Embry-Riddle Aeronautical University
SCHOLARLY COMMONS

Journal of Aviation/Aerospace Education & Research

Volume 33
Number 4 *Special Issue: Artificial Intelligence in Aviation*

Article 4

2024

Machine Learning - Hail Awareness Spatial Analysis Toolkit (HASAT)

Haoruo Fu M.S.
Purdue University, fu361@purdue.edu

Joseph P. Hupy Ph.D.
Purdue University, jhupy@purdue.edu

Chien-tsung Lu Ph.D.
Southern Illinois University Carbondale, chientsung.lu@siu.edu

Zhenglei Ji M.S.
New York University, zj2399@nyu.edu

Scholarly Commons Citation

Fu, H., Hupy, J. P., Lu, C., & Ji, Z. (2024). Machine Learning - Hail Awareness Spatial Analysis Toolkit (HASAT). *Journal of Aviation/Aerospace Education & Research, 33*(4). DOI: https://doi.org/10.58940/2329-258X.2041

JAAER

Machine Learning - Hail Awareness Spatial Analysis Toolkit (HASAT)

Haoruo Fu[1a], Joseph P. Hupy[1b], Chien-tsung Lu[2c], Zhenglei Ji[3d]
[1]Purdue University, IN 47907 USA
[2]Southern Illinois University-Carbondale, IL 62901 USA
[3]New York University, NY 10012 USA
[a]fu361@purdue.edu, [b]jhupy@purdue.edu, [c]ctluiapctip@gmail.com, [d]zj2399@nyu.edu

Abstract

The National Airspace System (NAS) is a sophisticated network of air traffic control, navigation, and communication systems that play a critical role in ensuring the safe and efficient flow of air traffic across the United States. However, the occurrence of severe weather conditions, particularly hailstorms, poses a significant threat to flight safety within the NAS. To mitigate the risks associated with hail, aviation organizations have implemented a range of safety measures. This study utilized Esri's ArcGIS as a mapping software to conduct a geospatial analysis of the impact of severe weather, particularly hail, on the NAS. The Hail Awareness Spatial Analysis Toolkit (HASAT), developed as part of this research, leveraged Machine Learning (ML) as a forecasting method to predict the occurrence of severe hail events. The results of the analysis revealed that states such as Texas, Oklahoma, Kansas, and Nebraska emerged as the epicenter of these hailstorms. The Hail Awareness Spatial Analysis Toolkit (HASAT) possessed an additional capability to provide localized hail data to pilots, empowering aviation operators with critical information for flight safety. By incorporating this tool into existing systems, pilots can access real-time, location-specific hail data, enabling them to make informed choices regarding flight routes and potential hazards associated with hailstorms.

Keywords: *Hail Forecast, Severe Weather, Airport Safety, Machine Learning, ArcGIS, Neural Network*

Introduction

The National Transportation Safety Board (NTSB) has consistently found human errors to be the main cause of accidents, with weather-related factors accounting for about 23% of incidents (Kulesa, 2010). Prior to September 11, 2001, weather issues caused around 70% of delays in the National Air Space (NAS). While advanced weather forecast systems have improved decision-making, over 45.8% of NAS delays are still weather-related (Bureau of Transportation Statistics, 2022). Severe weather events like high winds, flooding, heavy downpours, and power outages present challenges to aviation operations, necessitating vigilant monitoring and response from airlines, air traffic control, and other stakeholders to ensure safety.

During natural disasters, airports restrict public access and reroute flights for safety, causing widespread flight disruptions. For pilots, staying informed about Temporary Flight Restrictions (TFRs) and Notices to Airmen (NOTAMs) is vital to navigate severe weather conditions. The Federal Aviation Administration (Federal Aviation Administration, 2022) recommends measures like two-pilot operations, traffic avoidance systems, and readiness for mechanical issues during severe weather. Unlike airborne aircraft, airports are stationary and take extreme precautions during severe weather. Understanding weather patterns is crucial for effective preparation. Historical data analysis and meteorological monitoring enable airports to develop comprehensive plans. Regular training and drills enhance staff and response teams' preparedness. Collaboration among aviation stakeholders, meteorological agencies, and authorities is essential. This need for comprehensive weather preparedness underscores the critical role of advanced analytical tools and technologies in aviation safety. The development and implementation of such tools, including predictive models and spatial analysis systems, are vital for enhancing the accuracy of weather forecasts and the efficacy of operational responses.

Literature Review

The Impact of Hail

Severe weather, often called inclement weather, encompasses hazardous conditions in the NAS, such as thunderstorms, snowstorms, wind shear, icing, and fog. These weather events could significantly impact airline operations and increase operational costs. According to the Federal Aviation Administration, 2023b, a single hour of delay can cost airlines anywhere from $1,400 to $4,500. Winter storms, though dangerous, tend to develop and move slowly. Conversely, summer storms form rapidly, cluster together, and move swiftly, covering extensive airspace. When these storms travel long distances and impact major international airports, the entire nation's aviation system can be affected. Hail, a form of solid ice precipitation, develops within thunderstorm updrafts. Large hailstones can cause damage to building structures, windows, and aircraft, and pose a threat to livestock and human safety (Luo et al., 2022; National Oceanic and Atmospheric Administration, n.d.-b). Hail, as a specific severe weather phenomenon, can be particularly damaging, especially when it comes to larger, denser hailstones. Understanding the characteristics of various severe weather events is crucial for effective risk management and safety measures, allowing aviation stakeholders to better prepare and respond to these challenging weather conditions.

Meteorological Service and Their Impact of Aviation Safety

The National Weather Service (NWS) meteorologists are assigned to Air Route Traffic Control Centers (ARTCC) and part of the Center Weather Service Units (CWSU), as well as the Air Traffic Control System Command Center (ATCSCC). They monitor surface, upper air, and radar weather observations and provide forecasting services nationwide Federal Aviation Administration, 2016. NWS's METARs offer automated observing reports at airports, providing data on temperature, dew point, wind speed, visibility, clouds, and ceilings (National Oceanic and Atmospheric Administration, n.d.-a). Terminal Aerodrome Forecasts (TAFS) provide weather forecasts for over 700 airports, valid for 30 hours and issued every six hours. They include information on wind speed and direction, visibility, ceiling, and precipitation type (National Oceanic and Atmospheric Administration, 2013).

Two essential types of advisories are AIRMETs and SIGMETs. AIRMETs amend surface and cloud forecasts for weather phenomena of operational interest and potential hazards to all aircraft. They cover moderate icing, turbulence, winds of 30 knots or more at the surface, widespread low ceilings, and mountain obscurement. On the other hand, SIGMETs are issued for weather conditions significantly affecting aircraft safety, including severe turbulence, icing, and reduced visibility due to dust or sandstorms (Federal Aviation Administration, 2023a). Accurate weather forecasts are pivotal for aviation safety, enabling informed decision-making to mitigate severe weather risks. The use of METARs, TAFs, AIRMETs, and SIGMETs allows aviation professionals to optimize flight safety and efficiency, highlighting the essential role of meteorological services in flight planning and operations. This integration of advanced forecasting into aviation underscores the critical relationship between meteorology and the industry's commitment to safety.

Geographic Information System

A geographic information system (GIS) is a versatile system that harnesses the power of location data to enhance decision-making and understanding across various disciplines (Esri, 2023). By integrating location data, which indicates the geographical positions of objects, with descriptive information that details the characteristics of these objects, GIS provides a comprehensive and dynamic way to understand spatial relationships and patterns. GIS allows users to visualize and interpret data in a spatial context, making it easier to comprehend complex relationships and make informed decisions (Chen et al., 2015). Moreover, GIS could play a crucial role in connecting flood data with epidemiological information, resulting in improved public health responses following severe weather events that impact the community (Waring et al., 2005).

Prediction Using the Machine Learning and Neural Network

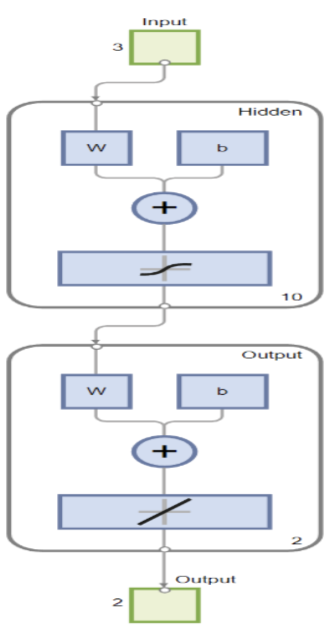
The advanced algorithms of Machine Learning have proven to be highly effective in uncovering complex empirical relationships between various meteorological variables, facilitating climate attribution studies, and improving El Nino forecasting (Scher & Messori, 2018). Machine Learning has revolutionized meteorology by revealing hidden patterns and interactions among atmospheric factors. The neural network, trained on extensive weather data, can predict future conditions by understanding complex relationships between atmospheric variables. This approach is advantageous for modeling the chaotic nature of the atmosphere and offers more accurate predictions than traditional methods due to its ability to handle non-linear datasets (Abhishek et al., 2012). They can directly represent both linear and non-linear relationships from the data being modeled. However, a major drawback is that neural networks demand significant computational power to solve atmospheric equations (Baboo & Shereef, 2010).

Methodology

The main purpose of this study is to find out whether it is feasible to use neural networks to predict hail events in proximity to airport areas, and if so, how accurate are these predictions based on limited variables like time, latitude, and longitude. This study uses a neural network to predict hail impact on airports, focusing on data collected after 1951. The GIS analysis targets Class B, C, and D airspaces with substantial air traffic, randomly splitting the data into 80% for training and 20% for validation tests. Translated into machine learning practices, this split is widely recognized as an effective heuristic to ensure a robust training dataset that accounts for most data variations (the 80%), while setting aside a significant proportion (the 20%) to validate the model's performance and generalize its predictive capabilities (Joseph, 2022). This balance is critical for avoiding model overfitting to the training data and underfitting by ensuring the model is exposed to a comprehensive array of data scenarios (Joseph, 2022; Nguyen et al., 2021). The training utilizes Feedforward Neural Networks, optimizing with the Lavenberg-Marquardt (LM) backpropagation method (Gavin, 2024). The LM method balances gradient-descent and Gauss-Newton methods to minimize the sum of squares of errors between the model function and data points. The Machine Learning processes are performed using MathWorks' MATLAB, a powerful tool with pre-programmed algorithms used by engineers and mathematicians for modeling, testing, and computation. MATLAB's various tool packages aid in selecting suitable algorithms for more accurate predictions. Figure 1 illustrates an example of the Feedforward Neural Network diagram in MATLAB. The adoption of Feedforward Neural Networks and the Lavenberg-Marquardt backpropagation method was guided by their suitability for handling complex, nonlinear data patterns, such as those encountered in meteorological events. This approach, coupled with the robust computational framework provided by MATLAB, allows for the nuanced analysis of spatial and temporal variables, offering a promising avenue for enhancing the accuracy of hail event predictions in critical airport vicinities.

Figure 1

Feedforward Neural Network Diagram



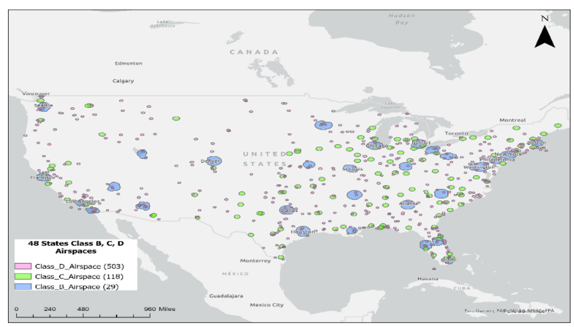
Data for this study was sourced from NOAA's Storm Prediction Center (SPC), encompassing hail events recorded from 1951 to 2021 National Oceanic and Atmospheric Administration, n.d.-c. The dataset includes information such as time, size (expressed in magnitude and diameter in inches), latitude, and longitude of hail occurrences in each state (National Oceanic and Atmospheric Administration, 2023). The data was downloaded and synchronized with Esri's ArcGIS Pro. Following the Machine Learning process, MATLAB generated a list of outputs comprising predicted locations of hail events for accuracy testing.

Results

The contiguous United States (48 states) contains 29 Class B, 118 Class C, and 503 Class D airspaces (refer to Figure 2).

Figure 2

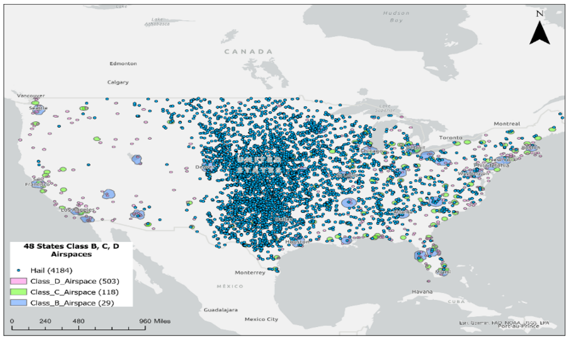
Contiguous United States (48 states) Class B, C, and D Airspaces



With all the selected hail events with a size of 3 inches or larger visualized in the program, the map shows that the Central United States is the epicenter of the severe hail events. Figure 3 represents the location of hail events in the 48 states. Most of the severe hail events occurred in the "Tornado Alley" area, which correlates with geographic conditions for hail formation. In addition, hail frequently occurs before and after tornado events.

Figure 3

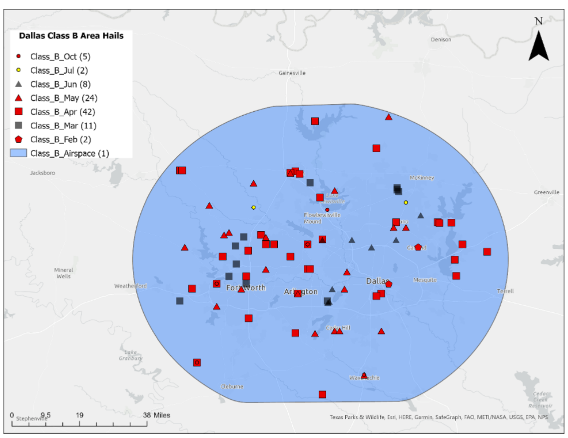
48 States and Hails



Detailed analysis of the airspace's impact by hail events involved selecting specific airspaces. In Figure 4, the spatial distribution of severe hail events within the Dallas Fort Worth Class B Airspace is illustrated.

Figure 4

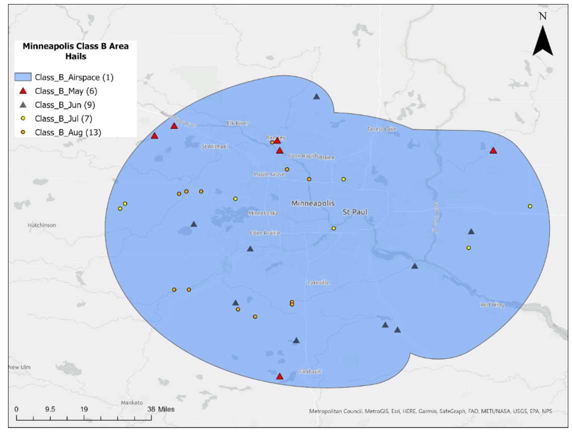
Dallas Class B Airspace Severe Hail Events



Dallas Fort Worth International Airport (DFW/KDFW), one of the largest airports in the central United States, has experienced numerous severe hail events. Historical data reveals that most hail events occurred in spring and summer, with 42 in April (red square) and 24 in May (red triangle). In contrast, no severe hail events were recorded during winter (November, December, and January). Minneapolis Saint Paul International Airport (MSP/KMSP) is a major airport in the Central United States. Figure 5 displays the impact of severe hail events in Minneapolis Class B airspace. Compared to Dallas Class B airspace, Minneapolis had fewer severe hail events. Only four months recorded severe hail events: 13 in August (orange dot), 9 in June (grey triangle), 7 in July (yellow dot), and 6 in May (red triangle).

Figure 5

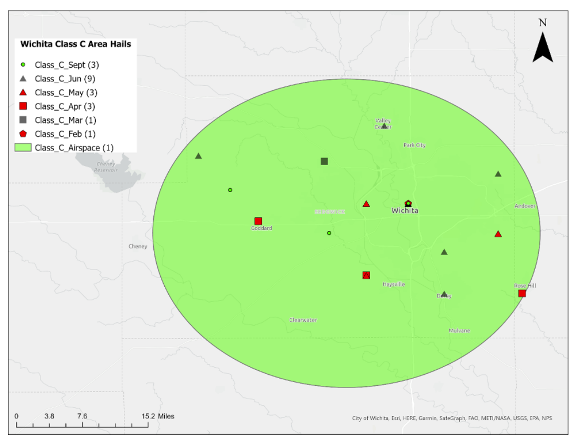
Minneapolis Class B Airspace Severe Hail Events



Although Class C Airspace may be smaller than Class B airspace, it still encompasses several crucial international airports. Wichita Class C airspace (see Figure 6) has June as the month with the highest number of recorded hail events (9).

Figure 6

Wichita Class C Airspace Severe Hail Events

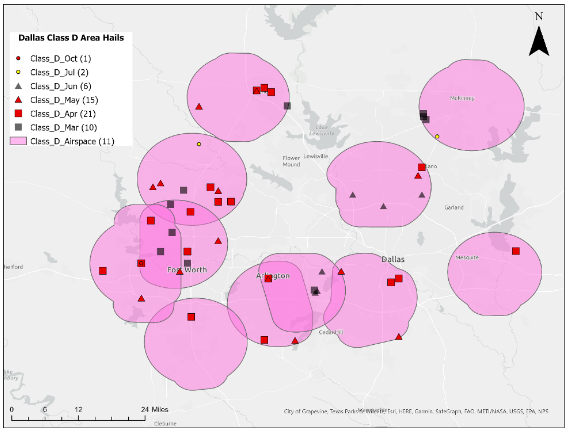


Small airports are mostly situated within Class D airspace, which can accommodate regional flights and student pilot training facilities. Understanding the impact of severe hail events in Class D airspace is crucial. In Figure 7, the impact of severe hail events in Dallas Class D airspaces is depicted. Compared to Class B and C airspaces, Class D airspace is smaller, with 11 selected for analysis. Like the larger Dallas Class B airspace, the most severe hail events occurred in April (21, red square) and May (15, red triangle), indicating that even the smaller, regional segments of the airspace under the umbrella of

Dallas Class B are not immune to the severe weather challenges faced by their larger counterparts.

Figure 7

Dallas Class D Airspaces Severe Hail Events



Machine Learning Prediction Analysis

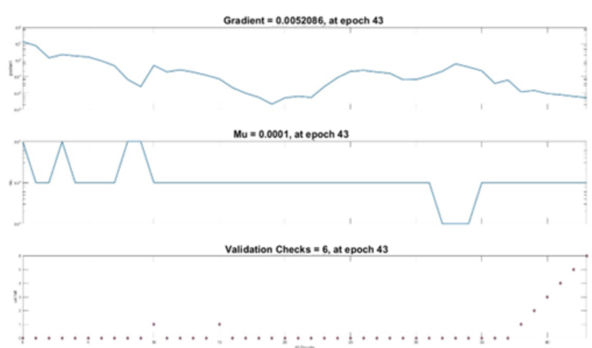
After geospatial visualization of previous severe hail events, machine learning was employed to predict hail occurrences in three tested states. MATLAB generated graphs for each machine learning process, including Training State, Best Validation, and Error Histogram. The predicted hail events were visualized using Esri's ArcGIS Pro based on latitude and longitude. For this study, only hail events larger than 1.5 inches were selected for analysis due to the magnitude of the potential damage. Following the machine learning of severe hail events, the selection of hail events greater than 1.5 inches is strategic, aiming to augment the dataset size for robust machine learning analysis. This threshold aligns with meteorological guidelines that denote hail of this caliber as significantly impactful, particularly concerning aviation safety and operations. By incorporating events of this size, the study ensures a comprehensive dataset, enhancing the predictive model's capacity to generalize and thus improving its utility in practical scenarios. The training state illustrates the neural network's training process and parameter changes over epochs. It includes three graphs:

1. Gradient graph: This represents changes in the gradient of the error function concerning network weights. It indicates the direction and magnitude of changes required to reduce the loss function.
2. Mu graph: This shows changes in the adaptive learning rate parameter, Mu, which dynamically adjusts weight update size during training. Mu is decreased for large gradients to take smaller steps and increased for small gradients to take larger steps.
3. Validation graph: This displays the number of validation checks conducted during training to monitor network performance on a validation set and prevent overfitting. The training process stops if the validation error does not decrease for a specified number of epochs.

The validation graph displays the best validation performance of the process, showcasing the neural network's training progress over time. It measures the mean squared error (MSE) on the training, validation, and test sets. The blue line represents the MSE on the training set, the green line represents the MSE on the validation set, and the red line represents the MSE on the test set. The dashed line indicates the best validation performance achieved by the network thus far. As the neural network undergoes training, the MSE on the training set generally decreases, indicating improved data fitting. The error histogram visualizes the distribution of errors in a prediction model by displaying the number of predictions falling within specific error ranges. It consists of 20 bins, each representing a different error range. The x-axis represents the error range, while the y-axis shows the number of predictions falling within each range. The shape of the histogram resembles a bell curve. When most errors are close to zero, the model demonstrates more accurate predictions. Kansas experienced a total of 10,676 hail events, and the machine-learning process stopped at 43 epochs. Figure 8 illustrates the training state of Kansas.

Figure 8

Kansas ML Training State



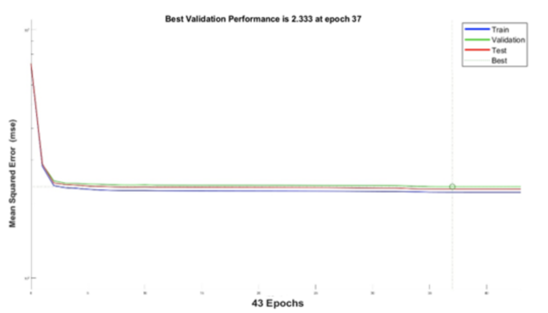
Note: The top graph, labeled 'Gradient', shows the gradient magnitude on the Y-axis, ranging from 10^{-3} to 10^{-2}. The X-axis represents the epoch count. The middle graph, labeled 'Mu,' displays the mu parameter on the Y-axis, with a range from 10^{-5} to 10^{-3}. The X-axis corresponds to the epoch count as in the Gradient graph. The bottom graph, labeled 'Validation Checks,' indicates the count of validation checks on the Y-axis, which varies from 0 to 6.

The X-axis, consistent with the above graphs, denotes the epoch count.

Figure 9 depicts the validation performance of the machine learning process. Among all 43 epochs, the best performance was achieved at the 37th epoch.

Figure 9

Kansas Hail Validation

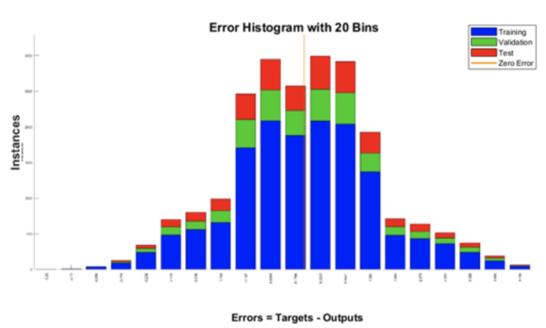


Note: The Y-axis represents the Mean Squared Error (MSE) on a logarithmic scale, ranging from 10^0 to 10^1, which corresponds to 1 to 10 in linear scale. The X-axis denotes the epochs, which is the number of complete passes through the training dataset.

Figure 10 displays the error histogram of the machine learning process for Kansas, featuring 20 bins representing different error ranges. Although it resembles a bell-shaped curve, it is not a perfect one, and some errors occurred during this machine learning process.

Figure 10

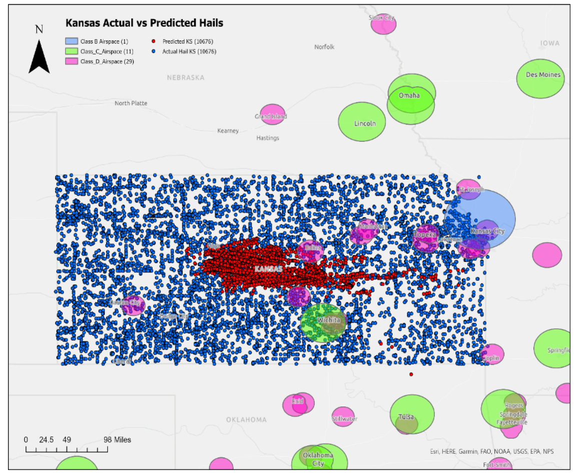
Kansas Hail Error Histogram



After the machine learning validation, ArcGIS Pro was used to visualize all the predicted hail events. Figure 11 compares the predicted (red dots) and actual (blue dots) hail events in Kansas. The results indicate that most predicted hail events are concentrated in the state's middle region. The Root Mean Square Error (RMSE) represents the location differences. In Kansas, the latitude RMSE is 0.827 (approximately 50 miles), and the longitude RMSE is 1.95 (approximately 102 miles).

Figure 11

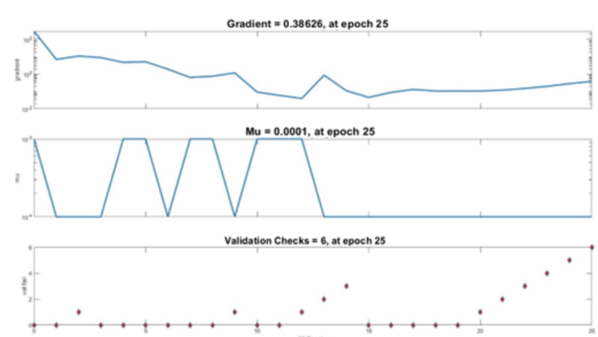
Kansas Actual vs. Predicted Hails



Texas recorded a total of 19,157 hail events, and the machine-learning process concluded at 25 epochs. Figure 12 illustrates the training state of the machine learning process for Texas.

Figure 12

Texas ML Training State

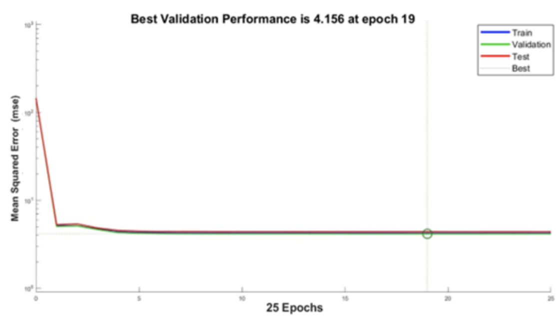


Note: The top graph, labeled 'Gradient', shows the gradient magnitude on the Y-axis, ranging from 10^{-2} to 10^2 the X-axis represents the epoch count. The middle graph, labeled 'Mu,' displays the mu parameter on the Y-axis, with a range from 10^{-4} to 10^{-3}. The X-axis corresponds to the epoch count as in the Gradient graph. The bottom graph, labeled 'Validation Checks,' indicates the count of validation checks on the Y-axis, which varies from 0 to 6. The X-axis, consistent with the above graphs, denotes the epoch count.

Figure 13 depicts the validation performance of the machine learning process. Among all 25 epochs, the best performance was achieved at the 19th epoch.

Figure 13

Texas Hail Validation

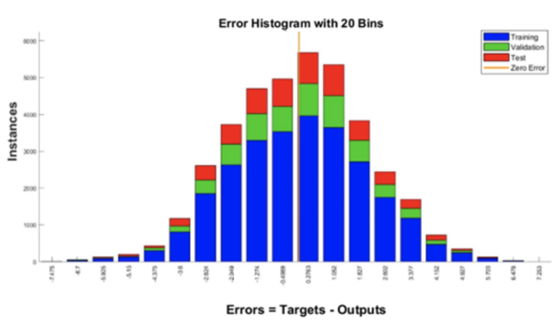


Note: The Y-axis represents the Mean Squared Error (MSE) on a logarithmic scale, ranging from 10^0 to 10^3. The X-axis denotes the epochs, which is the number of complete passes through the training dataset.

Figure 14 displays the error histogram of the machine learning process for Texas, with 20 bins representing different error ranges. It resembles a bell-shaped curve, with the orange line falling in the middle. Compared to Kansas, fewer errors occurred in this machine-learning process.

Figure 14

Texas Error Histogram

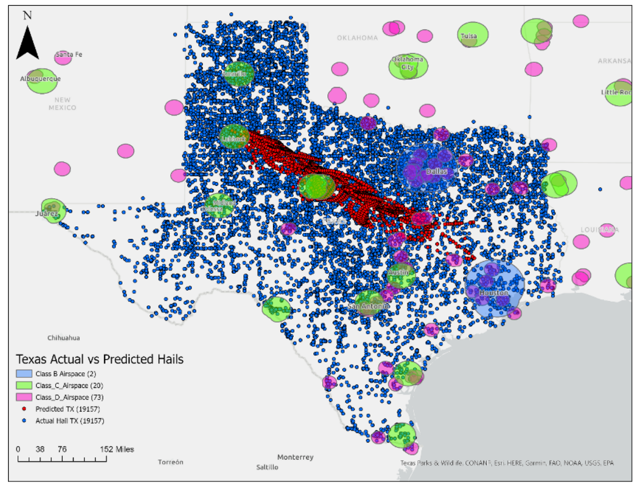


Note: The X-axis represents the range of error values, calculated as Targets minus Outputs, extending from -7.475 to 7.253. The Y-axis quantifies the number of instances (frequency) within each error bin, with a range from 0 to 6,000 instances.

Following the machine learning validation process, ArcGIS Pro visualized all the predicted hail events. Figure 15 compares the predicted (red dots) and actual (blue dots) hail events in Texas. The result indicates that most predicted hail events are concentrated in the middle of the state. The Root Mean Square Error (RMSE) represents the location difference. For Texas, the latitude RMSE is 1.88 (approximately 130 miles), and the longitude RMSE is 2.22 (approximately 135 miles).

Figure 15

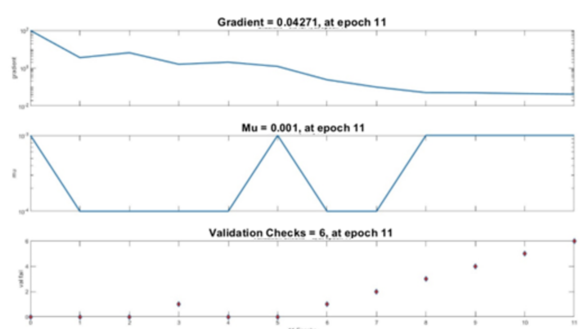
Texas Actual vs. Predicted Hail Events



Minnesota witnessed a total of 19,157 hail events, and the machine-learning process concluded after 11 epochs. Figure 16 showcases the training state of the machine learning process for Minnesota.

Figure 16

Minnesota ML Training State

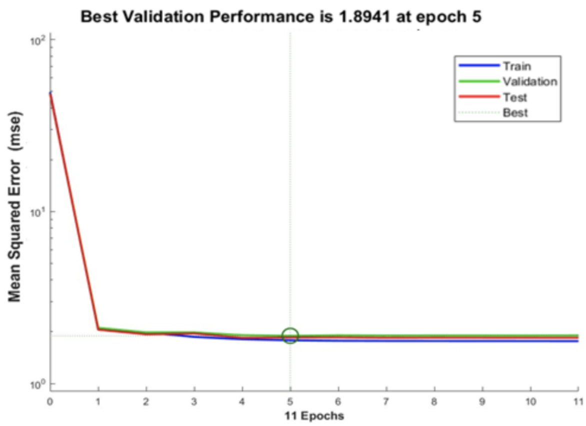


Note: The top graph, labeled 'Gradient', shows the gradient magnitude on the Y-axis, ranging from 10^{-2} to 10^2. The X-axis represents the epoch count. The middle graph, labeled 'Mu,' displays the mu parameter on the Y-axis, with a range from 10^{-4} to 10^{-3}. The X-axis corresponds to the epoch count as in the Gradient graph. The bottom graph, labeled 'Validation Checks,' indicates the count of validation checks on the Y-axis, which varies from 0 to 6. The X-axis, consistent with the above graphs, denotes the epoch count.

Figure 17 illustrates the validation performance of the machine learning process. Among all 11 epochs, the best performance was achieved at the 5th epoch.

Figure 17

Minnesota Hail Validation

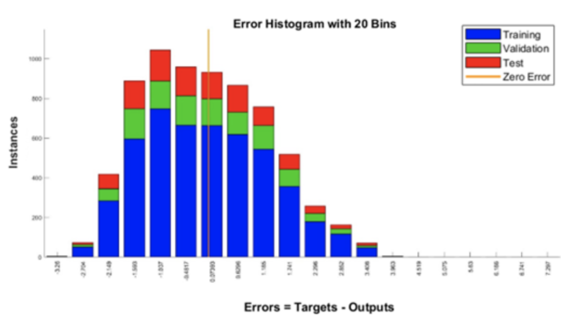


Note: The Y-axis represents the Mean Squared Error (MSE) on a logarithmic scale, ranging from 10̂0 to 10̂2. The X-axis denotes the epochs, which is the number of complete passes through the training dataset.

Figure 18 displays the error histogram of the machine learning process for Minnesota, with 20 bins representing different error ranges. The results show a right-skewed bell-shaped curve, indicating more errors compared to Kansas and Texas.

Figure 18

Minnesota Error Histogram

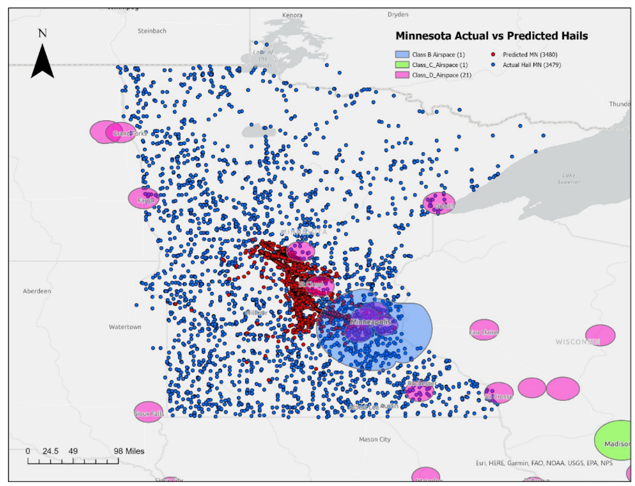


Note: The X-axis represents the range of error values, calculated as Targets minus Outputs, extending from -3.26 to 7.297. The Y-axis quantifies the number of instances (frequency) within each error bin, with a range from 0 to 1,000 instances.

Following the machine learning validation process, ArcGIS Pro visualized all the predicted hail events. Figure 19 compares the predicted (red dots) and actual (blue dots) hail events in Minnesota. The result indicates that most predicted hail events are concentrated in the middle of the state. The Root Mean Square Error (RMSE) represents the location difference. For Minnesota, the latitude RMSE is 1.37 (approximately 94 miles), and the longitude RMSE is 1.31 (approximately 57 miles).

Figure 19

Minnesota Actual vs. Predicted Hail Events



Discussion

Geospatial analysis revealed that the central United States is the epicenter of hail events due to the geographical conditions of the great plain area, promoting hail formation. In contrast, the west coast experienced fewer hail events. Southern states like Texas, and Kansas had higher frequencies of severe thunderstorms, resulting in more severe hail events, while northern states like Minnesota, located farther from the storm track, experienced fewer such events. Warm air from the Gulf of Mexico, coupled with drier air from the western United States, created an unstable atmosphere in the southern states, crucial for strong updrafts and severe hail formation. Additionally, the flatter terrain in the southern central United States allowed air masses to move more freely, supporting thunderstorm development and powerful updrafts conducive to hail formation.

The southern central United States experiences an earlier transition from winter to spring compared to the northern central United States, leading to an early onset of

DOI: 10.58940/2329-258X.2041

thunderstorms and hail formation. In contrast, the northern region takes longer to transition into spring, resulting in a later onset of severe weather events, including hail. Most states have fewer severe hail events in winter compared to spring and summer, with no severe hail events occurring in the selected states. Winter weather conditions tend to be more stable, suppressing the development of strong updrafts necessary for hail formation. Instead, winter precipitation typically takes the form of snow or freezing rain, associated with different atmospheric conditions and precipitation processes.

The neural network machine learning process used past hail events' latitude, longitude, and time to predict possible hail events in three selected states. Machine learning successfully provided predicted hail events concentrated in the middle of each state. Kansas had a longitudinal difference of about 102 miles and a latitudinal difference of about 50 miles. Texas had a longitudinal difference of about 135 miles and a latitudinal difference of about 130 miles, making it the state with the largest variance. Minnesota had a longitudinal difference of about 57 miles and a latitudinal difference of about 94 miles. The error histogram showed that both Kansas and Minnesota had more errors than Texas, which could be attributed to their state size and actual hail event distribution. Texas, despite its larger size and higher variance, had the lowest error in the machine learning process. This study's observed discrepancies between actual and predicted hail events highlight the complexity of severe weather phenomena. Future iterations of this research will extend the scope of investigation to include a wider range of atmospheric variables and leverage more sophisticated algorithms designed to capture complex meteorological interactions. The objective is to develop a model that not only accurately predicts hail occurrences but also reflects the multifaceted factors that contribute to such events.

The findings of this study have profound geospatial implications, offering enhanced visualization tools that empower practitioners with a deeper understanding of the spatial-temporal dynamics of hail event occurrences. Such visualizations facilitate a more nuanced comprehension of hail risks, enabling stakeholders in aviation and related sectors to anticipate and mitigate these events more effectively. For machine learning weather forecasting, the research provides a foundational approach to weather forecasting that encourages proactive planning for air operators. By demonstrating the potential of neural networks to predict hail events with a degree of accuracy based on limited variables, the study showed an idea for the development of more sophisticated models. These advancements are crucial for enhancing operational safety and efficiency, suggesting a future where air operators can rely on advanced predictive models for strategic decision-making, thereby minimizing the adverse impacts of severe weather on air operations.

Conclusion

The study analyzed the impact of hail events on the NAS from a geographical perspective, revealing that the central United States is the epicenter of severe hail events occurring mainly from March to August. The southern central United States experiences earlier hail events compared to the northern states. The first part of the study focused on airports, where these areas are highly sensitive to weather disruptions, which directly pertains to the operational facets of the National Airspace System (NAS). This localized approach facilitates a more granular examination of safety and efficiency impacts on flight operations. Future research will extend this analysis to encompass broader NAS implications, including economic and scheduling considerations that hail events precipitate. This progression will allow for a comprehensive understanding of hail's multifaceted impact on the entire NAS ecosystem. The study utilized a neural network as a machine learning tool to predict hail events, focusing on three states and using variables like time, latitude, and longitude for testing. The results indicated that most predicted hail events were concentrated in the center of the selected states. However, to enhance accuracy, additional variables such as temperature, humidity, wind shear, and pressure from other sources are recommended for forecasting. Future studies could explore additional states in the United States to further advance the prediction model.

References

Abhishek, K., Singh, M. P., Ghosh, S., & Anand, A. (2012). Weather forecasting model using artificial neural network. *Procedia Technology*, *4*, 311–318. https://doi.org/10.1016/j.protcy.2012.05.047

Baboo, S. S., & Shereef, I. K. (2010). An efficient weather forecasting system using artificial neural network. *International Journal of Environmental Science and Development*, *1*(4), 321–326. https://doi.org/10.7763/IJESD.2010.V1.63

Bureau of Transportation Statistics. (2022). *Understanding the reporting of causes of flight delays and cancellations*. U.S. Department of Transportation. https://www.bts.gov/topics/airlines-and-airports/understanding-reporting-causes-flight-delays-and-cancellations

Chen, P.-C., Dokic, T., Stokes, N., Goldberg, D. W., & Kezunovic, M. (2015). Predicting weather-associated impacts in outage management utilizing the gis framework. *2015 IEEE PES Innovative Smart Grid Technologies Latin America (ISGT LATAM)*. https://doi.org/10.1109/ISGT-LA.2015.7381191

Esri. (2023). *What is GIS? A spatial system that creates, manages, analyzes, and maps all types of data*. https://www.esri.com/en-us/what-is-gis/overview

Federal Aviation Administration. (2016). *Meteorology*. Skyhorse Publishing.

Federal Aviation Administration. (2022). *Chapter 15: Airspace*. U.S. Department of Transportation. https://www.faa.gov/regulationspolicies/handbooksmanuals/aviation/phak/chapter-15-airspace

Federal Aviation Administration. (2023a). *Airmen's meteorological information (AIRMET)*. In Pilot/Controller Glossary. U.S. Department of Transportation. https://www.faa.gov/air_traffic/publications/atpubs/pcg_html/glossary-a.html#$AIRMET

Federal Aviation Administration. (2023b). *Navigating around bad weather*. U.S. Department of Transportation. https://www.faa.gov/holiday-travel/weather

Gavin, H. P. (2024). *The levenberg-marquardt algorithm for non-linear least squares curve-fitting problems*. Duke University Department of Civil; Environmental Engineering. https://people.duke.edu/~hpgavin/ExperimentalSystems/lm.pdf

Joseph, V. R. (2022). Optimal ratio for data splitting. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *15*(4), 531–538. https://doi.org/10.1002/sam.11583

Kulesa, G. (2010). *Weather and aviation: How does weather affect the safety and operations of airports and aviation, and how does faa work to manage weather-related effects?* U.S. Department of Transportation. https://www.transportation.gov/sites/dot.gov/files/docs/kulesa_Weather_Aviation.pdf

Luo, G., Zhang, F., Cao, W., Zhang, H., Liu, L., & Chen, W. (2022). Numerical and experimental verification of spherically layered ice model for analysis on hail impacting aviation structures. *International Journal of Aerospace Engineering*, 1–21. https://doi.org/10.1155/2022/1806623

National Oceanic and Atmospheric Administration. (n.d.-a). *Aviation weather data*. Aviation Weather Center. https://aviationweather.gov/help/data/#metar

National Oceanic and Atmospheric Administration. (n.d.-b). *Hail*. NOAA National Severe Storms Laboratory. https://www.nssl.noaa.gov/research/hail/

National Oceanic and Atmospheric Administration. (n.d.-c). *Nssl research: Hail*. NOAA National Severe Storms Laboratory. https://www.nssl.noaa.gov/research/hail/

National Oceanic and Atmospheric Administration. (2013). *Aviation weather center - terminal aerodrome forecasts (tafs)*. Aviation Weather Center. https://aviationweather.gov/data/taf/

National Oceanic and Atmospheric Administration. (2023). *Storm prediction center*. NOAA/National Weather Service. https://www.spc.noaa.gov/

Nguyen, Q. H., Ly, H.-B., Ho, L. S., Al-Ansari, N., Le, H. V., Tran, V. Q., Prakash, I., & Pham, B. T. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. *Mathematical Problems in Engineering*, 1–15. https://doi.org/10.1155/2021/4832864

Scher, S., & Messori, G. (2018). Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, *144*(717), 2830–2841. https://doi.org/10.1002/qj.3410

Waring, S., Zakos-Feliberti, A., Wood, R., Stone, M., Padgett, P., & Arafat, R. (2005). The utility of geographic information systems (gis) in rapid epidemiological assessments following weather-related disasters: Methodological issues based on the tropical storm allison experience. *International Journal of Hygiene and Environmental Health*, *208*(1-2), 109–116. https://doi.org/10.1016/j.ijheh.2005.01.020