THEME ARTICLE: LARGE SCALE SCIENCE ON NSF'S FRONTERA SYSTEM

Modeling the World's Most Violent Thunderstorms

Leigh Orf, *University of Wisconsin, Madison, WI, 53706, USA*

Each year tornadoes are responsible for loss of life and property as they wreak destruction across large swaths of the United States. The storms that cause the most damaging tornadoes, supercell thunderstorms, have been the subject of numerical modeling since the dawn of the supercomputing age in the late 1970s. In 2020, models run on supercomputers can simulate individual thunderstorms with unprecedented realism where well-resolved multiple-vortex tornadoes form as the result of physical processes occurring within the model. Research conducted on the Frontera supercomputer includes identifying specific processes occurring in supercells that result in the most devastating tornadoes, those ranked EF4/5 on the Enhanced Fujita scale. While such tornadoes account for a small fraction of observed tornadoes, they cause the bulk of the damage and fatalities, and only by increasing our knowledge of what discerns these rare but powerful storms from less damaging storms will better forecasts be possible.

Each spring across broad regions of the central and southeastern United States, atmospheric conditions occasionally become favorable for the formation of long-lived, rotating thunderstorms called supercells. Supercell thunderstorms are the most prolific creators of long-lived violent tornadoes; in the unusually active year of 2011, tornadoes in the United States killed 553 people and caused $28 billion dollars in property damage. Cities such as Joplin, MO; Tuscaloosa, AL; and El Reno, OK experienced tornadoes rated EF4 or EF5, the top two categories of the Enhanced Fujita scale that is used to estimate tornado wind speeds based upon resultant damage. These extreme tornadoes were spawned from supercells, and a main focus of my team's research conducted on National Science Foundation (NSF) supercomputers is understanding the specific processes that occur in the small fraction of observed supercells that produce the most damaging tornadoes.

1521-9615 © 2021 IEEE
Digital Object Identifier 10.1109/MCSE.2021.3069771
Date of publication 30 March 2021; date of current version 15 June 2021.

NUMERICAL WEATHER PREDICTION AND SEVERE WEATHER FORECASTING

The US Storm Prediction Center (SPC) and National Weather Service (NWS) are tasked with forecasting severe weather, on both long and short time scales. Globe-spanning numerical weather prediction (NWP) models in conjunction with regional models focused on the United States ingest observational data from a whole host of sources including weather satellites, weather radar, commercial aircraft, and surface weather stations to provide forecasts ranging from several days to several hours into the future. These models serve as guidance for forecasters who issue convective outlooks (that include the location, timing, and magnitude of anticipated severe weather outbreaks) from one to several days in advance. These convective outlook forecasts have improved dramatically over the past decades; it is now common for outlooks days in advance to forecast accurately, albeit over a broad region, when and where a severe weather outbreak will occur and the probabilities of certain hazard types (such as hail, straight-line winds, and tornadoes) occurring within a certain area.

Forecasting Improvements Driven by Research and Technology

Our ability to forecast these outbreaks more accurately is largely due to improvements in NWP model forecasts, which are the result of advancements on several fronts. Models provide more accurate forecasts due to improvements in model physics and numerical techniques, and have been adapted to modern parallel supercomputing architectures, enabling simulations to be run over larger areas and at higher resolution. Ensemble forecasting, an approach that involves running many slightly different simulations of the same event, provides probabilistic weather forecasts that are not otherwise possible only utilizing single deterministic simulations. There are more observational sources of data available now than ever, including modern weather satellites viewing the world through dozens of spectral windows, and these additional observations provide more accurate initial conditions for models. Data assimilation, the computationally intensive process by which the myriad of different observational sources are blended to initialize the forecast models, has benefited from improved algorithms as well as advances in supercomputing hardware.

Many Remaining Forecasting Challenges

Despite these and other advances, such as the increasing use of unmanned aerial vehicles (UAVs), such as drones and quadcopters to take atmospheric measurements, accurate short-term forecasting (on the order of tens of minutes to several hours in advance) of the formation and paths of individual observed storms, including the accurate paths of any tornadoes, is something that remains elusive in 2020. The resolution of operational prediction systems (~3 km) is too low to resolve some hazards such as tornadoes; however, these models can explicitly predict other hazards such as high winds and hail, although accurately predicting the timing and location of these hazards remains a challenge. Even if hardware allowed it, current forecasting techniques are limited by the lack of widespread accurate fine-scale observational data to serve as the initial conditions for tornado-resolving operational forecast models. Lacking these accurate initial conditions, simulations at these resolutions would not be able to properly predict the timing and location of individual tornadoes. These simulations would likely still be useful to forecasters, however, as they would provide examples of the overall convective mode that characterizes the types of storms that are likely to occur, information that could be exploited by operational forecasters.

Even if tornado-resolving (on the order of 50 m horizontal grid spacing) regional forecast models spanning the United States could be initialized ideally today, modern supercomputing hardware would not be able to run nearly fast enough to provide timely forecasts. While operational models of the future will undoubtedly be run at a higher resolution and use improved algorithms exploiting artificial intelligence and machine learning technology, the accuracy of these models will always be limited by the availability of accurate observational data to initialize them. Remote sensing technology that can sample large volumes of the atmosphere quickly, such as radar and satellite, are very costly and physically constrained in their sampling abilities, practically limiting the availability of widespread, timely atmospheric data.

IF ALL FORECASTS PROVIDE A SIMILAR SOLUTION, IT PROVIDES HIGH CONFIDENCE IN THAT SOLUTION

Ensemble Forecasting

Ensemble forecasting is a powerful tool that is used to partly overcome our inability to run accurate forecasts at very high resolution. Ensembles are sets of simulations that all forecast over the same region and time, but are initiated with slightly different atmospheric conditions, or run with different numerics or physics parameterization options or even different models altogether. Because NWP models contain many available physical parameterization options as well as tunable parameters that have no one "correct" setting, it is valuable to run ensembles of simulations to see what kind of forecast spread occurs between the simulations. If all forecasts provide a similar solution, it provides high confidence in that solution; if forecasts differ, how they differ between simulations is useful and can be described in probabilistic language, where the amount of certainty of different outcomes is provided in the forecast.

Ensembles are not just useful for making better probabilistic forecasts; carefully constructed ensembles of idealized thunderstorm simulations can provide a "spectrum of possibilities" for storm outcomes

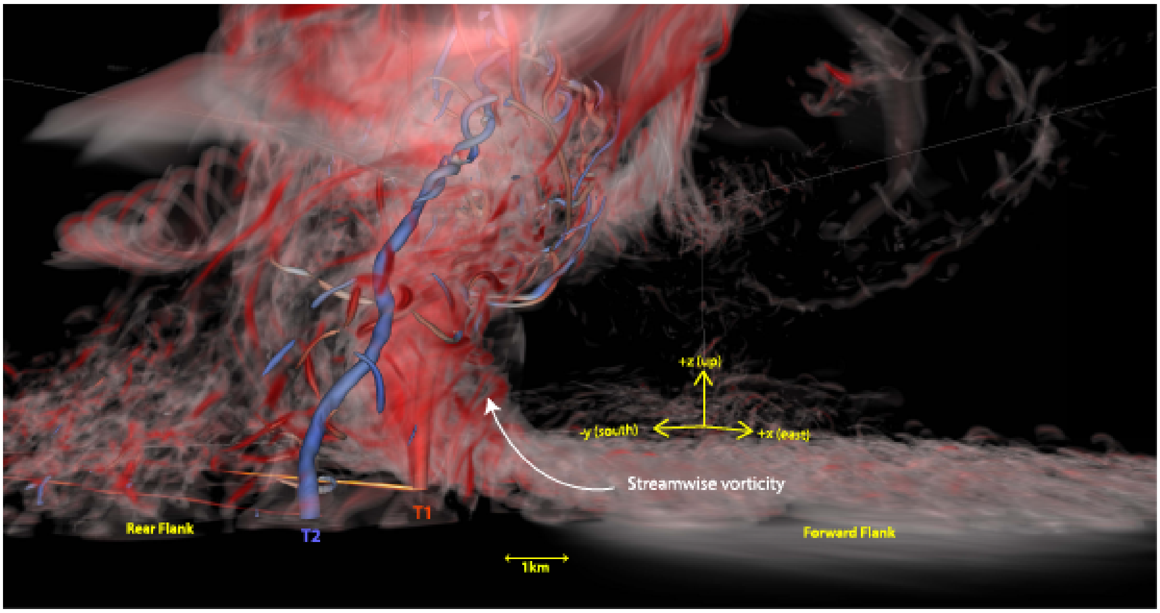

FIGURE 1. Tornado supercell simulation of 30 m resolution conducted on Frontera. The volume rendered field is positive streamwise vorticity (red indicates larger values), and the 0.6 s^{-1} vorticity magnitude isosurfaces are shown, indicating two tornadoes. Tornado paths are indicated by maximum surface wind speed tracks. View is looking toward the northwest. Two tornadoes are annotated: T1, the stronger of the pair, is rotating cyclonically, while T2 is rotating anticyclonically. The curved white arrow indicates the motion of a feature we named the SVC, a "tube" of horizontally oriented streamwise vorticity near the ground that it is abruptly tilted upwards into the storm's updraft. The SVC, thought to strengthen the storm's low-level updraft, making it more conducive to tornado formation and maintenance, has been found to form roughly parallel to buoyancy gradients along the storm's forward flank cold pool boundary, where thermodynamic processes such as evaporation and melting result in negatively buoyant air.

in a given environment in basic research on supercells and tornadoes. Ensembles can help answer questions about the intrinsic predictability of tornadoes; given ideal initial conditions, how do tiny changes in the simulation parameters, initial forcing, or background environment affect the simulation, such as whether a significant tornado forms? Recent research in this area suggests that accurately predicting the timing, duration, and intensity of individual tornadoes in storms that are forming or just about to form may not be feasible.[1] Ensemble forecasting can, however, provide a statistical picture of the range and nature of possibilities.

MODERN RESEARCH ADVANCING KNOWLEDGE OF SUPERCELLS AND TORNADOES

Field Study

While improved forecasting is always an end goal of severe weather research, there is active research dedicated toward answering fundamental questions on the internal workings of thunderstorms, including how they produce and sustain tornadoes, and why storms in similar environments can result if very different outcomes (such as no tornado versus strong tornado). Research field projects dedicated to understanding thunderstorms use both *in situ* and remote sensing technology such as portable Doppler radars to sample storms throughout their life cycles. A primary goal of many of these studies is to capture features associated with tornado formation in severe thunderstorms such as supercells.

Recent field projects such as the NSF-sponsored Targeted Observations by Radars and UAS of Supercells (TORUS) include the use of unmanned aerial systems (UAS) to provide *in situ* sampling of thermodynamic data (temperature, pressure, humidity) in targeted regions of supercells thought to be associated with tornadoes.[2] A feature my research team dubbed the streamwise vorticity current (SVC; see Figures 1 and 2), discovered first in simulations of supercell thunderstorms conducted on the NSF-sponsored Blue Waters supercomputer,[3] is one of the targets of the TORUS field study, and SVCs have recently been identified in the field (e.g., the paper by Schueth *et al.*[4]). The discovery of the SVC in simulation

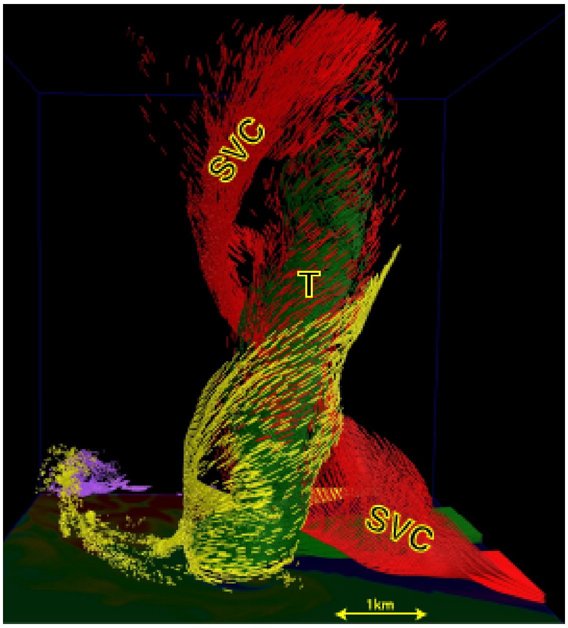

FIGURE 2. Trajectories released every 2 s from four different regions near the ground during the maintenance phase of the tornado for a 30 m simulation. View is looking toward the west. The red trajectories, originating near the ground along the leading edge of the cold pool in the storm's forward flank, trace the SVC. Dark green trajectories originate deep within the storm's forward flank cold pool where they are abruptly lifted into the outer circulation of the tornado. Yellow and blue trajectories originate further rearward in the storm's cold pool.

data that prompted scientists to seek it, and ultimately find it, in nature, is a strong testament to the power of HPC technology to further our understanding of the physical world.

Simulating Violent Supercells, Past and Present

The aforementioned operational NWP models used by the SPC and NWS run on dedicated hardware and are tuned specifically to provide timely forecasts, providing results sufficiently in advance to be of use to forecasters. Basic research utilizing numerical models does not suffer from this constraint, and researchers are free to run computationally expensive models at higher resolution even if it means a single simulation may take days or even weeks to complete. In fact, it has only been relatively recently that computing technology and numerical model sophistication advanced to the point where idealized simulations of supercells conducted on research supercomputers produced tornadoes that were sufficiently resolved.

Since the 1970s, just at the dawn of the supercomputer era, computers such as the CDC-7600 and Cray 1 were utilized by atmospheric scientists who could, for the first time, run simulations in three dimensions (3-D), albeit at very coarse resolution. Huge strides in scientists' understanding of the fundamental properties of supercell thunderstorms came as a result in the 1970s and 1980s as researchers were able to conduct 3-D simulations using the state-of-the-art cloud models of the time. Even though the models of the time-contained crude parameterizations for cloud microphysics and were run at coarse resolution over a small domain, much of today's basic understanding of supercells, such as how and why the updraft rotates at different levels and the locations of downdrafts, came from these early simulations.

OVER TIME, AS COMPUTER HARDWARE BECAME MORE POWERFUL, MODELS BECAME MORE SOPHISTICATED IN ORDER TO EXPLOIT THIS POWER

Over time, as computer hardware became more powerful, models became more sophisticated in order to exploit this power, utilizing higher order accurate numerical techniques and more sophisticated physical parameterizations. The additional power provided by advances in supercomputing hardware also enabled researchers to run models at higher resolution, enabling smaller and smaller features to be resolved. For a historical review of thunderstorm modeling research spanning, the dawn of 3-D cloud modeling in the 1970s through the present time, the reader is referred to the paper by Wilhelmson and Wicker[5] and the paper by Orf.[6]

While there is no universally agreed upon limit, simulations with grid spacings greater than ~100 m are very likely insufficient to properly resolve tornadoes and some of the underlying processes that instigate and maintain them. With modern supercomputing hardware, such as Blue Waters and Frontera, we have been able to conduct simulations with grid spacings as small as 10 m, resolving tornadoes and smaller, weaker whirls of spinning air oriented in both vertical and horizontal orientations, that are abundant throughout the simulation (see Figure 3).

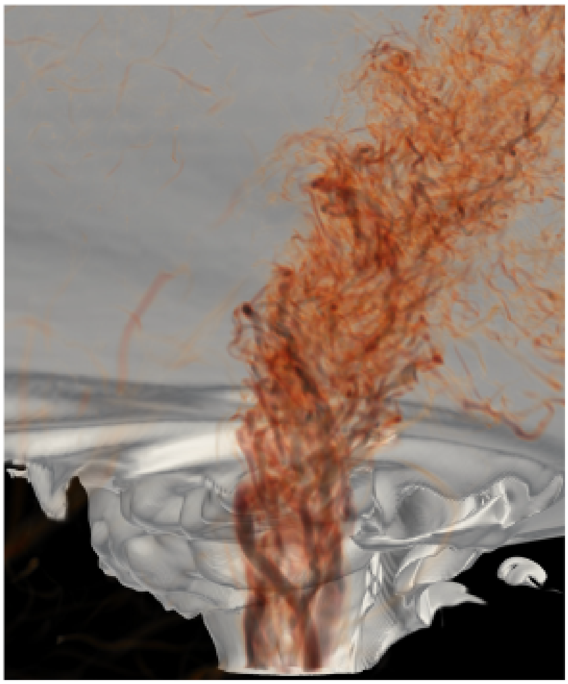

FIGURE 3. Cloud (grey transparent) and vorticity (brown) fields, volume rendered, during the maintenance phase of the multiple vortex EF5 strength tornado in the 10 m simulation. The tornado is approximately 1 km in diameter.

Breaking Through the Resolution Barrier

The work my research team has focused on over the past several years hinges upon the efficient utilization of modern supercomputing hardware by both a numerical model as well as analysis tools that will operate on model output long after the simulation completes. In 2010, I obtained early access, via the NSF proposal process, to the newly built Blue Waters supercomputer, with the goal of efficiently adapting a widely used cloud model to the hardware in order to simulate tornado-producing supercell thunderstorms at unprecedented resolution. CM1, a cloud model written and maintained by the National Center for the Atmospheric Sciences (NCAR) scientist George Bryan, is open-source and contains a permissive license.[7] CM1 was written to exploit distributed memory supercomputers, utilizing a hybrid nonblocking MPI/OpenMP parallel decomposition where communication overlaps with computation when possible. As such, CM1 has been shown to exhibit excellent weak and strong scaling performance to hundreds of thousands of MPI ranks spanning tens of thousands of multiprocessor compute nodes. Typical tornado-resolving simulations (with 30 m grid spacing) utilize 256 Frontera nodes (each of which contains 56 compute cores), spanning nearly 2 billion grid volumes. These simulations typically take less than 1 day to complete. The largest simulations being run on Frontera are run with 10 m grid spacing, spanning 1/4 trillion grid volumes across 7,700 Frontera nodes. Simulations of this size, spanning nearly the entire machine, are estimated to take a week or two to complete. A 10 m tornadic supercell simulation on Blue Waters took nearly a full month to complete, and the 250 TB of data from this simulation are currently being analyzed on Frontera.

Simple File System Enables New Discoveries

Despite its excellent scaling performance for computation, early benchmarks of CM1 on Blue Waters indicated that the frequent I/O required to achieve science goals would vastly dominate the runtime of simulations that we intended to conduct, greatly restricting the scope of the research that could be done. The following years were dedicated toward reducing I/O overhead of CM1, resulting in a new I/O and postprocessing framework collectively called Lack Of a File System (LOFS).[8] LOFS was developed primarily via trial and error, with an approach that made it possible to save model output at an extremely high temporal frequency (as high as the model's fraction-of-a-second time step, the smallest possible unit of time in the model) without taking an inordinate amount of runtime. This was achieved via the use of the HDF5 file format, exploiting the library's plugin architecture for utilizing modern compression approaches as well as its ability to grow HDF5 files in buffered memory over dozens of write cycles before flushing to disk.

BY INTELLIGENTLY COMPRESSING DATA, WE WERE BOTH ABLE TO REDUCE THE AMOUNT OF BUFFERED MEMORY REQUIRED TO HOLD THE DATA, AS WELL AS THE AMOUNT OF DISK SPACE FOR LONG-TERM STORAGE/ANALYSIS

We chose the lossy ZFP algorithm to compress our 3-D, 32 b floating point arrays.[9] ZFP compression provides a simple way to compress data based upon a single floating point value for each saved 3-D array that describes the maximum acceptable error (called the accuracy parameter). We have achieved compression ratios anywhere from 3:1 to over 100:1 with ZFP while maintaining enough accuracy for postprocessing and analysis. By intelligently compressing data, we were both able to reduce the amount of buffered memory required to hold the data, as well as the amount of disk space for long-term storage/analysis. Additional runtime performance was achieved by decomposing the physical domain of the model to match the underlying hardware in such a way as to minimize communication both during model execution and the assembly and buffering of I/O files.

On the read-side, LOFS provides a straightforward API that makes it easy to directly read model data, as well as a flexible routine called lofs2nc that reads any subregion of LOFS data at a given time, calculates new diagnostic quantities if requested, and writes output to Climate and Forecasting (CF) Compliant netCDF files that can be read by most visualization and analysis tools in the atmospheric sciences. While all I/O in LOFS is serial, excellent performance can be achieved by running hundreds or more individual routines concurrently across multicore nodes, reading from a shared parallel file system such as Lustre, with each process operating on a different model time. Similarly, while LOFS utilizes serial I/O when writing data, by writing hundreds to thousands of files concurrently (spread across many different directories) on a modern supercomputer with a parallel file system, excellent I/O throughput is achieved.

Hours of high-definition video made up of hundreds of thousands of individual rendered PNG files, spaced between 0.2 and 1 s apart in time, have been created by farming out render jobs across hundreds of compute nodes. Jobs are typically split up such that each process operates over the same model volume with the same visualization parameters, but each job renders at a unique model time. Software such as VisIt and Paraview can be run in batch mode, following scripts that describe how to render a scene, temporarily utilizing a section of the supercomputer as a render farm. These applications either access the LOFS API to retrieve the compressed model data directly or read from netCDF files generated with lofs2nc. The video animations made from sequencing these PNG files have revealed many fascinating aspects of supercell evolution including tornado formation, maintenance, and decay, and the attendant features such as the SVC and many small subtornatic vortices that move along density boundaries within the storm. By saving data at such high temporal frequency, the rapidly changing regions of the simulation (such as the rapid translation and rotation of the tornado and its multiple vortices) can be visualized smoothly and without aliasing artifacts.[10]

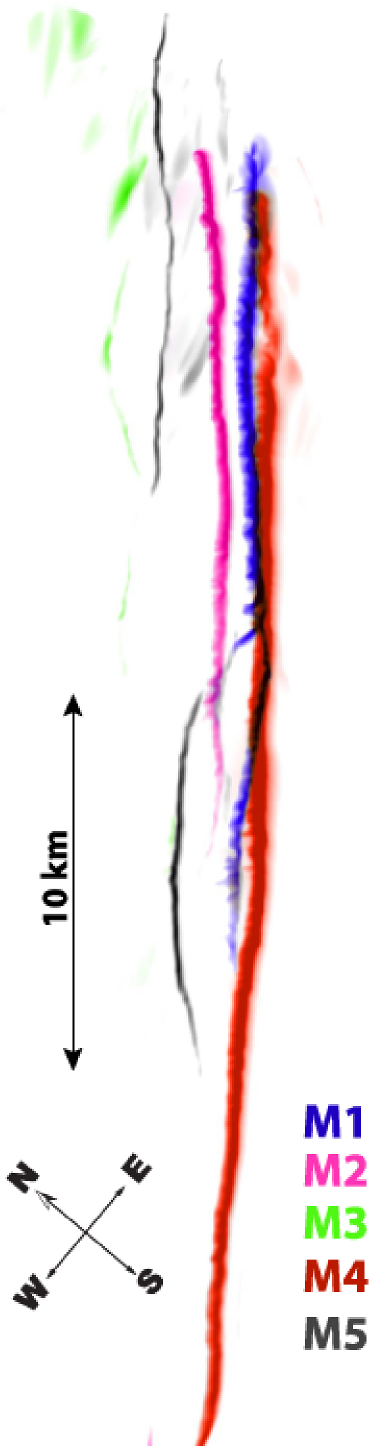

FIGURE 4. Overlaid locations of instantaneous ground-relative maximum winds exceeding EF3 in five members of a 30 m ensemble, labeled M1 through M5. The linear features represent tornado paths of each ensemble member, with variation in width, strength, and onset clearly evident across simulations.

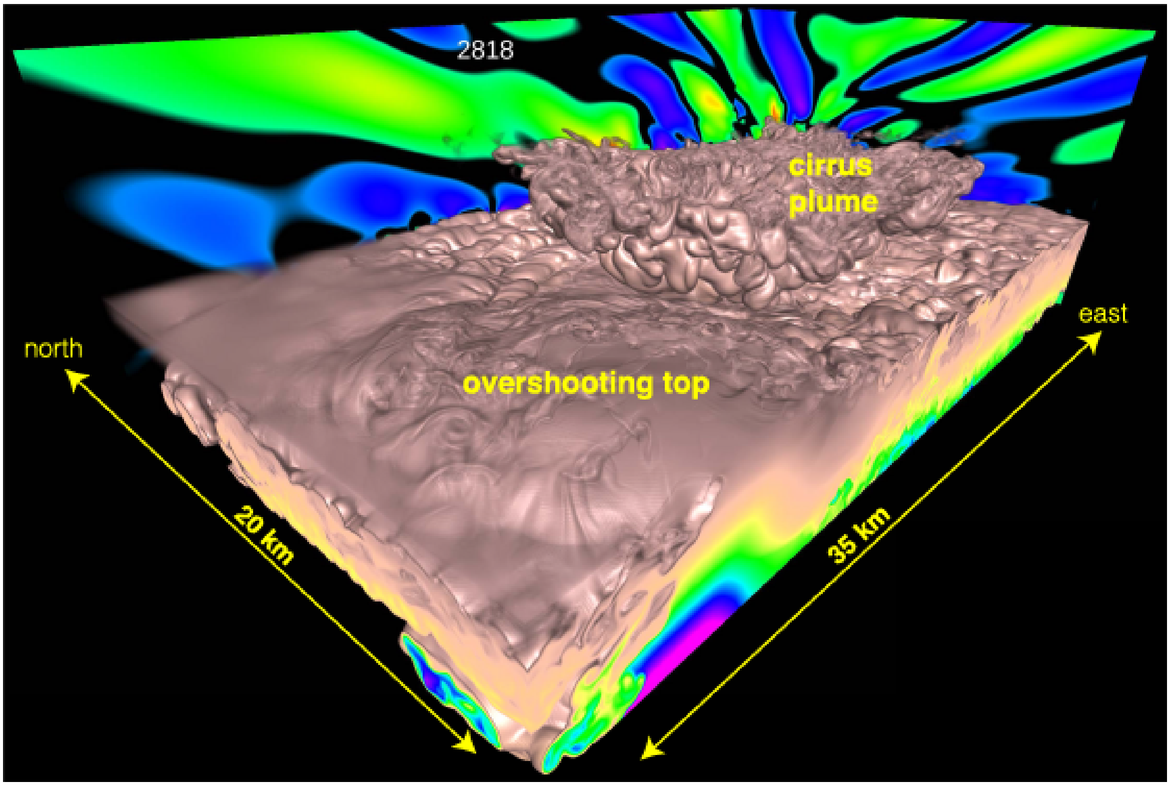

FIGURE 5. Subvolume of a 50 m supercell simulation, focused on the overshooting top and surrounding anvil 12–30 km above ground. The grey field is the volume rendered water vapor perturbation mixing ratio at the top of the supercell, focusing on air moister than its surrounding environment. A plume of cirrus embedded within moistened air extends upwards into the stratosphere downwind from the storm's overshooting top. The lateral walls are shaded by vertical velocity, with green (blue) indicating positive (negative) values and are associated with stratospheric gravity waves.

RESEARCH OBJECTIVES ON FRONTERA

More Supercells, More Tornadoes!

In 2019, after the decommissioning of Blue Waters for general science use, we obtained access to the NSF-sponsored Frontera supercomputer at the Texas Advanced Computing Center (TACC). In addition to having more compute cores per node, Frontera contains modern GPU hardware that can be exploited both during model run time and during postprocessing and analysis. Our research goals on Frontera include simulating supercells at tornado-resolving resolution (30 m or smaller grid spacing) in many different environments, where a spectrum of tornado intensities is captured, along with a collection of storms that do not produce tornadoes. By simulating a range of supercell types that approximates the diversity seen in nature at this resolution, we hope to better identify the key features that determine strength and duration of any tornadoes in simulated supercells, and whether such features can be related to the environment in which the storm formed or are more of a function of chaotic internal storm variability.

In order to explore the role of internal variability of supercells in determining tornado behavior, we have begun to investigate the sensitivity of simulations to model parameter settings at various resolutions by running ensembles of simulations in identical environments. In Figure 4, the paths of tornadoes in five 30 m ensemble members (M1–M5), based upon the 24 May 2011 El Reno EF5 supercell environment, are overlaid. Simulations varied by turbulence closure parameterization and some microphysical parameter settings of the CM1 model. A spectrum of outcomes is evident, varying from only marginal EF3 strength winds over a short narrow path to wide, violent, persistent EF5 tornado (that most closely matches the observed event). While there is some evidence of solution clustering, it is clear that the consequences

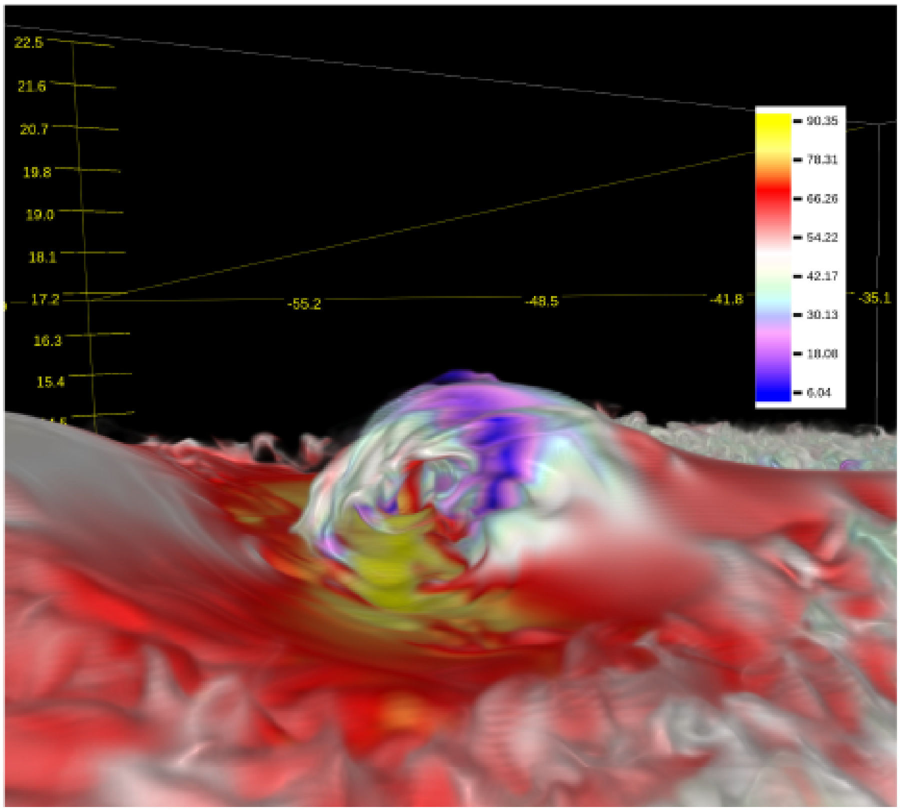

FIGURE 6. Cloud ice field (shaded by storm-relative wind speed) immediately downwind of the overshooting top. The feature resembling a breaking wave is associated with a thin, intense horizontal jet beneath its base, and precedes the formation of turbulent flow that results in a formation plume of cirrus in the stratosphere. Tick marks are in kilometer.

of a storm such as M4 encountering a region of high population density are dramatically more serious than a storm such as M3. Furthermore, there is cyclic behavior in some ensemble members, and the onset of tornado strength winds varies substantially between simulations. By comparing simulations, we hope to tease apart the physics of what is happening in the most violent simulations with the ultimate goal of identifying specific precursors to the most violent tornadoes that are not evident in less violent storms. Such precursors, if visible in real storms on the weather radar, could aid forecasters in providing more accurate warnings.

NEW ANALYSIS APPROACHES ON SAVED MODEL DATA ARE BEING EXPLORED TO ASSIST IN ANSWERING FUNDAMENTAL QUESTIONS ABOUT SUPERCELLS AND TORNADOES

HPC resources such as Frontera enable us to conduct dozens to hundreds of simulations of supercell thunderstorms at tornado-resolving resolution, and to save many terabytes of data for analysis, some of which will require additional supercomputing resources. New analysis approaches on saved model data are being explored to assist in answering fundamental questions about supercells and tornadoes. One such method, described below, involves following air parcels throughout the domain both forward and backwards in time using GPU technology. We are also exploring the use of an analysis method based upon Bayesian analysis and information field theory called entropy field decomposition (EFD). EFD analysis has been used successfully to identify modes that describe underlying patterns in data ranging from functional magnetic resonance imaging (fMRI) to mobile Doppler radar data of tornadic supercells. We are exploring the use of EFD analysis on simulation data as well as Doppler weather radar from real storms, seeking a unified analysis framework for observed and simulated data.[11]

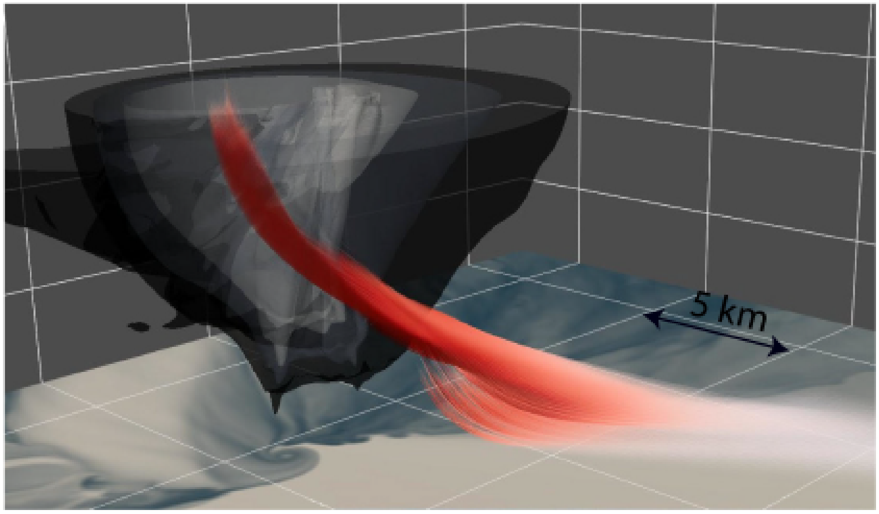

FIGURE 7. LOFT output for a 30 m tornadic supercell simulation, showing the path of thousands of parcels originating ahead of the storm near the ground and being drawn into the storm's updraft. The parcels are shaded by the stretching term in the vertical vorticity equation, with dark red indicating the intensification of rotation about a vertical axis in the same direction as the storm's rotating updraft. Updraft isosurfaces and surface buoyancy are also shown. LOFT is being used to analyze the source of rotation in both the SVC and tornado, as well as the forcing terms in the model's prognostic equations to determine what is responsible for significant updraft intensification near the ground, a process that appears to be linked to the SVC and its associated drop in pressure and is a precursor to the formation of the strongest tornadoes in numerical simulations. Image courtesy of Kelton Halbert.

Focusing on the Tops of Supercells

Because they are embedded within a large precipitating cloud and focused near the ground, tornadoes are not visible from weather satellites. What weather satellites can see are the cold tops of thunderstorms, including a broad region known as the anvil, that can hint to what is happening at lower levels. A satellite-identified feature called above an anvil cirrus plume (AACP) has been found to occur in about 75% of storms categorized as severe thunderstorms, storms that contain large hail and/or damaging winds.[12] These plumes are made up of cloud ice and appear as an elevated, elongated plateau of cloud ice trailing downwind from the storm's overshooting top that is situated directly above the updraft. Because AACPs are associated with severe weather, their appearance on satellite imagery, viewable in near-real time, provides another piece of information that can aid in operational weather forecasting. One of our science goals on Frontera is to better understand the dynamics of AACPs in supercells and their relationship to the formation of severe weather at the ground level.

It is typical for 3-D thunderstorm simulations in the published literature to utilize a stretched vertical mesh that focuses resolution near the ground, decreasing upwards to the top of model domain. This mesh geometry results in a poorly resolved cloud top. Our recent Frontera supercell simulations designed to better understand features found at the tops of supercell thunderstorms are run at 50 m isotropic grid spacing in order to resolve turbulent flow such as that which occurs in regions of strong wind shear (see Figures 5 and 6). Furthermore, the simulation is run in a large domain that extends 240 km in both lateral dimensions and 30 km vertically, well into the stratosphere, in order to avoid potential artifacts arising from situating the model's rigid top near the storm's dynamic overshooting top. A simulation of this size (spanning 14 billion grid volumes) can only be run with supercomputing resources such as Frontera. A recently identified feature found in one simulation currently under investigation is a shallow, but intense horizontal jet (with winds exceeding 90 m/s) that forms downwind of the storm's overshooting top, and which precedes the formation of a cirrus plume that extends into the stratosphere. This jet is shallow enough such that it would be unable to be resolved in previously published work, further underscoring the importance of powerful supercomputing resources such as Frontera.

Lagrangian Parcel Tracking Tool Using LOFS Data

The Lagrangian framework takes the point of reference of a massless particle moving with the air and is useful for tracing the air and its properties as it flows. Time series data of selected air parcels analyzed in this manner can show how the forces acting on parcels vary over time and space, telling a story about the evolution of different regions of the storm.

UW-Madison PhD student Kelton Halbert has developed Lagrangian offline flow trajectories (LOFT), CUDA C++ code that utilizes GPU hardware to conduct Lagrangian analysis from LOFS data (saved every model time step) tracing the paths of thousands to millions of individual air parcels, seeded at any arbitrary location in space and time (see Figure 7). This powerful form of analysis can be done repeatedly for a given simulation using only modest hardware. Typically, this form of analysis is done during model run time only, as parcels are updated every model time step. However, by saving data in LOFS format every model time step, repeated analysis can be conducted quickly that would normally require many expensive supercomputer simulations. Furthermore, parcels can be run backwards in time to trace their origin, something that cannot be achieved during a model simulation.

CONCLUSIONS

Advances in supercomputer hardware and the software that exploits this hardware have increased fundamental scientific understanding across many scientific disciplines. This is abundantly clear in the field of atmospheric science, with weather forecasting models in 2020 regularly producing accurate regional forecasts several days into the future, something impossible only two decades ago. While these models have shown steadily improving skill in forecasting the environments in which severe weather such as tornado-producing thunderstorms occur, they cannot currently forecast individual tornado-producing storms more than a few hours in advance. Basic research on supercell thunderstorms, however, can be conducted by simulating individual storms in an idealized framework.

Thunderstorm simulations conducted on the recently decommissioned Blue Waters supercomputer and its successor, Frontera, have advanced to the point where fully resolved tornadoes can occur in supercell simulations spanning domains that encapsulate the full storm and its immediate surrounding environment. Because of the damage and fatalities tornadoes cause, their accurate, timely prediction is a major goal of forecasters. The basic, highly idealized simulations such as those described in this article serve to elucidate processes occurring within supercells that result in a spectrum of outcomes, from weak, short-lived tornadoes to strong, persistent tornadoes. Supercomputers such as Frontera make it possible to simulate many individual storms in different environments at tornado-resolving resolution. It is hoped that by identifying specific factors that precede the formation of the most devastating supercell tornadoes, this research will result in improved forecasting of these events, ultimately reducing the loss of life caused by these powerful storms.

ACKNOWLEDGMENTS

This work was supported by National Science foundation grants AGS-1832327, OAC-1663954 and OAC-2031921, and the UW-Madison Cooperative Institute for Meteorological Satellite Studies. Frontera is supported by NSF grant OAC-1854828. The author would like to thank George Bryan for his freely available CM1 model, as well as all of the authors of open source software such as Linux, Python, HDF, NetCDF, ZFP, as well as visualization tools Ncview, VAPOR, VisIt, and Paraview. The author would also like to thank the four anonymous reviewers whose helpful comments improved the manuscript.

REFERENCES

1. P. M. Markowski, "What is the intrinsic predictability of tornadic supercell thunderstorms?," *Monthly Weather Rev.*, vol. 148, no. 8, pp. 3157–3180, May 2020, doi: 10.1175/MWR-D-20-0076.1.
2. J. Elston, B. Argrow, A. Houston, and E. Frew, "Design and validation of a system for targeted observations of tornadic supercells using unmanned aircraft," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 101–106, doi: 10.1109/IROS.2010.5649940.
3. L. Orf, R. Wilhelmson, B. Lee, C. Finley, and A. Houston, "Evolution of a long-track violent tornado within a simulated supercell," *Bull. Amer. Meteorol. Soc.*, vol. 98, no. 1, pp. 45–68, 2017. doi: 10.1175/BAMS-D-15-00073.1
4. A. Schueth, C. Weiss, and J. M. L. Dahl, "Comparing observations and simulations of the streamwise vorticity current and the forward flank convergence boundary in a supercell storm," *Monthly Weather Rev.*, vol. 1, Mar. 2021. [Online]. Available: https://journals.ametsoc.org/view/journals/mwre/aop/MWR-D-20-0251.1/MWR-D-20-0251.1.xml, doi: 10.1175/MWR-D-20-0251.1.

5. R. B. Wilhelmson and L. J. Wicker, "Numerical modeling of severe local storms," in Severe Convective Storms (series Meteorological Monographs). Boston, MA, USA: American Meteorological Society, 2001, pp. 123–166, doi: 10.1007/978-1-935704-06-5_4.
6. L. Orf, "High-resolution thunderstorm modeling," in *Oxford Research Encyclopedia of Climate Science*. London, U.K.: Oxford Univ. Press, Jan. 2020, doi: 10.1093/acrefore/9780190228620.013.667.
7. G. Bryan, CM1 Homepage. Accessed: Nov. 10, 2020. 2020. [Online]. Available: http://www2.mmm.ucar.edu/people/bryan/cm1
8. L. Orf, "A violently tornadic supercell thunderstorm simulation spanning a quarter-trillion grid volumes: Computational challenges, I/O framework, and visualizations of tornadogenesis," *Atmosphere*, vol. 10, no. 10, p. 578, Sep. 2019, doi: 10.3390/atmos10100578.
9. P. Lindstrom, "Fixed-rate compressed floating-point arrays," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 2674–2683, Dec. 2014, doi: 10.1109/TVCG.2014.2346458.
10. L. Orf, "A 10-m resolution quarter-trillion gridpoint tornadic supercell simulation," in *Proc. 100th Amer. Meteorological Soc. Annu. Meeting*. 2020. [Online]. Available: https://youtu.be/fXPCpEkjejo
11. L. R. Frank, V. L. Galinsky, L. Orf, and J. Wurman, "Dynamic multiscale modes of severe storm structure detected in mobile Doppler radar data by entropy field decomposition," *J. Atmos. Sci.*, vol. 75, no. 3, pp. 709–730, Dec. 2017, doi: 10.1175/JAS-D-17-0117.1.
12. K. Bedka, E. M. Murillo, C. R. Homeyer, B. Scarino, and H. Mersiovsky, "The above-anvil cirrus plume: An important severe weather indicator in visible and infrared satellite imagery," *Weather Forecasting*, vol. 33, no. 5, pp. 1159–1181, Oct. 2018, doi: 10.1175/WAF-D-18-0040.1.

LEIGH ORF is an atmospheric scientist who utilizes supercomputers to simulate damaging thunderstorms at eddy-resolving resolutions in order to better understand the behavior of damaging downbursts and tornadoes. He has held faculty positions at the University of North Carolina - Asheville and Central Michigan University. Since 2015, he has been a research scientist at the Space Science and Engineering Center at the University of Wisconsin, where his current work focuses on tornado behavior in supercell thunderstorms. He is the author of LOFS, a simple file system that enables numerical model data on massively parallel supercomputers to be saved at very high temporal frequency where it can be subsequently analyzed and visualized with very high fidelity. He received the Ph.D. degree from the University of Wisconsin in 1997. His research website can be found at http://orf.media. Contact him at leigh.orf@wisc.edu.