist of an initial warning, followed by several updates leading up to and through the event. For each update, forecasters can alter and refine polygons within the official zone-based warning area. In general, polygons typically shrink in areal coverage with time to more closely match the expected dimensions of the LES event as forecaster confidence increases. Polygon timing may also change as forecasters attempt to focus on the area and time of highest LES impact. Due to this evolving nature of the LES polygons, we had to choose how to verify potentially hundreds of polygons issued during the course of an event. In this study, we have chosen to spatially verify just the first-issuance warning polygons. First-issuance polygons are those issued when any new zone-based warning is issued (VTEC of NEW) or when a warning is expanded to include additional zones (VTEC of EXB) (Figure 4) (NOAA 2017c). These initial polygons are issued at the longest lead time and will not benefit from the additional refinements as forecast confidence improves later into the event (polygons that were issued with a WSW VTEC of CON). In some long-duration events, these warning polygons may have been issued 2 to 3 days in advance, and in some instances may be the worst verifying polygons. If the initial polygon-based warnings are at least comparable in skill to their zone-based counterparts, then they presumably provide better service.

In order to perform the spatial verification, local storm reports (LSR) of snowfall and radar reflectivity data from the Buffalo, NY (KBUF) and Montague, NY (KTYX) radars were time matched to each initial issuance polygon. LSRs were gathered from 1 hour after the start time of the polygon to 4 hours after the end time of the polygon from the archive on the Iowa State website (https://mesonet.agron.iastate.edu/lsr/). The time-delays were intended to help account for some of the latency in snowfall reports. Radar data from KBUF was used for polygons covering Cayuga County, NY (Figure 1) west and south, while KTYX was used for areas covering Monroe County, NY east. This results in a little overlap, with both radars used for lake bands from Monroe east to Cayuga County in an attempt to handle low topped multi-banded events that may be picked up by one radar and not the other. With this in mind, hourly Level II reflectivity data was downloaded to the NCEI Weather and Climate Toolkit (WCT) (Ansari et al. 2009) via the NOAA Big Data Amazon archive (Ansari et al. 2018) from the start to end time of each polygon warning. The math tool in the WCT was used to create an average reflectivity image for the polygon duration and exported as a KMZ file. Finally, archived polygons in the form of a shapefile, snow spotter reports in the form of CSV files, and archived radar data in the form of KMZ files, both hourly and averaged, were viewed in Google Earth. An example of this process can be viewed in Figure 5.

One can then compare the local storm report information and radar displays in Google Earth to determine the area impacted by heavy lake effect for every event. The local storm report information provides ground truth for the snowfall accumulations. However, they can be limited in helpfulness by the fact they are plotted by report time and do not indicate when the heavy snow actually fell. The average radar reflectivity provides a continuous field that can help to define the edges of high impact snowfall. When also accounting for trustworthy LSRs, one can calibrate a reflectivity threshold to a meaningful snowfall value. The combination of the average radar reflectivity and hourly reflectivity can be helpful in eliminating untimely LSRs that do not represent the location of heavy snowfall during the valid time of the polygon. Weighing this information, one can draw a verification polygon in Google Earth that corresponds to the area of high impact during the valid time of the warning polygon, which is saved as a KML file.

A python program was written for ArcGIS Desktop to perform the spatial verification. The program takes inputs of the polygon-based warning as a KML file, the verification polygon KML file, and the corresponding zones within the warning. All polygons are clipped to the land area of the WFO Buffalo county warning area (CWA). The intersection is then found between the polygon warning and the area impacted (the verification polygon); the intersection is also found between the area impacted and the zone-based warning. With this information, the spatial verification can be computed as follows for both polygon-based warnings and zone-based warnings:

$$\boldsymbol{POD_{spatial}} = \frac{\textbf{Total Area Correctly Warned}}{\textbf{Total Area Impacted}}$$

$$\boldsymbol{FAR_{spatial}} = \mathbf{1} - \left(\frac{\textbf{Total Area Correctly Warned}}{\textbf{Total Area Warned}}\right)$$

$$\boldsymbol{CSI_{spatial}} = \left[\left(\boldsymbol{POD_{spatial}}\right)^{-1} + \left(\mathbf{1} - \boldsymbol{FAR_{spatial}}\right)^{-1} - \mathbf{1}\right]^{-1}$$

In cases where an event occurred but no warning was in effect, no spatial statistics are calculated, and the event will earn a POD of zero and a FAR of zero. Similarly, if no event occurs and a warning was in effect, this will earn a POD of zero and a FAR of one.

In addition to spatial statistics, the python program also compared the respective areas defined above to the 2010 census data (United States Census Bureau 2012) to determine the population impact of polygon and zone-based warnings. Using this information, not only may population-based POD, FAR and CSI be calculated, but also the correct and incorrect reduction in population warned by polygon versus zone-based products.

Given the human subjectivity introduced by the spatial verification method, where one must make a well informed, yet still subjective decision, on the area of impact, five "verifiers" independently produced statistics for this study. The five people were a General Forecaster and the Science and Operations Officer at WFO Buffalo, and three summer internship students. Students were chosen in order to provide less bias in studying the utility for experimental polygons when compared to the forecast staff. Due to staffing limitations, the five verifiers were only available to verify the 2016-2017 season.

2.2 Warning Time Reduction Statistics

Since the polygon-based warnings are not static when compared to their zone-based counterparts, and instead move with expected the heavy lake effect snow, we expect there should be a reduction in the time under warnings for the polygon-based warnings. For example, over the course of a lake-effect snow event, a specific location may see several hours of heavy lake effect snow before the snow moves out of the area and then returns to the area later in the event. In a zone-based warning scheme, this location may be under a warning for the entire time. But for a polygon-based warning scheme, this location may be under a warning initially, have a period of no-impact, and then be under a warning again later in the event. This component of the verification methodology seeks to measure the warning time reduction produced by a polygon-based warning scheme over the zone-based scheme.

In contrast to the spatial verification discussed in the previous section, for warning time reduction calculations, the zero-hour lead time polygon warnings were used; that is, the polygons that were active at each hour from the time the zone-based warning went into effect until the time the zone-based warning was either canceled or expired. Polygon based warnings have a time resolution as small as an hour (time limitation of the Graphical Forecast Editor (GFE) used to produce the product). In order to find the zero-hour lead time polygon, a script was written to step hour-by-hour through the active warning times and extract the polygon that was in effect at the hour the zone-based warning period started. As opposed to the spatial verification method described earlier, which used the first issuance set of polygons only; this approach should be the best case set of polygons since it incorporates all of the forecast refinement that occurred based on later forecast information and increasing forecaster confidence. This set of polygons best illustrate how long any one location was actually under a warning, as polygons updated after the initial issuance may have been updated to exclude areas from the earlier warning. This is similar to cancelling a zone-based warning before its expiration when a lake band moves out of a zone earlier than originally forecast. Also, by utilizing the same start time, the total possible number of warning hours will be the same between the polygon-based warning and the zone-based warning, which gives an equal footing to the compare the two methods.

Once all of the zero lead-time warning polygons were collected into a KML file, this file was imported into ArcGIS Desktop for analysis. A procedure was run to count the number of overlapping polygons. Since there is one polygon per hour, this results in a dataset that represents the total number of hours any given point spent under a polygon-based warning. An example of the polygon-based warning hours can be seen in Figure 6 from Lake Effect Storm "D" from the 2016-2017 season. Zone-based warning hours are assigned to each zone counting from the time the warning went into effect to the time it was either canceled or expired. Continuing the above example from Lake Effect Storm "D" (2016-2017), the corresponding zone-based warning times can be seen in Figure 7. The warning time information for zone and polygon-based warnings were then joined with the 2010 census data (United States Census Bureau 2012). From this combination of data, a population weighted average warning time was calculated that describes the number of hours the average person was under a warning. This calculation is as follows, where the summation is over each polygon segment or zone-based warning (n):

$$\boldsymbol{Average\ Hours\ Warned} =$$

$$\frac{\sum_{i=1}^{n} \boldsymbol{HoursWarned} * \boldsymbol{Population}(\boldsymbol{n})}{\sum_{i=1}^{n} \boldsymbol{Population}(\boldsymbol{n})}$$

For example, consider one hypothetical zone-based warning that was in effect for 10 hours and had a population of 10 people. This equation would yield an average hours warned of 10 hours per person, as we would expect. Now consider a slightly more complex application where that 10-hour zone-based warning was sub-divided into two polygon-based warnings that are smaller than the size of the zone. Let's assume that the first polygon-based warning was in effect for 3 hours and had a population of 5 people, while the second polygon-based warning was in effect for 7 hours and had a population of 2 people. To apply this equation, ((3 hours warned * 5 people) + (7 hours warned * 2 people)) / (5 people + 2 people) will tell us that the average person in this population is warned for 4.14 hours. Also note that this does not account for the reduction in hours warned for the 3 people that were entirely excluded from the polygon-based warning but were included in the zone-based warning. Thus, the reduction in time warned between a zone-based and a polygon-based warning is actually conservative, with the effects of the reduction in population warned are better accounted for in the spatial statistics in part one.

This calculation was performed on an event-by-event basis for the 2016-2018 seasons, which included 21 total lake effect events (labeled as 2016-2017 A through J and 2017-2018 A through L). The comparison is made between the polygon-based approach and the zone-based approach, which yields an average warning time reduction for each event as well as the percentage of time reduced.

3. Results

3.1 Spatial Verification

Using a single verifier (the General Forecaster), spatial verification showed a net advantage to the first issuance polygon-based warnings over the corresponding zone-based warnings (Table 1, Figure 8). For all 21 events of the 2016-2018 seasons, featuring 130 polygons, the average spatial combined skill score (CSI) rose from 0.39 for zone-based warnings to 0.51 for polygon-based warnings. This net improvement did not come without some cost though, as the average spatial probability of detection (POD) fell from 0.87 for zone-based warnings to 0.65 for polygon-based warnings. However, the more significant reduction in spatial false alarm rate (FAR) from 0.59 for zone-based warnings to 0.28 for polygon-based warnings explains the net improvement in the CSI.

While not every polygon-based warning outperformed its zone-based counterpart, the net effect was an improvement in skill. Comparing the distribution of the scores for all 21 events (Figure 8), 75% of the events had a zone-based CSI between 0.25 and 0.52, meanwhile 75% of the events had a polygon-based CSI between 0.39 and 0.68. In addition, of the 130 polygons scored, 100, or about 77%, scored better than their zone-based counterpart. This is visualized in Figure 9, which shows a direct comparison of the polygon-based to zone-based CSI scores for all 130 polygon events verified. Clearly many more polygon warnings verified better than their zone-based counterparts, some significantly so, however it is important to note that 30 of the 130, or just 23%, scored worse than the zone-based warnings.

To add confidence that the polygon-based warnings overall outscored their zone-based counterparts, a T-Test was performed on the polygon-based and zone-based CSI datasets, containing the 130 verifications of the single verifier. This dataset is represented by the two rightmost distributions in Figure 8. The p number of the two-sided T-Test was found to be 1.18e-8, which means the hypothesis that the single-verifier polygon-based warnings CSI overall scored better than the zone-based warnings CSI can be accepted with greater than 99% confidence.

The results above came from the one verifier; however, we also examined the sensitivity of these results to the subjectivity of the verifier. Due to resource limitations, this sensitivity test of verifier subjectivity was only performed on the data from the 2016-2017 season. While we did find the verifier influenced the results some, we found the overall trend of the results was not significantly influenced by the verifier, giving credence to the verification method. For the 5 verifiers, we found the same drop in both POD and FAR, with a mean increase in CSI for polygon-based warnings (Figure 10). With 5 separate verifiers, the standard error of the scores was only ±0.05 to ±0.07 across POD, FAR, and CSI (Table 2). In addition, while there is some overlap of the polygon and zone CSIs from the 5 verifiers, the average of all CSI scores for the 76 polygons evaluated showed 55, or 72% of them, scored better than their zone-based counterpart. Note that this ratio is very similar to the single-verifier result discussed above, of 77% of polygons scoring better than their zone-based counterpart across two seasons.

To add confidence that the result of improved CSI for polygon-based warnings is not significantly affected by the subjectivity of the verifier, a T-Test was also performed on the average polygon-based and zone-based CSI datasets of the 5 verifiers. First, the polygon-based and zone-based CSI scores for the 5 verifiers were averaged for each of the 76 polygons. The T-Test was then performed on these two datasets, which are roughly represented by the two rightmost distributions in Figure 10, to determine if the difference between the two are statistically significant. The p number of the two-sided test was found to be 0.0035, which means the hypothesis that the polygon-based warnings CSI overall scored better than the zone-based warnings CSI can be accepted with 99% confidence.

Another interesting statistic that falls out of the spatial verification is the number of people that were correctly excluded from each warning versus the number of people that were incorrectly excluded from each warning (Figure 11). While the spatial statistics above show we are providing some added skill with the polygon-based warnings, it really comes down to how people are impacted by these warnings. The median first issuance polygon correctly excluded 196,888 people that would have been included in the corresponding zone-based warning. Meanwhile, the median first issuance polygon incorrectly excluded just 1,624 people that would have been in the corresponding zone-based warning. The ratio of people correctly to incorrectly excluded in the median first-issuance polygon is approximately 121:1.

The results shown here provide clarity that not only do polygon-based warnings out perform their zone-based counterparts in spatial verification, but that any subjectivity in the verification process does not significantly impact the results. The major benefit to the added spatial skill of the polygon warning is the considerable reduction in over-warned population compared to the zone-based warning.

3.2 Warning Time Reduction Statistics

The reduction of time under-a-warning produced by the polygon-based warning method was significant over the zone-based warning method (Table 3). For all 20 events of the 2016-2018 season, the average time a person was under a zone-based warning was 29.3 hours, while for a polygon-based warning the average person only spent 16.1 hours under the warning. This resulted in an average warning time reduction of about 42% over the course of the two seasons. Impressively, the distribution of warning times shifted down markedly. When looking at all 20 events, 75 percent of the events had an average polygon warning time of 20 hours per person or less, while 75 percent of the zone-based warning times were 19 hours per person or more (Figure 12).

When examining long-duration events, which are arbitrarily defined here as events with an average zone-based warning time longer than 36 hours, the time reductions are actually even greater. During the 2016-2018 seasons, there were 6 events that qualified as long-duration events (Table 3). The zone-based warning time of these events was 47.7 hours per person, while their corresponding polygon-based warning time was only 22.7 hours per person. This is a savings of 52%, and equates to nearly an entire day. The greatest time savings was found in event H from 2016-2017, where the polygon-based warnings saved the average person 36.4 hours of warning time, which was a reduction of 62.8%.

One event, I of 2016-2017, resulted in a net zero-time savings (Table 3). This event was a single polygon, short duration event that lasted 8 hours. No time savings can be realized from a single polygon event, as the polygon warning time and zone warning time will always be equal. While this event shows no time savings, the main utility of the polygon warning was the greatly reduced number of people warned; the polygon warning for this event reduced the warned population from 863,534 to just 88,290, or a savings of about 90%.

There is also a fairly linear relationship ($y=0.56x-3.26$, $r^2=0.68$) between the total length of hours warned in an event and the total number of hours saved per person by the polygon-based warning (Figure 13). This linear relationship shows we can expect that for every zone-based warning hour, there are about half as many polygon warning hours. It makes sense that this relationship should be possible, as the longer the duration of an event the more warning hours are capable of being saved. As discussed above, the short duration event, 2016-17 Event I, showed no savings, meanwhile the longest duration event, 2016-17 Event H, showed the most savings at 36.4 hours. Figure 13 shows all the events in between fall fairly neatly along the trend line. This also shows that the added specificity of the polygons moving in time has a clear advantage over zone-based warnings, especially in long duration events.

4. Discussion and Conclusion

The primary goal of issuing polygon-based Lake Effect Snow Warnings is to provide enhanced information focusing on the highest impact areas bordering Lake Ontario and Lake Erie, while reducing the area of false alarm and increasing the effectiveness of the warning. We believe this will result in more organized and cost-effective use of public resources to minimize the effects of these high-impact LES events on transportation and commerce.

The verification results from the 2016-2018 seasons, in which WFO BUF issued experimental lake effect polygon warnings off of both Lake Erie and Lake Ontario, showed that polygon-based warnings effectively reduced false alarm area while providing enhanced information to high impact areas. Two forms of verification were conducted. One focused on a spatial verification of all first-issuance polygon-based warnings and their corresponding zone-based warnings, while the other focused on a population-weighted warning time savings calculation of zero-hour polygon-based warnings compared to their corresponding zone-based warnings.

The spatial verification demonstrated considerable reduction in false alarm area in the polygon-based warnings over the zone-based warnings. For the 21 events of the 2016-2018 seasons, a 59% false alarm area in the zone-based warnings was reduced to just 28% for the polygon-based warnings. This reduction did come with a cost though, as the probability of detection fell from 0.87 (zone-based) to 0.65 (polygon-based). However, the CSI indicates that the benefit of the FAR reduction outweighs the POD reduction, as the CSI increased from 0.39 (zone-based) to 0.51 (polygon-based). This result shows an effective reduction in false alarm area with skill value added over the zone-based warnings. When compared to the spatial savings of the polygon-based method to the populations impacted, the number of people that were correctly removed from the zone-based warning was much greater than the number of people that were incorrectly removed from the zone-based warning. On average over all 21 events, the ratio of correctly reduced population to incorrectly reduced population 121:1. So while the reductions in warned area were not perfect, there was skill and added value to the reductions that were made.

Regarding the subjectivity of the verification methodology where the verifier has to make a decision about the area of impact, the POD, FAR, and CSI from each of the 5 verifiers were found to only vary by between ±0.05 to ±0.07 on average for the 2016-2017 season. In addition, both the single-verifier and multiple-verifier datasets showed the same statistically significant improvement in CSI for polygon-based warnings over zone-based warnings.

It is also important to keep in mind that this spatial verification only took into account the first-issuance polygons, and does not reflect the refined polygons subsequently issued. In theory, these first-issuance polygons should be the worst verifying set of polygons issued as they do not benefit from any later adjustments due to new forecast information or increased forecaster confidence, and are issued at the longest lead time ahead of the onset of lake-effect snow. The benefit of the polygon-based warnings over the zone-based warnings that we see in these results will likely only increase if considering the updated and refined suite of polygons issued with warning continuances.

In addition, it is believed the results of the verification will help the forecasters to understand their strengths and weakness in issuing polygon-based warnings. The data collected by this verification will be very helpful in discovering any tendencies that may be able to be corrected. For example, on several cases the POD may be able to be increased by simply increasing the inland extent of the polygon warning, or by adding a little more width to the polygon to account for uncertainty, especially for the longer lead time polygons. This data will be reviewed more carefully and presented to the forecasters to hopefully increase the accuracy of the polygon warnings even more with coming seasons.

The time-savings statistics found a considerable reduction in total warning time experienced by the average person under a warning. The average warning time per person fell from 29.3 hours to just 16.1 hours during the 2016-2018 seasons' events. Most importantly, the longer the total duration of the event, the more substantial the time savings became. For the longer duration events, where zone-based warnings were in place longer than 36 hours, the polygon-based warning saved the average person nearly an entire day of unnecessary warning time. The time-saving information here is even more impressive when you consider this is only valid for people that were actually under a polygon-based warning at some point, and does not factor in the zero-warning time population that were excluded entirely from the polygon warning. This information is very important because it speaks to two main parts of the goal of the product: enhancing the forecast information and increasing the effectiveness of the warning.

This polygon warning experiment ties in very well with NWS efforts to increase focus on Impact Decision Support Services. Consider the impact these new and innovative warnings may have on an individual or perhaps emergency managers and highway departments. As an example, for a long duration lake effect event that lasts 48 hours, in theory a person could be warned by a zone-based warning for all 48 of those hours, but only be impacted for the first 12 hours and the last 12 hours of the event, with a 24-hour break of no weather in the middle. The polygon-based warning could not only reduce their total warning time from 48 to 24 hours, but also provide more detailed timing information. Consider this same theoretical event from a county, state or highway department perspective. This added information of timing and location from the polygon-based warning could help direct resources where they are most needed to reduce the disruption on transportation and commerce.

The verification results from the 2016-2018 seasons have shown the polygon-based warnings have an overall improvement in skill over the zone-based warnings by reducing the false alarm area and increasing the effectiveness of the warning by introducing significant total warning time savings over the zone-based warnings. The verification statistics presented show that there is sufficient forecast skill and significant value in switching to a polygon-based warning scheme from the current zone-based warning scheme.

Acknowledgments

The author would like to thank a team within WFO Buffalo for helping take this project from an idea originally put forth by Tom Niziol and David Zaff to a full multi-year operational test. This includes former and current staff members (in alphabetical order): Jon Hitchcock, Fred Pierce, John Rozbicki, Shawn Smith, and Jeff Wood. Jason Franklin and Judy Levan worked with NWS Eastern Region Scientific Services Division (ER SSD) and Meteorological Services Division (ER MSD) to make the experimental available to the public in real-time. Finally, the author would like to thank Brian Miretzky for his invaluable suggestions for this publication.

References

Ansari, S., C. Hutchins, S. Del Greco, N. Stroumentova, and M. Phillips, 2009: The weather and climate toolkit. 25th Conf. on Interactive Information and Processing Systems (IIPS) for Meteorology, Oceanography, and Hydrology, Phoenix, AZ, Amer. Meteor. Soc., 6A.4, https://ams.confex.com/ams/89annual/techprogram/paper_146485.htm

Ansari, S., S. Del Greco, E. Kearns, O. Brown, S. Wilkins, M. Ramamurthy, J. Weber, R. May, J. Sundwall, J. Layton, A. Gold, A. Pasch, and V. Lakshmanan, 2018: Unlocking the Potential of NEXRAD Data through NOAA's Big Data Partnership. *Bull. Amer. Meteor. Soc.*, 99, 189–204, https://doi.org/10.1175/BAMS-D-16-0021.1

Niziol, T.A., W.R. Snyder, and J.S. Waldstreicher, 1995: Winter Weather Forecasting throughout the Eastern United States. Part IV: Lake Effect Snow. *Wea. Forecasting*, 10, 61-77, https://doi.org/10.1175/1520-0434(1995)010<0061:WWFTTE>2.0.CO;2

NOAA, 2017a: National Weather Service directives system: 10-513 WFO winter weather products specification. http://www.nws.noaa.gov/directives/sym/pd01005013curr.pdf Cited 28 September 2017.

NOAA, 2017b: National Weather Service directives system: 10-1601 Verification. http://www.nws.noaa.gov/directives/sym/pd01016001curr.pdf Cited 09 October 2017.

NOAA, 2017c: National Weather Service directives system: 10-1703 Valid Time Event Code (VTEC). https://www.nws.noaa.gov/directives/sym/pd01017003curr.pdf

U.S. Census Bureau, 2010 Census of Population and Housing, Population and Housing Unit Counts, CPH-2-1, United States Summary U.S. Government Printing Office, Washington, DC, 2012

Table 1. Spatial verification statistics for all polygon-based and zone-based warnings issued during the 2016-2017 season and 2017-2018 season, grouped by the 10 lake effect events and 11 lake effect events respectively. Note: Statistics for 2017-2018 Event I are omitted here because verification statistics were not able to be computed. For Event I, the Lake Effect Snow Warning was upgraded to a Blizzard Warning, and Lake Effect Snow Warning Polygons stopped being produced by the Buffalo office.

Season	Event	Polygon POD	Zone POD	Polygon FAR	Zone FAR	Polygon CSI	Zone CSI
2016-2017	A	0.62	0.85	0.06	0.17	0.59	0.70
2016-2017	B	0.58	0.96	0.13	0.42	0.50	0.57
2016-2017	C	0.67	0.84	0.14	0.46	0.59	0.50
2016-2017	D	0.50	0.94	0.16	0.47	0.47	0.50
2016-2017	E	0.76	0.87	0.37	0.69	0.53	0.31
2016-2017	F	0.76	0.97	0.14	0.57	0.67	0.42
2016-2017	G	0.34	1.00	0.01	0.64	0.34	0.36
2016-2017	H	0.50	1.00	0.30	0.69	0.38	0.31
2016-2017	I	0.72	0.99	0.02	0.54	0.71	0.46
2016-2017	J	0.73	0.89	0.19	0.53	0.63	0.45
2017-2018	A	0.49	0.64	0.29	0.59	0.42	0.33
2017-2018	B	0.32	0.63	0.45	0.64	0.29	0.35
2017-2018	C	0.49	0.69	0.55	0.76	0.33	0.24
2017-2018	D	0.96	1.00	0.22	0.41	0.75	0.59
2017-2018	E	0.45	0.87	0.39	0.56	0.33	0.37
2017-2018	F	0.82	0.90	0.39	0.63	0.55	0.37
2017-2018	G	0.78	0.91	0.28	0.60	0.58	0.37
2017-2018	H	0.61	0.63	0.60	0.84	0.37	0.16
2017-2018	I	-	-	-	-	-	-
2017-2018	J	0.85	0.95	0.42	0.61	0.52	0.38
2017-2018	K	0.82	0.97	0.19	0.78	0.64	0.22
2017-2018	L	0.90	1.00	0.54	0.86	0.43	0.14
Season Average 2016-2017		**0.62**	**0.93**	**0.15**	**0.52**	**0.54**	**0.46**
Season Average 2017-2018		**0.68**	**0.84**	**0.39**	**0.66**	**0.47**	**0.32**
Multi-season Average		**0.65**	**0.87**	**0.28**	**0.59**	**0.51**	**0.39**

Table 2. Spatial verification statistics describing the mean and standard error of all 5 verifiers during the 2016-2017 season.

	Polygon POD	Zone POD	Polygon FAR	Zone FAR	Polygon CSI	Zone CSI
Season Average 2016-2017	0.63 ± 0.056	0.93 ± 0.058	0.26 ± 0.077	0.55 ± 0.06	0.50 ± 0.058	0.43 ± 0.058

Table 3. Time reduction statistics for polygon-based versus zone-based warnings. The polygon warned time and zone warned time are the number of hours the average person was warned. The savings warned time shows the number of hours the polygon-based warning saved the average person over the zone-based warning. The savings warned percent shows the relative time reduction as a savings ratio. An asterisk denotes a long-duration event, where the zone warned time was 36 hours or greater. Note: Statistics for 2017-2018 Events I and J are omitted here because time-saving statistics were not able to be computed. For Event I, the Lake Effect Snow Warning was upgraded to a Blizzard Warning, and Lake Effect Snow Warning Polygons stopped being produced by the Buffalo office. For Event J, there was an operational error in producing Lake Effect Snow Warning Polygons during the middle portion of the event, therefore there were not continuously available polygons to compute the time-saving statistics.

		Polygon	**Zone**	**Reductions**	
Season	**Event**	**Warned Time (hr)**	**Warned Time (hr)**	**Warned Time (hr)**	**Warned Percent**
2016-2017	A*	27.12	47.38	20.26	42.77%
2016-2017	B	17.79	31.65	13.86	43.80%
2016-2017	C	9.26	18.03	8.77	48.66%
2016-2017	D	13.83	26.04	12.21	46.90%
2016-2017	E*	19.44	44.58	25.14	56.40%
2016-2017	F*	15.11	40.08	24.98	62.31%
2016-2017	G	14.89	20.00	5.11	25.57%
2016-2017	H*	21.56	57.94	36.38	62.79%
2016-2017	I	8.00	8.00	0.00	0.00%
2016-2017	J	11.20	22.15	10.95	49.49%
2017-2018	A	10.9	14.3	3.4	23.75%
2017-2018	B	25.7	31.6	6.0	18.88%
2017-2018	C	10.6	18.6	8.0	42.90%
2017-2018	D*	36.6	38.1	1.5	3.93%
2017-2018	E	10.3	17.0	6.7	39.45%
2017-2018	F*	30.1	57.9	27.8	48.00%
2017-2018	G	16.5	25.6	9.1	35.62%
2017-2018	H	8.9	18.0	9.1	50.81%
2017-2018	I	-	-	-	-
2017-2018	J	-	-	-	-
2017-2018	K	7.3	21.0	13.7	65.44%
2017-2018	L	7.6	27.0	19.4	71.77%
Multi-Season	**Season Average**	**16.1**	**29.3**	**13.1**	**41.96%**

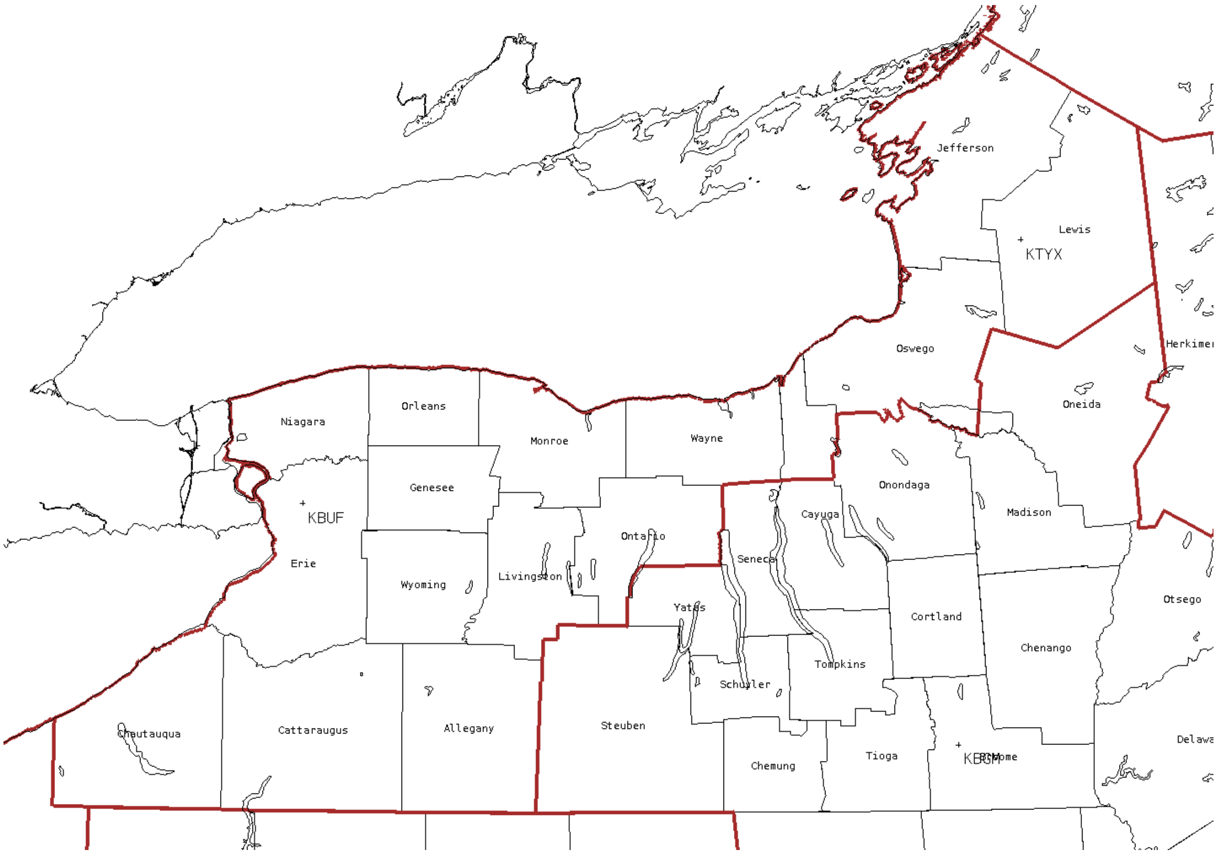
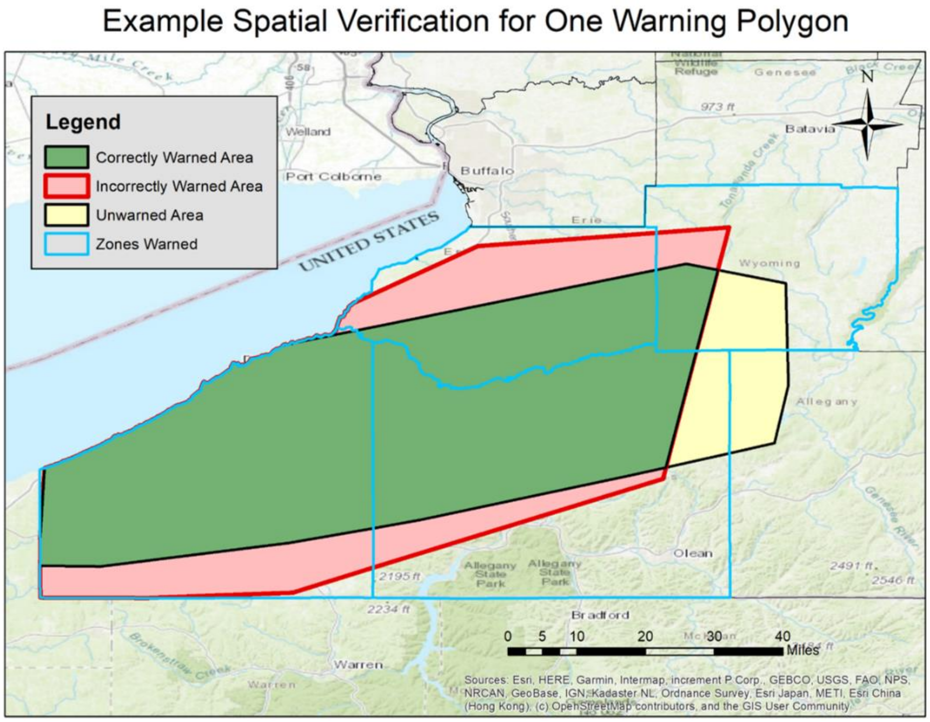
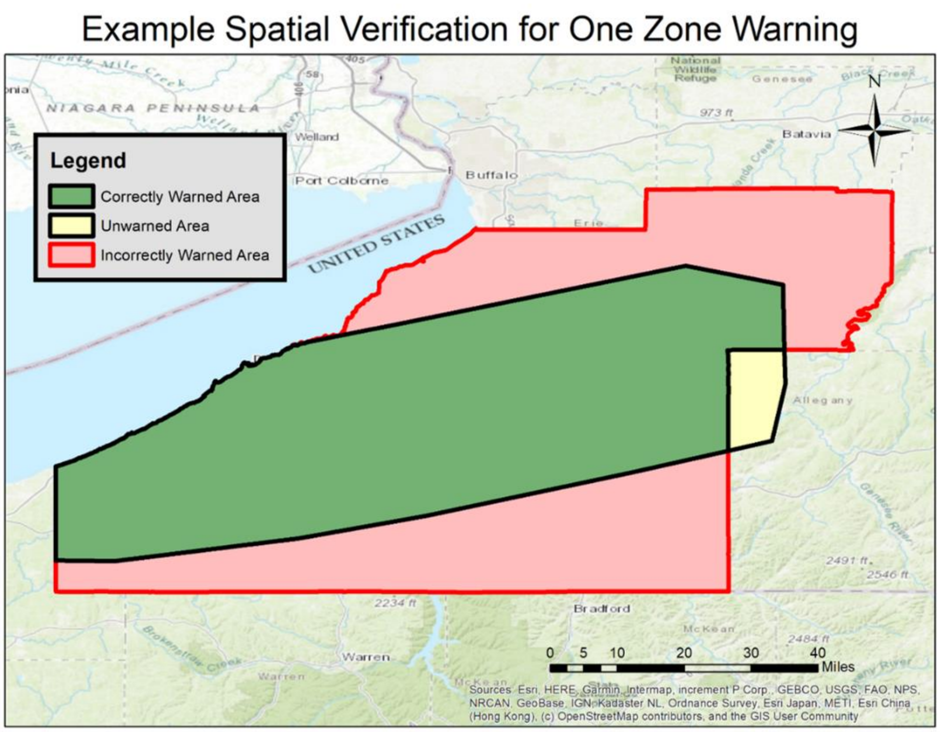


Figure 1. The WFO Buffalo area is comprised of 16 counties shown within the red outline. Cayuga County is split with WFO Binghamton. Zones are identical to the counties, with the exception of Erie County, which is divided into Northern and Southern Erie County (not shown). The KBUF radar is located in Erie County, while KTYX is located in Lewis County.



$$POD_{spatial} = \frac{Green\ Area}{Green+Yellow\ Areas} \qquad FAR_{spatial} = 1 - \left(\frac{Green\ Area}{Green+Red\ Areas}\right)$$

Figure 2. Example of spatial verification scheme produced in ArcGIS for one polygon warning off of Lake Erie, matching the same event as the zone warning verification in Figure 3.



$$POD_{spatial} = \frac{Green\ Area}{Green+Yellow\ Areas} \qquad FAR_{spatial} = 1 - \left(\frac{Green\ Area}{Green+Red\ Areas}\right)$$

Figure 3. Example of spatial verification scheme produced in ArcGIS for one zone warning off of Lake Erie, matching the same event as the polygon warning verification in Figure 2.

```
NYZ012-019-020-085-040400-
/O.UPG.KBUF.LE.A.0001.170104T1800Z-170106T2200Z/
/O.NEW.KBUF.LE.W.0001.170104T1800Z-170106T2200Z/
WYOMING-CHAUTAUQUA-CATTARAUGUS-SOUTHERN ERIE-
INCLUDING THE CITIES OF...WARSAW...JAMESTOWN...OLEAN...
ORCHARD PARK...SPRINGVILLE
301 PM EST TUE JAN 3 2017

...LAKE EFFECT SNOW WARNING IN EFFECT FROM 1 PM WEDNESDAY TO 5 PM
EST FRIDAY...

THE NATIONAL WEATHER SERVICE IN BUFFALO HAS ISSUED A LAKE EFFECT
SNOW WARNING...WHICH IS IN EFFECT FROM 1 PM WEDNESDAY TO 5 PM EST
FRIDAY. THE LAKE EFFECT SNOW WATCH IS NO LONGER IN EFFECT.

* LOCATIONS...CHAUTAUQUA...CATTARAUGUS...WYOMING AND SOUTHERN
  ERIE COUNTIES.

* TIMING...FROM EARLY WEDNESDAY AFTERNOON THROUGH LATE FRIDAY
  AFTERNOON.

* ACCUMULATIONS...UP TO 2 INCHES WEDNESDAY...2 TO 4 INCHES
  WEDNESDAY NIGHT...2 TO 4 INCHES THURSDAY...1 TO 3 INCHES
  THURSDAY NIGHT...AND 1 TO 2 INCHES FRIDAY...LEADING TO STORM
  TOTALS OF 7 TO 14 INCHES IN THE MOST PERSISTENT LAKE SNOWS.

* WINDS...WEST 25 TO 35 MPH WITH GUSTS UP TO 50 MPH. THE STRONGEST
  WINDS WEDNESDAY AFTERNOON AND EVENING.

* VISIBILITIES...NEAR WHITEOUT CONDITIONS AT TIMES.

* IMPACTS...ACCUMULATING LAKE EFFECT SNOW AND GUSTY WINDS WILL
  CREATE VERY DIFFICULT TRAVEL AT TIMES WITH POOR VISIBILITY AND
  SNOW COVERED ROADS.

PRECAUTIONARY/PREPAREDNESS ACTIONS...

IN LAKE EFFECT SNOW THE WEATHER CAN VARY FROM LOCALLY HEAVY SNOW
IN NARROW BANDS TO CLEAR SKIES JUST A FEW MILES AWAY. IF YOU WILL
BE TRAVELING ACROSS THE REGION BE PREPARED FOR RAPID CHANGES IN
ROAD AND VISIBILITY CONDITIONS.

STAY TUNED TO NOAA WEATHER RADIO OR YOUR FAVORITE SOURCE OF
WEATHER INFORMATION FOR THE LATEST UPDATES. ADDITIONAL DETAILS
CAN ALSO BE FOUND AT WWW.WEATHER.GOV/BUFFALO.

&&

EXPERIMENTAL CONTENT BELOW...DO NOT USE OPERATIONALLY
PLEASE SEE BELOW LINKS FOR MORE INFORMATION ON THIS EXPERIMENT
HTTP://WWW.WEATHER.GOV/BUF/POLYGON_PDD
HTTP://WWW.NWS.NOAA.GOV/OS/NOTIFICATION/
PNS15LAKE_EFFECT_SNOWAAA.HTM
TO VIEW THE EXPERIMENTAL POLYGONS PLEASE SEE:
HTTP://WWW.WEATHER.GOV/BUF/LESPOLYGON
COORD...4278 7831 4274 7884 4232 7975 4199 7977
        4201 7923 4225 7845
TIME Y17M01D04T1800Z-Y17M01D05T1200Z
COORD...4260 7822 4284 7848 4281 7890 4232 7975
        4199 7977 4220 7941 4227 7893
TIME Y17M01D05T1200Z-Y17M01D06T0000Z
COORD...4273 7826 4272 7906 4232 7975 4199 7977
        4228 7851
TIME Y17M01D06T0000Z-Y17M01D06T1800Z

$$
```

Figure 4. Example of a WSW product with polygon information. In this example, an advisory (LE.A) was updated to a warning (LE.W for four zones (Wyoming, Chautauqua, Cattaraugus, and Southern Erie. Experimental LES polygons were included at the bottom, as highlighted in yellow.

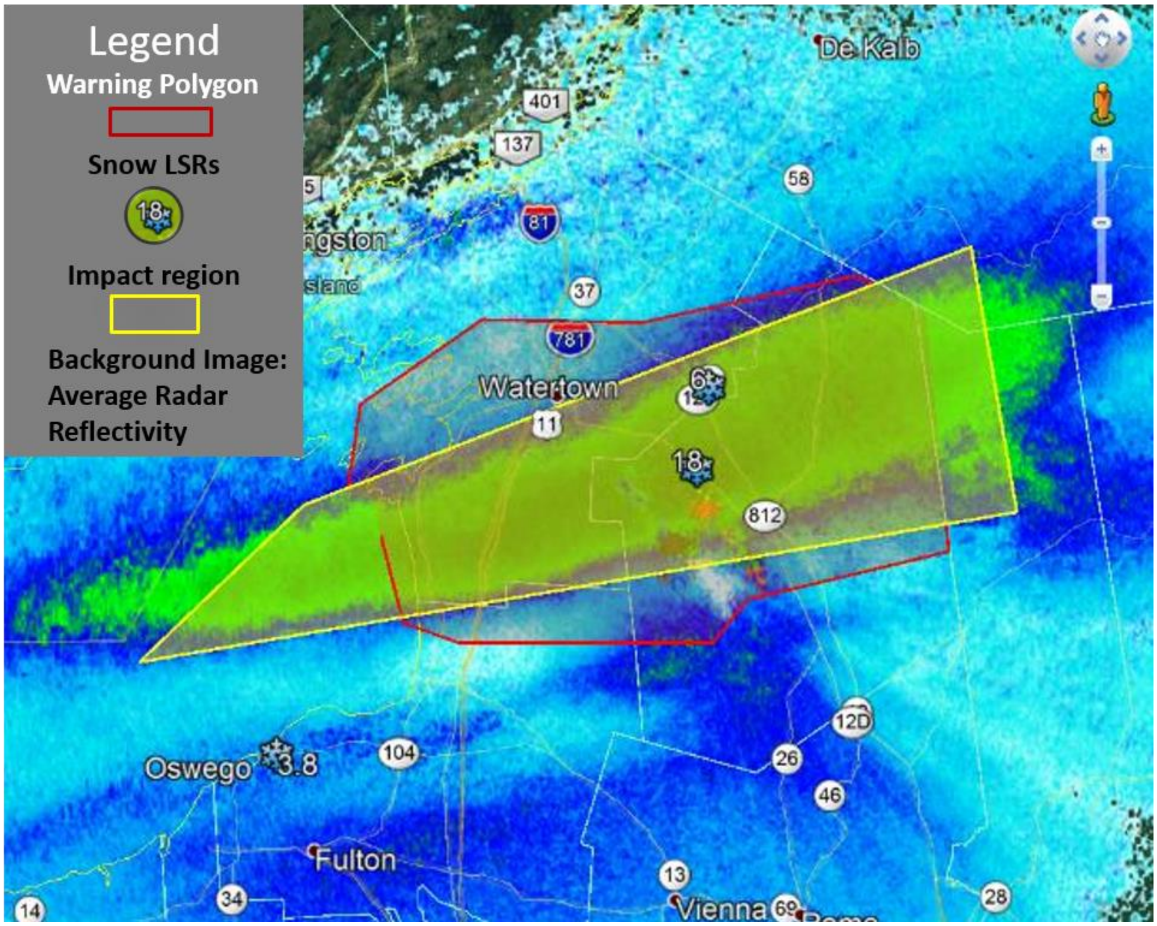


Figure 5. Example of the combined datasets used to verify each warning polygon, as well as the warning polygon (red) and verifier drawn impact region (yellow) that is used to spatially verify the polygon warning.

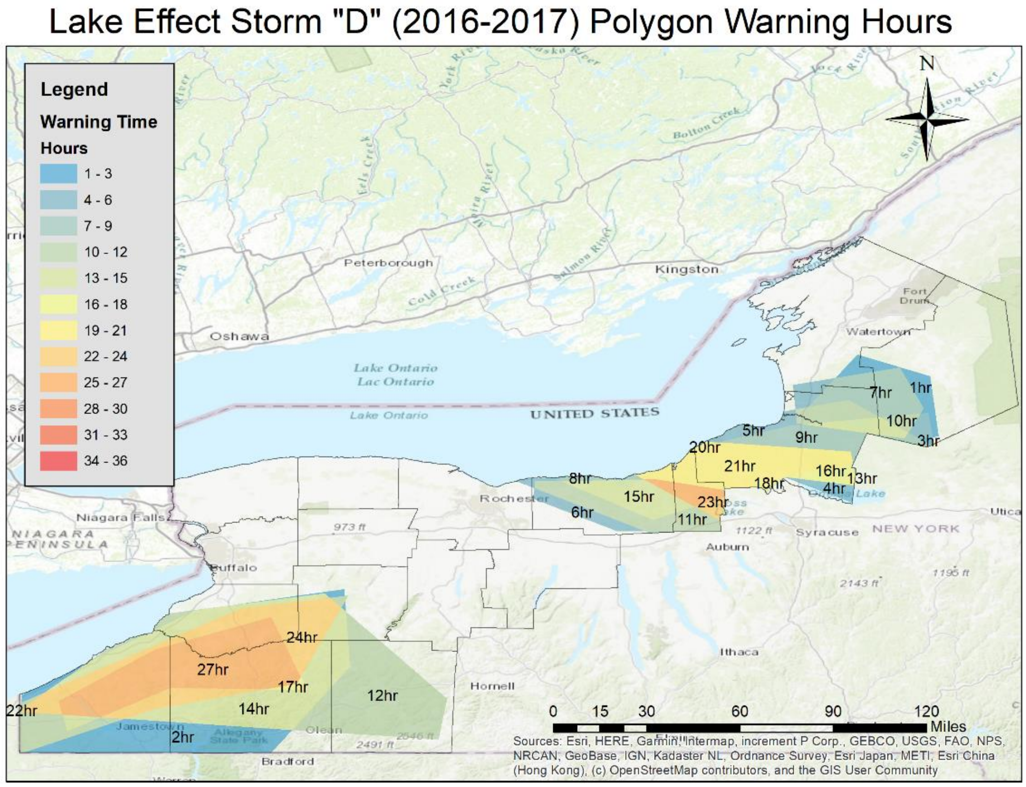


Figure 6. Example of polygon-based warning hours for Lake Effect Storm “D” of 2016-17. Corresponds to the zone-based warning hours displayed in Figure 6.

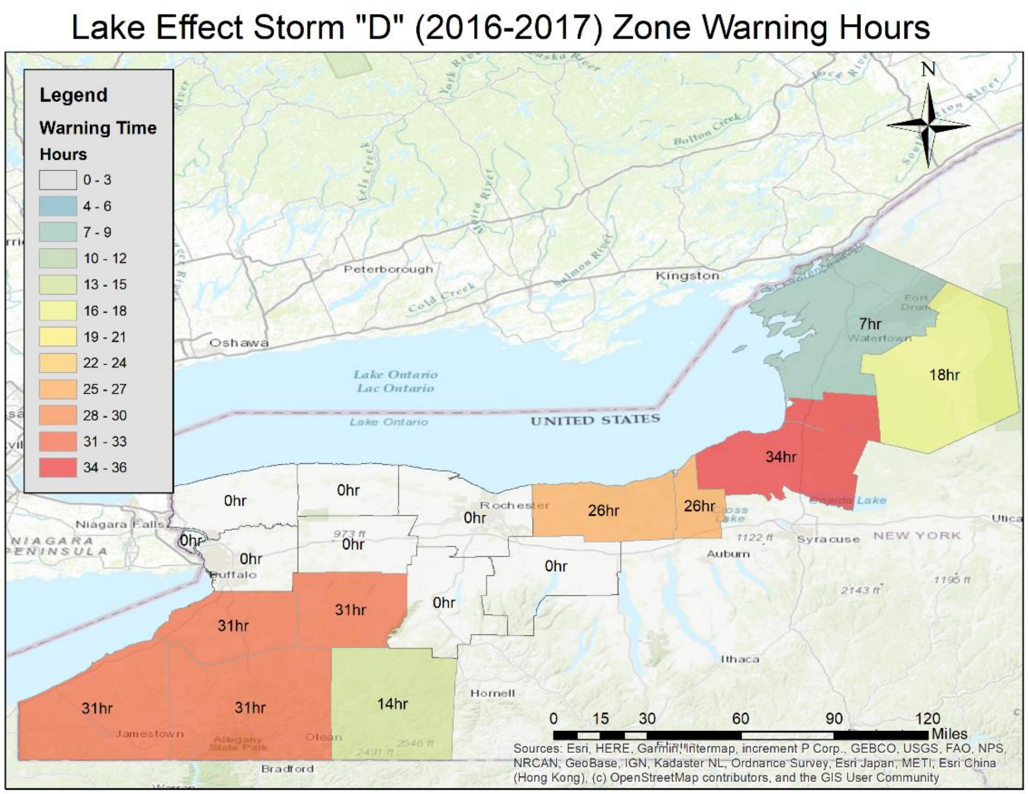


Figure 7. Example of zone-based warning hours for Lake Effect Storm “D” of 2016-17. This corresponds to the polygon-based warning hours shown in Figure 5.

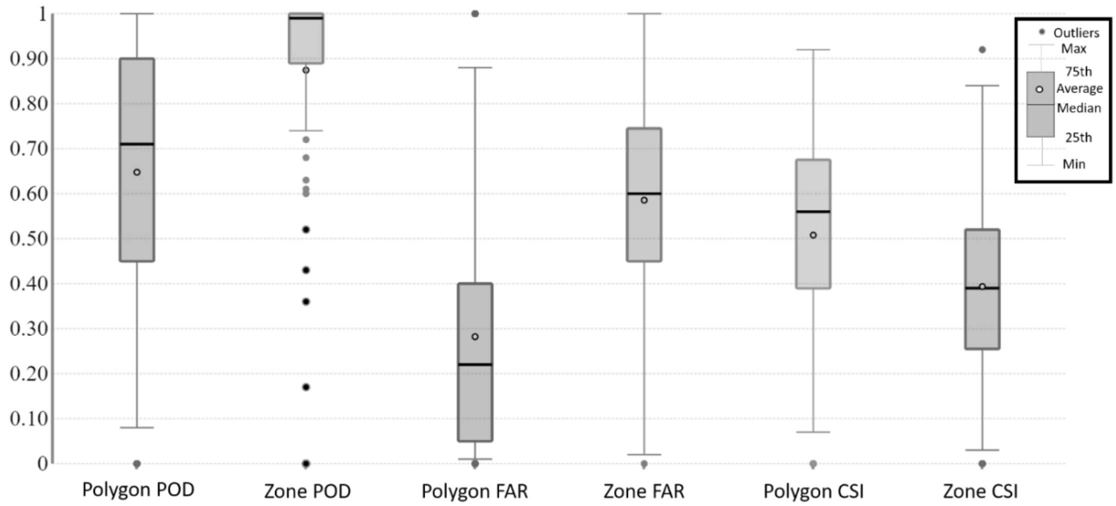


Figure 8. Box and whisker plot comparing polygon-based and zone-based POD, FAR, and CSI statistics for the 21 verified lake effect snow events, containing 130 polygons, of the 2016-2018 seasons.

.

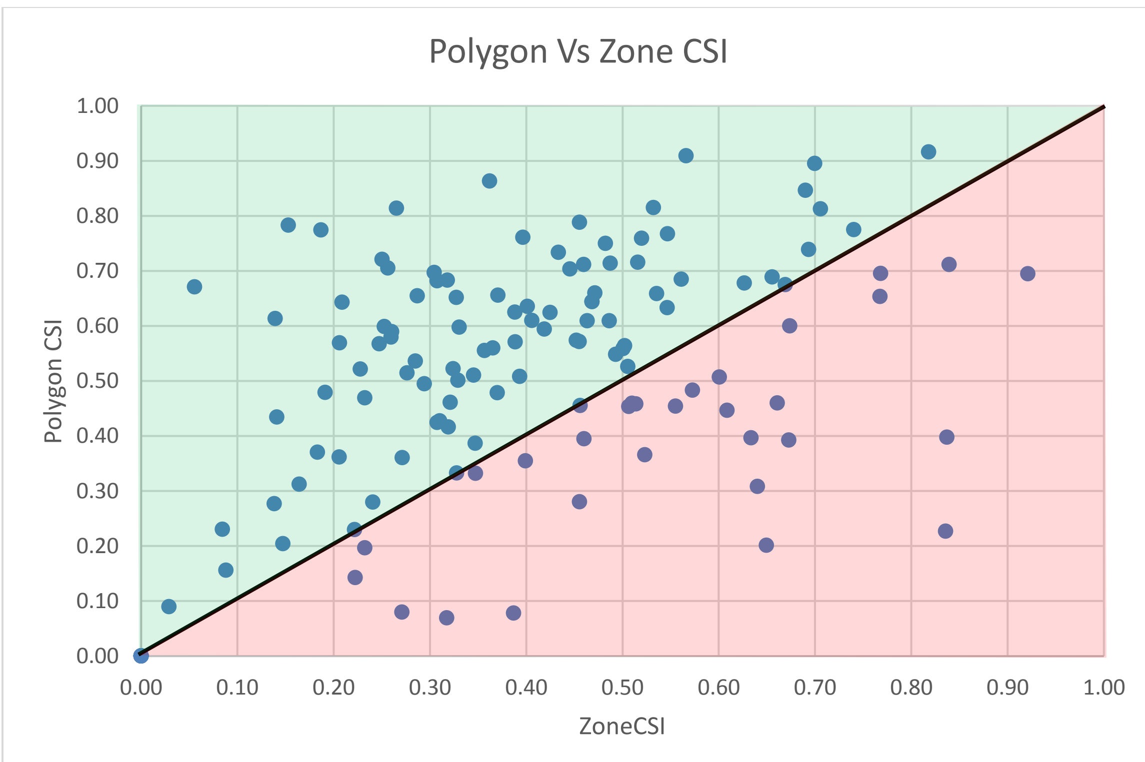


Figure 9. Scatter plot of polygon-based vs zone-based CSI scores for each of the 130 verified polygon-based warnings. Data points falling above the black line in the green-shaded area show polygon-based warnings that spatially verified better than their zone-based counterpart; meanwhile, points falling below the black line, in the red-shaded area, represent polygons that scored worse than the zone-based warning. The farther above (below) the black line, the better (worse) the polygon warnings scored against the zone-based warnings

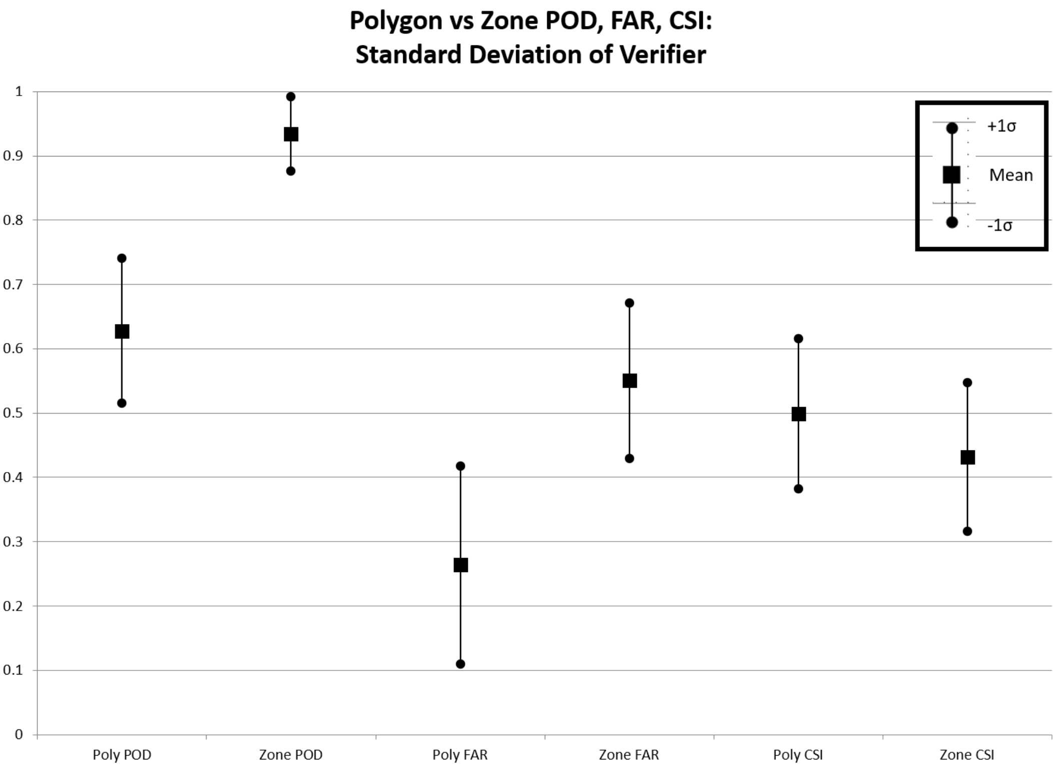


Figure 10. Plot comparing the mean polygon-based and zone-based POD, FAR, and CSI statistics for the 11 lake effect snow events, containing 76 polygons, of the 2016-2017 season, as well as the range of the average 2 standard deviation spread (-1σ to +1σ) of the 5 different verifiers.

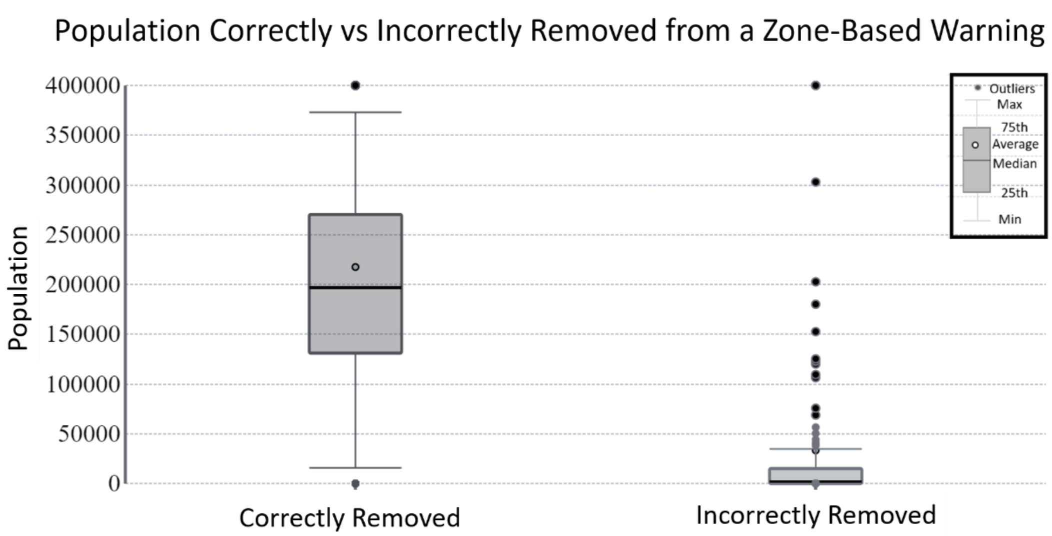


Figure 11. Box and whisker plot showing the distribution of the number of people correctly removed from a legacy zone-based warning by a polygon-based (left) and the number of people incorrectly removed from a legacy zone-based warning by a polygon-based warning (right) for the 21 lake effect snow events, containing 130 polygon-based warnings, of the 2016-2018 seasons.

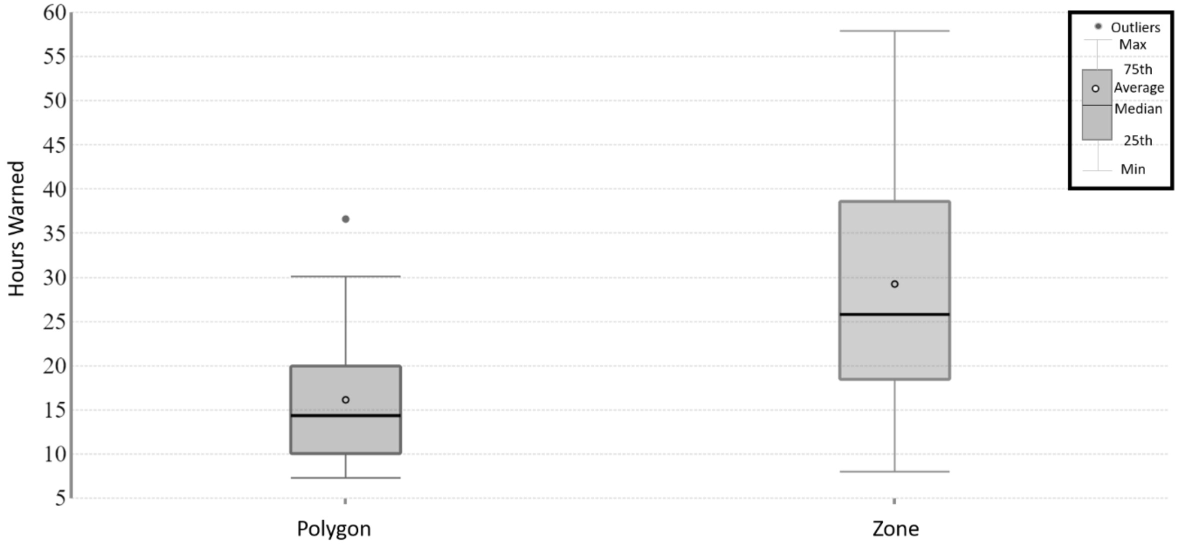


Figure 12. Box and whisker plot showing the distribution of the number of hours warned per person under the polygon-based scheme (left) and the zone-based scheme (right) for the 21 lake effect snow events of the 2016-2018 seasons.

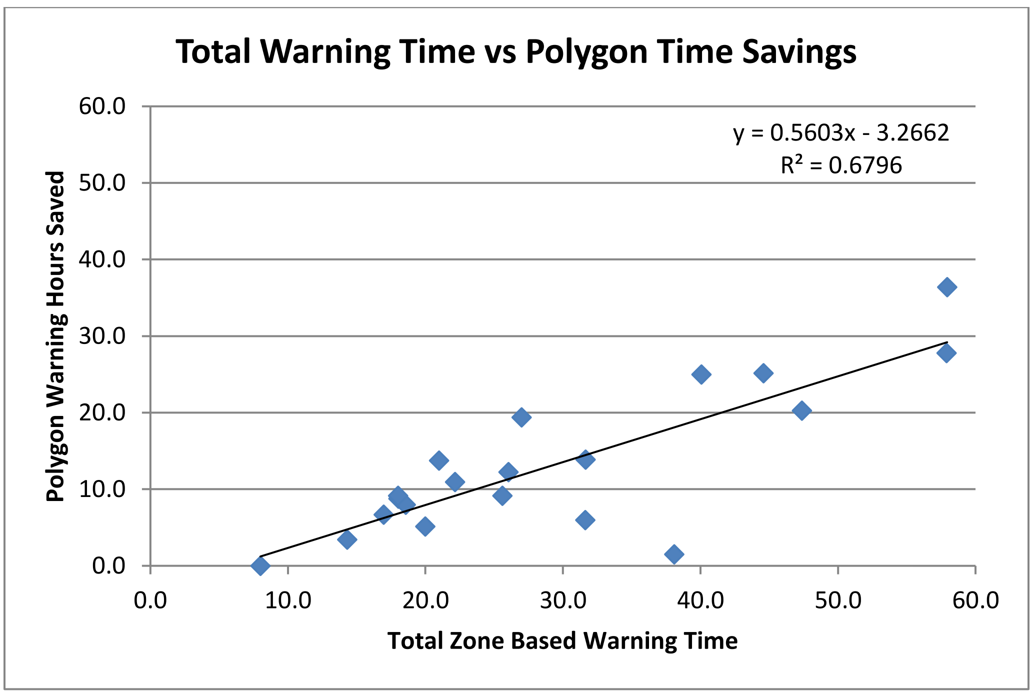


Figure 13. Scatter plot showing the total warning time compared to the number of hours saved per person by the polygon based warning for the 21 lake effect snow events of the 2016-2018 seasons. Linear trend line overlaid is described by the equation and R^2 value shown on the chart.