Thesis Topic Discussion

Hybrid Meteorological Forecasting: ML-driven Predictions of Lake Michigan's Lake-Effect Snow on Urban Preparedness.

Noctis Yamazaki Zhang

Master of Science in Computer Engineering Systems

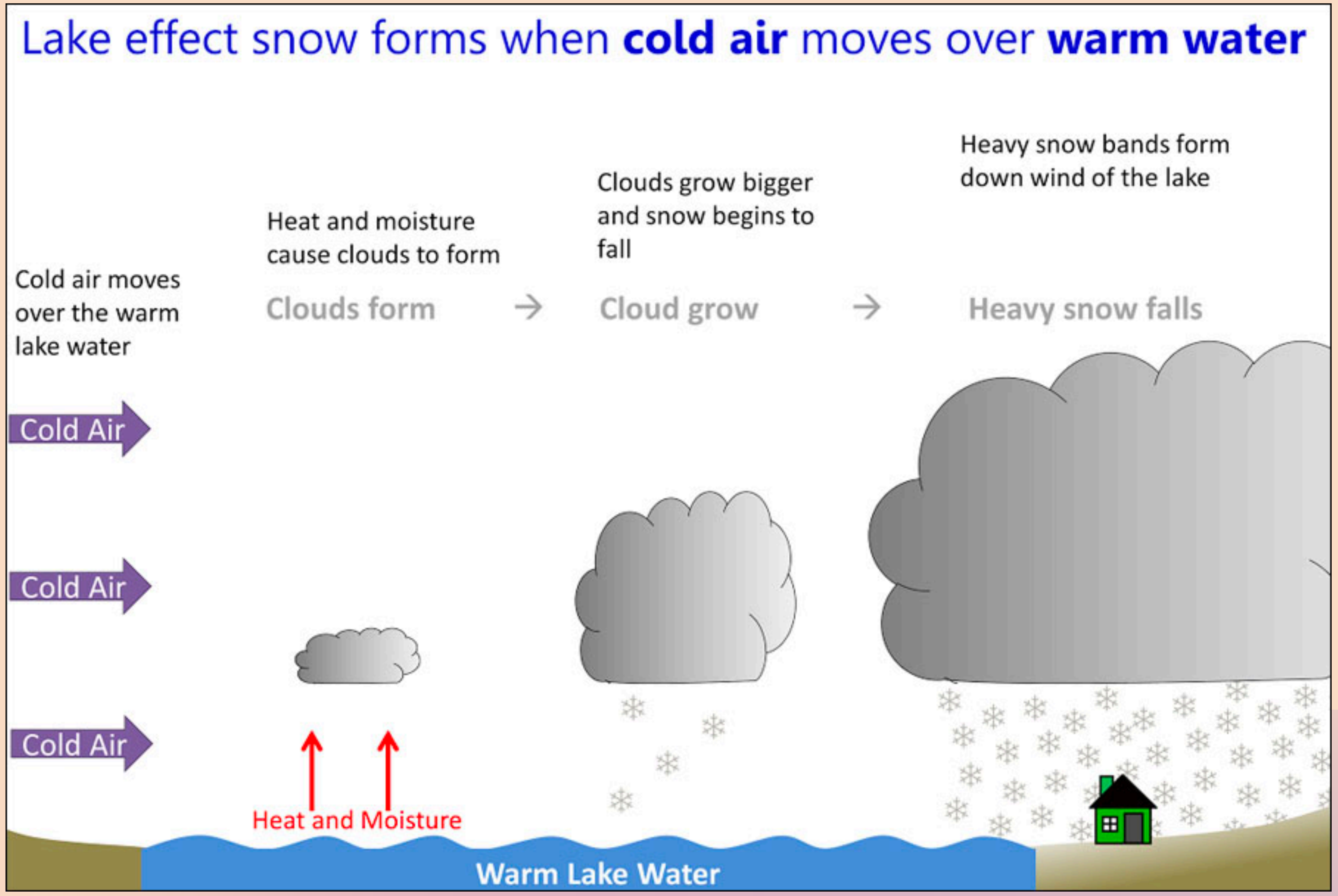
Prof. Dino Konstantopoulos as Advisor

Overview

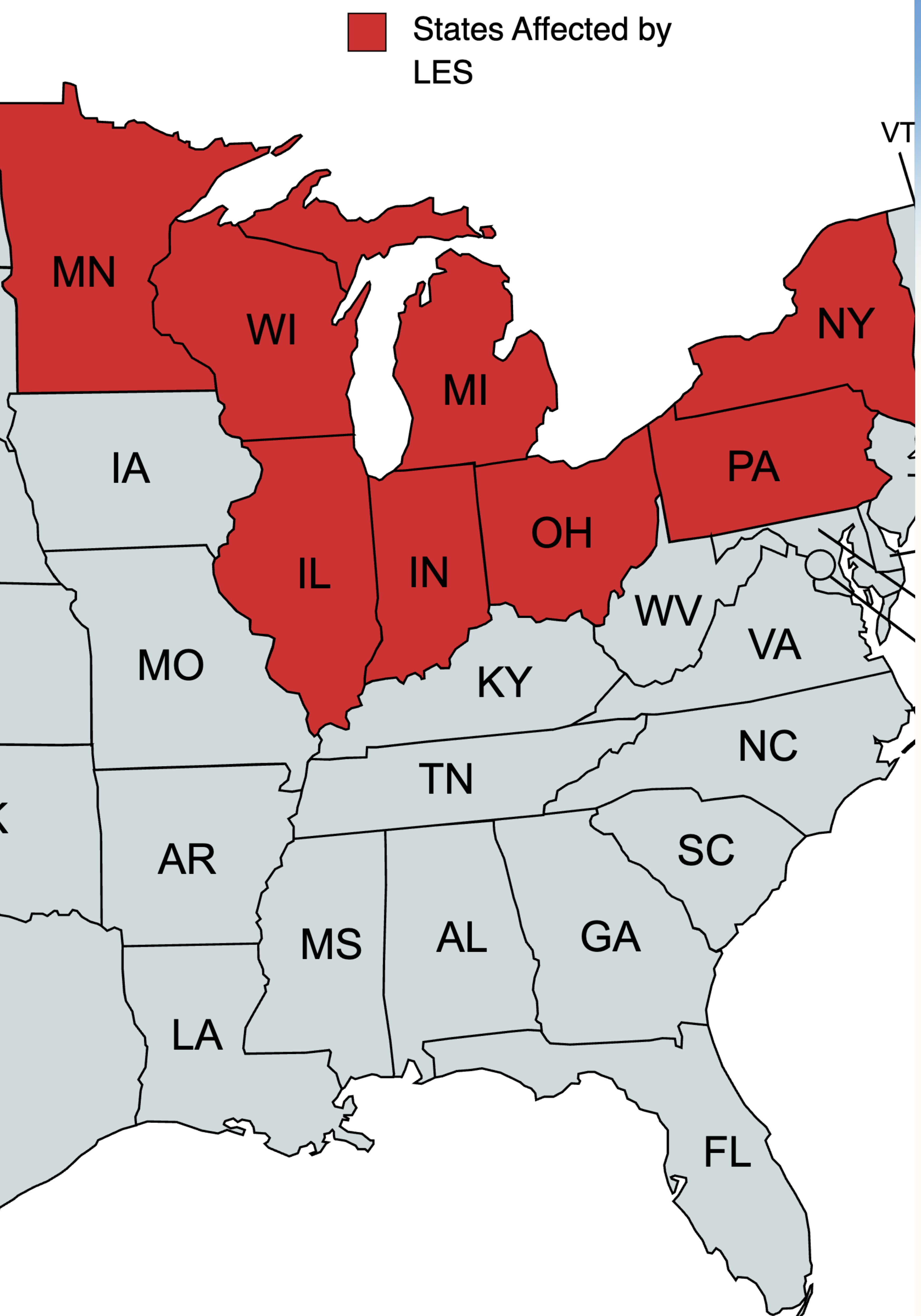
- Introduction
- Statement of the Problem
- Hypothesis
- Review of Related Literature
- Dataset
- References

What is Lake-Effect Snow

- Lake Effect Snow in the Great Lakes Region:
  - Common during late fall and winter.
  - Triggered by cold air moving across the open waters of the Great Lakes.
  - Origin often from Canada.
- Formation Process:
  - Cold air gathers warmth and moisture from the warm, unfrozen lake waters.
  - Leads to the rise of air, forming clouds.
- Forecasting Lake Effect Snow:
  - National Weather Service meteorologists consider various factors.
  - Includes wind direction, geographical factors, and more.



- Source: US Department of Commerce, N. (2018, March 23). *What is a lake effect snow?*. National Weather Service. https://www.weather.gov/safety/winter-lake-effect-snow#:~:text=Lake%20Effect%20snow%20occurs%20when,lowest%20portion%20of%20the%20atmosphere.



Impact of Lake-Effect Snow

- States Affected: 8
- Population Impacted: 85.8M
- Percentage of U.S. Total: 25.9%

Past Experience

Proposed to evaluate and advance the representation of lake-atmosphere interactions and heavy LES in the Great Lake Basin.

- Explored feasibility of expanding 1D lake models with Radar & Satellite Imagery data

- Designed image classifier for Lake-Effect Snow satellite images with simple ML method

Funded by NASA, $1,000,000, 2017-12/31/2018, *Heavy Lake-Effect Snowstorms in NU-WRF.*

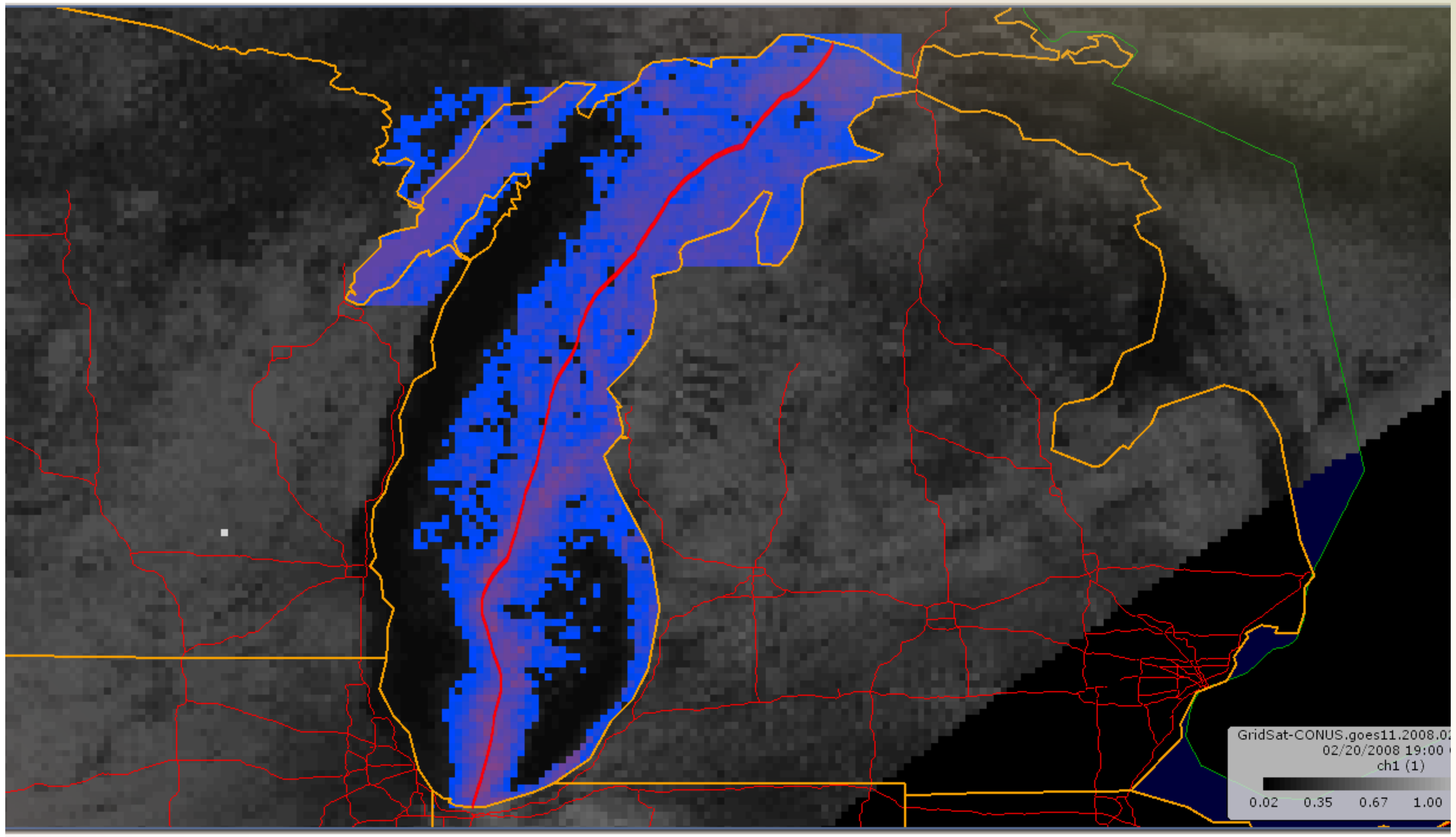
Previous Achievement

32% Improvement

Original Accuracy: 57.0%
Improved Accuracy: 75.2%

We witnessed a significant improvement from the 1D lake NWP model Flake proposed by G. Rooney and I.D. Jones in 2009, by adding explanatory variables based on satellite imagery data to predict precipitation near shore in 24 hours.

The identified cloud for image # 1

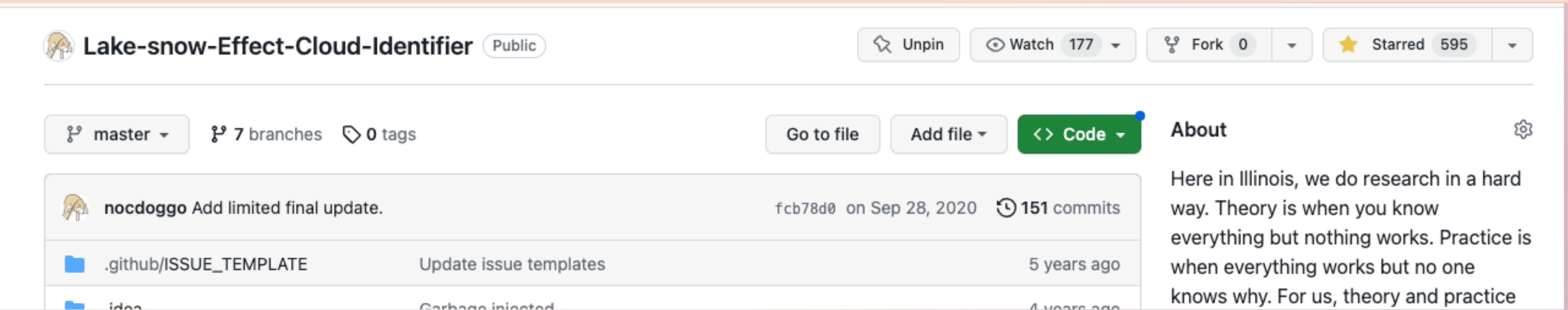


FLake is a one-dimensional (1D) freshwater lake model, focusing on vertical variability from the lake's surface to its bottom, while assuming horizontal consistency, with properties like temperature and currents only varying by depth.

Inspirations

- Utilize Deep Learning and Graphic Processing Units to craft forecast model

- Incorporating spatial-temporal data in machine learning can greatly improve traditional forecasting models



Statement of the Problem

There is a pressing need for alternative predictive methodologies that can offer both short-term accuracy and long-term forecasting capabilities.

Objectives

- Harness Hybrid Predictive Meteorology
- Integrate Multiple Data Sources
- Enhance Forecasting Capabilities

Hypothesis

1 Combine satellite and meteorological data with varied sample durations into a single model

Issue:
- Satellite data: 8 hours/day, 15-min intervals, nighttime data absent
- Meteorological data: Hourly samples

Goal:
- Develop neural network to process and train both datasets concurrently

2 Attain precise mid-range (1-3 days) precipitation forecasting

Issue:
- NWP and ML models limited to short-range accuracy (up to 24 hours)
- Lack of focus on specific climatic events

Goal:
- Utilize multi-day climatic datasets for forecasting
- Predict events at 24, 48, and 72-hour intervals

3 Create a model that can be applied to other geolocations post-training

Issue:
- Traditional models are location-specific
- Lack adaptability for different geolocations post-training
- Requires separate training for each new location, hindering scalability

Goal:
- Assess a forecasting model tailored for cross-geolocation adaptability
- Utilize local datasets to predict similar climatic events across varied locations

Review of Related Literature

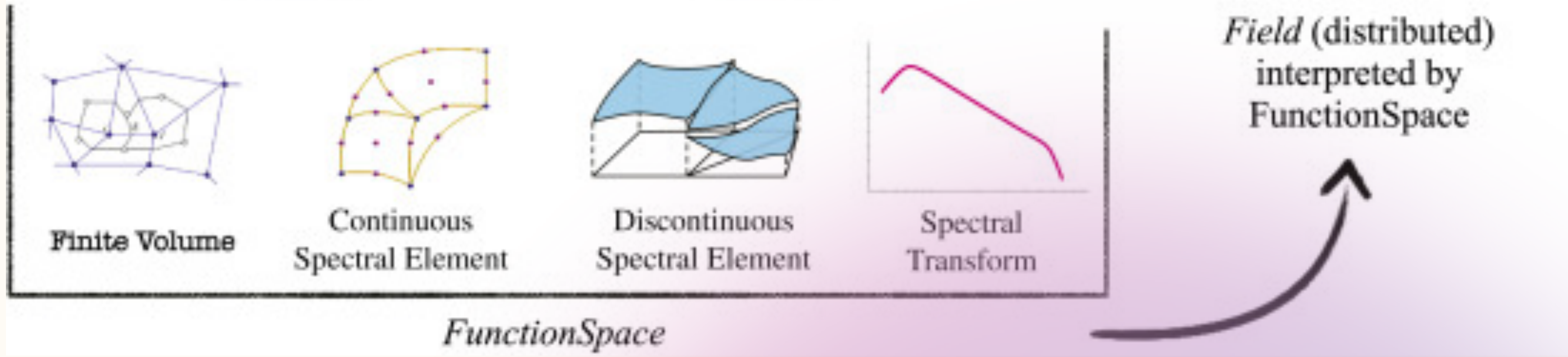
State-of-the-Art Practices in Meteorology

- Numerical Weather Prediction (NWP) Models
- Recent Machine Learning Methods

Foundations of NWP Models

Atmospheric Equations (Key: Navier-Stokes Equations)

- Continuity Equation
  - Principle: Mass conservation
  - Mass neither created nor annihilated
- Momentum Equations
  - Principle: Momentum conservation
  - Across three spatial dimensions
- Thermodynamic Energy Equation
  - Principle: Energy conservation
  - Energy transitions but remains conserved
- Water Substance Equations
  - Principle: Conservation of water phases
  - Vapor, liquid, and solid transitions



Challenges of NWP Models

In the meteorological domain, the efficacy of NWP models is widely recognized, yet their forecast accuracy, rooted in complex mathematics and physics, can be influenced by numerous factors.

- Forecast Horizon:
  - Short-term: 1-3 days with ~75% accuracy.
  - Medium-range: 3-10 days averaging ~60% accuracy.
  - Decline over time due to atmospheric chaos.

- Weather Parameter:
  - Variable accuracy across different climatic parameters.
  - Temperature predictions may be more reliable than precipitation.

- Geographical Region:
  - Complex terrains (mountains, coastlines) present greater challenges.
  - Flatter landscapes offer fewer prediction complexities.

- Seasonality:
  - Models' accuracy varies with seasons.
  - Atmosphere's predictability is not constant; influenced by specific atmospheric configurations during certain seasons.

Foundation of Machine-Learning Methods

Advanced sophisticated deep learning approaches in 2010s.

Techniques from Computer Vision, like deep and reinforcement learning, offer promising forecasting advancements.

- CNNs (Convolutional Neural Networks)
  - Ideal for satellite and radar images.
  - Excell at identifying patterns, e.g., lake-effect snow signs.
- RNNs (Recurrent Neural Networks) & LSTMs
  - Capture temporal sequences, predict events like snowfall evolution.
- Rich data sources:
  - Satellite imagery: cloud shapes, sea surface temps.
  - Terrestrial stations: temperature, humidity, wind.
  - Radar: detailed precipitation profiles.
  - Weather balloons: stratified atmospheric insights.

Research Trajectories in Deep Learning

- Remote Sensing
    - E.g., TITAN and NEXRAD - focus on radar-centric storm detection and tracking.
    - High-resolution radar data boosts precision.
    - Short-term predictive radar imaging offers cloud drift forecasts.
- TITAN: excels in real-time storm cell detection and tracking.
- Prediction accuracy:
    - High for immediate forecasts (1-2 hours).
    - Decreases over extended periods.
- Rapid Refresh model: reaffirms deep learning's potential in meteorology.

Challenges of ML Models

- Meteorological Data & Deep Learning
  - Vast data, but focus remains on short-term predictions.
  - Many models utilize just a fraction of available data.

- Nowcasting's Precision & Limits
  - Strength: Imminent weather changes, especially storm trajectories.
  - Challenge: Diminished accuracy beyond 1-hour forecasts.

- Adaptability Gap
  - Few models use "transfer learning" for regional adaptability.
  - Potential to reduce costs and fast-track deployment.

- Need for Specialization
  - General forecasts prevalent, but specific tools scarce.
  - Example: Limited tools for events like lake-effect snows.

Dataset

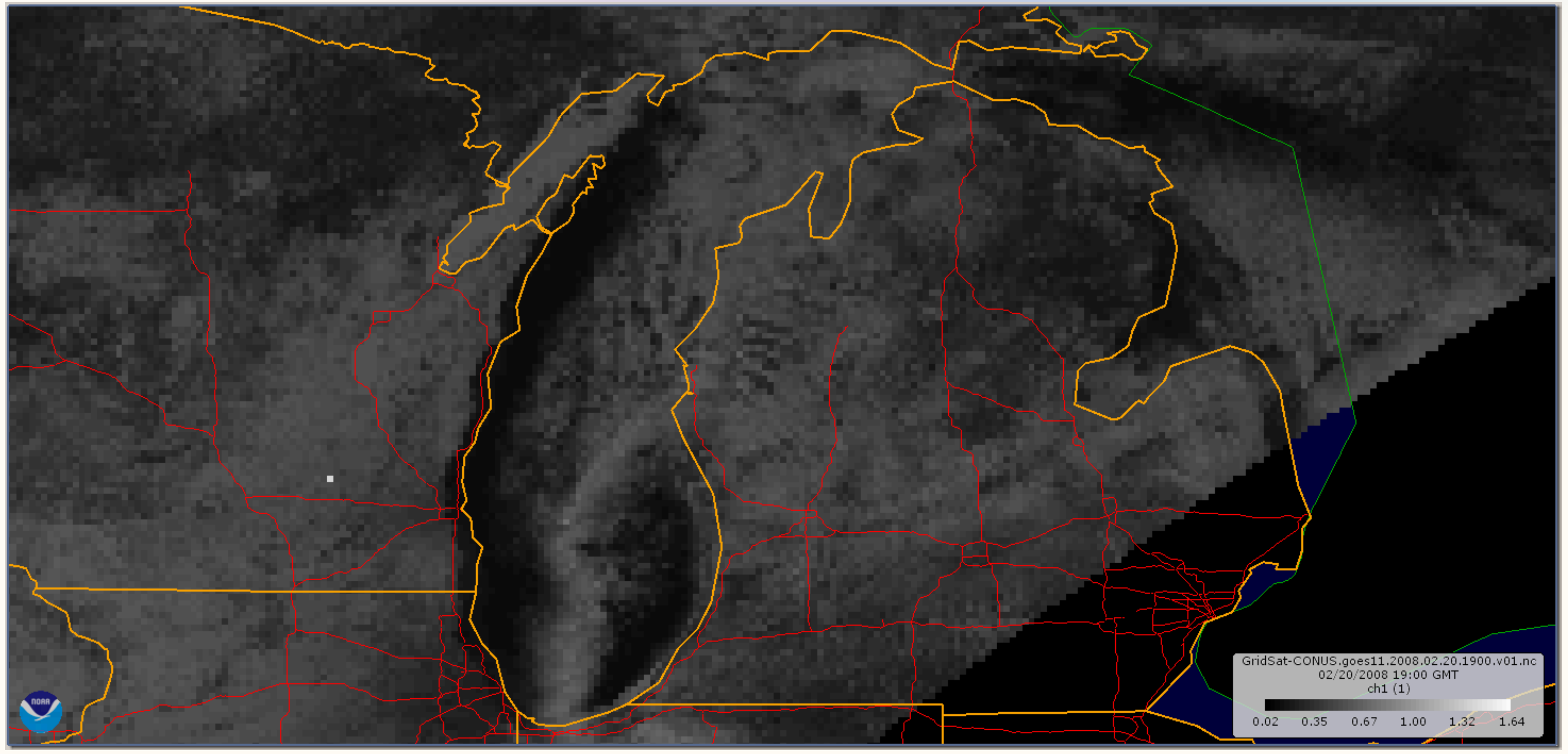
Understanding Lake-Effect Snow

- Essential to grasp meteorological conditions around lake regions.
- Research relies on satellite imagery of the Great Lakes and meteorological data up to 150 miles inland.
- Peak snowfall usually within 10-40 miles from shoreline, influenced by lake size and wind factors.
- Lakes like Superior and Michigan influence areas over 100 miles inland.

Key Tools in Research

- Geostationary Operational Environmental Satellite (GOES) for HD imagery.
- Data from airports, weather stations, and trained meteorological spotters.

GOES Imagery Data



Geostationary Operational Environmental Satellite (GOES)

- Initiative by NOAA
- Provides a detailed look at continental climates
- A decade of precise atmospheric data available for public use

NOAA GOES Weather Toolkit

- More than just weather data: a comprehensive suite
- Offers rich background maps, animations, and precise spectral band filtering
- Enables in-depth research and actionable insights

GOES Imagery Data

Data Segregation & Analysis: Lake Michigan

- Focus: Lake Michigan within the greater Great Lakes region.
- Time Frame: Oct 1, 2006 - March 31, 2007 (11 years).
- Data Type: Quarter-hourly cloud density indices. Method: Utilized longitude & latitude values for precise regional differentiation.
- Insight: Indices (0-1) reveal cloud coverage, critical for lake-effect snow study.

Data Transformation: Structuring Cloud Density Data

- Goal: Convert Lake Michigan data into a 2D array for analysis.
- Format: Rows: Latitude values (ascending order). Columns: Longitude values (ascending order).
- Result: Each array element indicates cloud density (0 for no cloud, near 1 for thick clouds).
- Size-Reduction: Results in 48 x 48 grayscale images with use of OpenCV

National Weather Service (NWS) Meteorological Data

NOWData

- Sourced from federal weather stations within the National Weather Service network.
- Hourly data collection spanning a wide range of parameters.
- Each parameter provides insights into the complexities driving lake-effect snow genesis and dynamics.

Key Meteorological Parameters

- Temperature (°F): Ambient conditions.
- Relative Humidity (%): Atmospheric moisture content.
- Dew Point (°F): Air saturation temperature.
- Wind Speed & Gust (mph): Force and surges of wind.
- Wind Direction: Orientation of winds (measured in 10° intervals).
- Cloud Heights (ft): Vertical cloud distribution.
- Visibility (miles): Atmospheric clarity.
- Pressures: Atmospheric, Altimeter, Sea Level (all in hPa).
- Precipitation (inches): Moisture deposition (rain, sleet, snow).

References

Abi bands quick information guides. ABI Bands Quick Information Guides. (n.d.). https://www.goes-r.gov/mission/ABI-bands-quick-info.html

Al-Yahyai, S., Charabi, Y., & Gastli, A. (2010). Review of the use of Numerical Weather Prediction (NWP) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, *14*(9), 3192–3198. https://doi.org/10.1016/j.rser.2010.07.001

Ansari, S. (n.d.). *Introduction*. NOAA's Weather and Climate Toolkit (Viewer and Data Exporter). https://www.ncdc.noaa.gov/wct/

Ayzel, G., Scheffer, T., & Heistermann, M. (2020). RainNet v1.0: A Convolutional Neural Network for radar-based precipitation nowcasting. *Geoscientific Model Development*, *13*(6), 2631–2644. https://doi.org/10.5194/gmd-13-2631-2020

Bauer, P., Thorpe, A., & Brunet, G. (2015). The Quiet Revolution of Numerical Weather Prediction. *Nature*, *525*(7567), 47–55. https://doi.org/10.1038/nature14956

Bulletin. (2017, November 27). *Innovations and new technology for improved weather services*. World Meteorological Organization. https://public.wmo.int/en/bulletin/innovations-and-new-technology-improved-weather-services

Data types. MRCC. (n.d.). https://mrcc.purdue.edu/data_serv/dataTypes

Dixon, M., & Wiener, G. (1993). Titan: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology. *Journal of Atmospheric and Oceanic Technology*, *10*(6), 785–797. https://doi.org/10.1175/1520-0426(1993)010<0785:ttitaa>2.0.co;2

Feistel, R., Naush, G., & Wasmund, N. (2008). *State and evolution of the Baltic Sea, 1952-2005: A detailed 50-year survey of Meteorology and climate, physics, Chemistry, Biology, and Marine Environment*. Wiley-Interscience.

Geer, A. J., Baordo, F., Bormann, N., Chambon, P., English, S. J., Kazumori, M., Lawrence, H., Lean, P., Lonitz, K., & Lupu, C. (2017). The growing impact of satellite observations sensitive to humidity, cloud and precipitation. *Quarterly Journal of the Royal Meteorological Society*, *143*(709), 3189–3206. https://doi.org/10.1002/qj.3172

References

Geostationary satellites. National Environmental Satellite, Data, and Information Service. (n.d.). https://www.nesdis.noaa.gov/current-satellite-missions/currently-flying/geostationary-satellite

Ide, K. (2010). *Multi-Model Ensemble Approaches to Data Assimilation Using the 4D-Local Ensemble Transform Kalman Filter*. https://doi.org/10.21236/ada542670

Johannsen, F., Dutra, E., & Magnusson, L. (2020). *Evaluation of Systematic Biases and Skill of Summer Subseasonal Forecasts in the ECMWF System*. https://doi.org/10.5194/egusphere-egu2020-21676

Kim, Y., & Hong, S. (2019). Deep learning-generated nighttime reflectance and daytime radiance of the Midwave Infrared Band of a geostationary satellite. *Remote Sensing*, *11*(22), 2713. https://doi.org/10.3390/rs11222713

Meyer, V. K., Höller, H., & Betz, H. D. (2013). The temporal evolution of three-dimensional lightning parameters and their suitability for thunderstorm tracking and nowcasting. *Atmospheric Chemistry and Physics*, *13*(10), 5151–5161. https://doi.org/10.5194/acp-13-5151-2013

Milani, L., Kulie, M. S., Casella, D., Kirstetter, P. E., Panegrossi, G., Petkovic, V., Ringerud, S. E., Rysman, J.-F., Sanò, P., Wang, N.-Y., You, Y., & Skofronick-Jackson, G. (2021). Extreme lake-effect snow from a GPM microwave imager perspective: Observational analysis and precipitation retrieval evaluation. *Journal of Atmospheric and Oceanic Technology*, *38*(2), 293–311. https://doi.org/10.1175/jtech-d-20-0064.1

Nostrand, V. (2006). Geostationary Operational Environmental Satellite (GOES). *Van Nostrand's Scientific Encyclopedia*. https://doi.org/10.1002/0471743984.vse8611

Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., Mozaffari, A., & Stadtler, S. (2021). Can deep learning beat numerical weather prediction? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *379*(2194), 20200097. https://doi.org/10.1098/rsta.2020.0097

Scott, R. W., & Huff, F. A. (1996). Impacts of the Great Lakes on regional climate conditions. *Journal of Great Lakes Research*, *22*(4), 845–863. https://doi.org/10.1016/s0380-1330(96)71006-7

Song, L., Schicker, I., Papazek, P., Kann, A., Bica, B., Wang, Y., & Chen, M. (2020). Machine learning approach to summer precipitation nowcasting over the eastern alps. *Meteorologische Zeitschrift*, *29*(4), 289–305. https://doi.org/10.1127/metz/2019/0977

Vandal, T., Kodra, E., & Ganguly, A. R. (2018). Intercomparison of Machine Learning Methods for statistical downscaling: The case of daily and extreme precipitation. *Theoretical and Applied Climatology*, *137*(1–2), 557–570. https://doi.org/10.1007/s00704-018-2613-3

Wiley, J., & Mercer, A. (2020). An updated synoptic climatology of Lake Erie and Lake Ontario Heavy Lake-effect snow events. *Atmosphere*, *11*(8), 872. https://doi.org/10.3390/atmos11080872

Xu, Z., Du, J., Wang, J., Jiang, C., & Ren, Y. (2019). Satellite image prediction relying on gan and LSTM Neural Networks. *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. https://doi.org/10.1109/icc.2019.8761462

Thank You for your Time and Interests!

Noctis Yamazaki Zhang