Improvements in numerical weather and climate modeling depend on accurate accounting of land–atmosphere interactions and the biological processes that mediate them at multiple spatial and temporal scales. Unfortunately, there is a problematic, persistent mismatch between the scales of observations and models. The substantial heterogeneity of the land surface means that observations do not always accurately reflect the entire model grid cell. Therefore, spatial and temporal scaling of surface fluxes is fundamental to how we evaluate theories on what happens within the subgrid of atmospheric models and how it feeds back onto larger-scale dynamics. This scaling is thus fundamental to assessing the parameterizations that represent land–atmosphere interactions in atmospheric models.
The notion that land-surface heterogeneity influences the surface energy balance—and resulting atmospheric responses—emerged from early model simulations showing the importance of soil moisture, vegetation, albedo, roughness, and heating to the atmosphere. There is no consensus on whether atmospheric boundary layer (ABL) responses to land-surface variations scale linearly or nonlinearly and whether they differ for dry versus moist dynamics. Scaling laws have been derived from numerical simulations, but a systematic regional-scale observational experiment to quantify multiscale subgrid scaling and patterning has long been needed. Now we have such an experiment: an ongoing National Science Foundation project based on an intensive field campaign that took place from June to October 2019. This is CHEESEHEAD19 (the Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors), which is designed to address longstanding puzzles in land–atmosphere exchange, including ABL responses to scales of spatial heterogeneity in surface–atmosphere heat and water exchanges.

One puzzle related to how heterogeneity influences transport in the ABL is the energy-balance closure problem: In eddy covariance (EC) flux measurements over sub-hourly time scales, the sum of incoming available energy [net radiation ($R_n$) minus ground heat flux ($G$)] tends to exceed surface turbulent sensible and latent heat fluxes. There is currently no definitive explanation for energy balance nonclosure, but one leading hypothesis is that surface heterogeneity generates mesoscale features not adequately resolved by traditional EC methods.
CHEESEHEAD19 was designed to disentangle how land-surface heterogeneity relates to atmospheric transport of mesoscale eddies, and thus enable a thorough investigation of the causes of energy balance non-closure. The field campaign concentrated an extensive suite of ground, tower, profiling, and airborne instrumentation over a 10 × 10 km domain of varied ecosystems of the Chequamegon-Nicolet National Forest in northern Wisconsin. Profiling observations characterized the mesoscale structure of atmospheric flows to an unprecedented degree. With CHEESEHEAD19 we strive to create a new class of observational flux data products that reconcile energy balance closure biases on the order of 10% and reveal actual surface emissions. For nonuniform exchange surfaces, like those featured in this project, this requires the evaluation of the conservation of mass and energy continuously in time and space throughout the study domain. However, even intensive field instrumentation campaigns such as CHEESEHEAD19 cannot produce observations everywhere, all the time, so the project evaluates data-driven methods for scaling surface energy fluxes, with the aim of improving model–data comparison and integration. The observations enable us to benchmark knowledge-guided machine-learning approaches and are being used alongside large-eddy simulation (LES) and scaling experiments to better understand submesoscale processes and improve formulations of subgrid-scale processes in numerical weather and climate models.

To this end, environmental response functions (ERFs) help attain the necessary information continuum from individual observation plots to model grid scale. Using multiple observation types, ERFs involve machine-learning to find relationships between measured fluxes and their meteorological and surface drivers, and georeference the results with unprecedented performance. This is a powerful approach not only for postfield data synthesis, but also for experiment planning—maximizing scientific return on experimental investment and thus closing the circle, turning knowledge into data and data into knowledge. Using an ERF to optimize experiment design allowed us not only to increase the scientific return on experimental investment but also to simplify the study’s necessary flight plans and increase crew safety.

**Experimental design**

CHEESEHEAD19 created one of the world’s highest-density networks of above-canopy EC measurements of surface energy fluxes. The project deployed a 20-tower EC network centered on the 447-m tower that anchors an AmeriFlux (US-PFa)/NOAA supersite (Park Falls WLEF) to gather data on humidity, temperature, greenhouse gases like carbon
dioxide, and more. The tower network was coupled with measurements of EC fluxes from aircraft; maps of leaf and canopy properties derived from airborne spectroscopy; ground-based measurements of plant productivity, phenology, and physiology; and atmospheric boundary layer profiles of wind, water vapor, and temperature using radar, sodar, lidar, microwave radiometers, infrared interferometers, and radiosondes.

**Selected preliminary results**

As is typical for EC measurements, we observed energy fluxes that were lower in magnitude than the available energy \( (R_N - G) \). The energy balance residual \( C_{EB} \) was largest during daytime, when incoming solar radiation was highest. The sign of \( C_{EB} \) reversed from day to night in part due to heat storage in the canopy, yet overall there was a daily mean imbalance.

Distinct seasonal shifting and spatial patterning in the surface energy fluxes appear to be directly related to surface characteristics. \( C_{EB} \) peaked during low turbulence—periods of calm wind and strongly unstable stratification—during which thermally induced mesoscale eddies resulting from landscape-scale heterogeneity are expected. This supports the hypothesis that mesoscale eddies are responsible for the energy balance nonclosure.

Landscape heterogeneity included vegetation, canopy height, surface temperature, and energy fluxes. Variability in surface sensible heat flux \( (H_s) \) was quantified by combining tower measurements with in situ measurements of air temperature and with land-surface temperature from a DJI S-1,000 (small uncrewed aircraft). For example, on 12 July, there were temperature differences of 10°C and \( H_s \) differences of 100 W m\(^{-2}\) over the 500 m \( \times \) 500 m area surrounding one of the EC towers. Spatial variability in temperature and \( H_s \) around the tower was directly related to surface heterogeneity. In short, datasets from the project show that changes in ABL development are closely tied to changes in the surface energy fluxes.

Two analysis approaches have been proposed to test the hypotheses of this study. The first is the application of ERF-VCV (environmental response function–virtual control volume)—a data-driven approach that can be used to account for the dispersive fluxes missed by single-tower EC measurements, and

---

**All 480 UWKA flight legs completed during the CHEESEHEAD19 field campaign. The yellow square represents the study domain and the red dots indicate the flux tower locations. During each IOP the UWKA flew over the study area to measure spatial EC fluxes of heat, water vapor, and CO\(_2\). The purpose of the airborne observations was to test flux tower scaling and observe atmospheric mesoscale patterning. The UWKA also measured cross-sectional profiles of water vapor and temperature below flight level using a downward-pointing Compact Raman Lidar (CRL) as well as ABL depth with the upward-looking Wyoming Cloud Lidar (WCL).**
Time series profiles at WLEF tall tower show the passing of a moist synoptic system on 24 Sep 2019 in H₂O mixing ratio and T measured by the ground-based MWR. Mesoscale eddies are seen in vertical wind speed from wind lidar LVS. Clouds are revealed in 532-nm backscatter from the ground-based HSRL. In the bottom panel, the energy balance components—available energy (Rₑ – G), and sensible and latent heat fluxes (Hₛ and Hₗ)—show that the energy balance residual (Cₑₑ), averaged across all EC towers, increases at the breakup of the morning inversion, is highest when mesoscale eddies are present midday, and dips when clouds form in the afternoon.

footprints, and storage fluxes. To simulate the physical processes as realistically as possible, we will assume realistic topography for the experiment site, and apply a land surface model (LSM) with a coupled soil and radiation model, as well as a plant canopy model (PCM). Both models are built into PALM. The LSM and PCM will allow us to study land–atmosphere feedbacks such as self-reinforcement of mesoscale circulations over the heterogeneous
study domain. The LSM will be set up for each IOP test case, with land-use classes, soil, and vegetation data as observed during the field experiment. Further, in order to account for synoptic-scale processes during the IOPs (e.g., advection of air masses), we will nest the LES domain into a larger-scale model.

One proposed goal is to derive a parametric heterogeneity correction of dispersive fluxes by setting up virtual towers within the LES, applying the correction to CHEESEHEAD19 tower flux field data, and evaluating the correction with ERF-VCV flux grids. Therefore, tower-level turbulence characteristics will be simulated in the LES as observed during the field campaign to investigate the energy balance nonclosure problem. Additionally, by emulating real-world measurements we intend to help interpret the observations—such as giving hints where secondary circulations occur or how far heterogeneity signals extend downwind.

Concluding remarks
In summary, CHEESEHEAD19 generates knowledge that advances the science of surface flux measurement and modeling, relevant to many scientific applications, including numerical weather prediction, climate change, energy resources, and computational fluid dynamics. Findings emerging from this project are expected to have broad implications for heterogeneous terrestrial regions beyond the specific study domain.

(a) ERFs (environmental response functions) augment sparse response observations (e.g., tower and aircraft EC) with abundant driver observations (e.g., meteorological stations and satellites) at minute and meter scales. Machine learning then extracts a driver-response process model. (b) During the experiment planning stage, we used LESs to create synthetic atmospheres over the CHEESEHEAD19 domain for different synoptic conditions. We simultaneously sampled the synthetic atmospheres as observed by different virtual experiment designs. Each experiment design resulted in a separate set of virtual observations that we independently processed through the ERFs [as seen in (c), a comparison with aircraft observing added in]. (c) We benchmarked the different experiment designs against their ability to reproduce the LES reference (top middle) in the form of flux grids that ERF reconstructed from the virtual observations alone.
are complex, with a great deal of spatial and temporal variation, which is difficult to capture with any single measurement technique. The CHEESEHEAD project shows how successfully a wide variety of instrumentation platforms can be used in coordination together, complementing each platform’s strengths and weaknesses.

Temple Lee (Cooperative Institute for Mesoscale Meteorological Studies and NOAA Air Resources Laboratory): The wealth of measurements obtained during CHEESEHEAD allowed us to document some pretty fascinating changes not only in the surface energy fluxes during the summer-to-fall transition period, but also the spatial patterns of these fluxes. Capturing the spatiotemporal variability in the fluxes is critical for helping to improve mesoscale meteorological models.

Luise Wanner (Karlsruhe Institute of Technologie): During studies for my master’s degree in geoeecology, I specialized in micrometeorology because I was so fascinated by the interaction between ecosystems and the atmosphere. I got to know eddy covariance measurements, which are a really great method to study the exchange of energy and gases like CO₂ between vegetation and atmosphere at the ecosystem level. But I also found it exciting that even this well-established and widely used method still has its little weakness: we don’t know exactly where a small fraction of the energy is going—the energy balance closure problem. And as a Ph.D. student working on this project, I hope to be able to contribute to solving this mystery.

Ankur Desai (University of Wisconsin—Madison): I came up with the original idea while on sabbatical in the Bavarian Alps, seven years ago, working with land–atmosphere modeling colleagues at KIT IMK IFU in Garmisch-Partenkirchen, Germany. At the time, NCAR was on the verge of expanding its eddy-covariance flux tower facilities and several other instruments were coming online, and it got me thinking about how we could resolve some outstanding problems.

BAMS: How did you become interested in the topic of this article?

BB: I became interested in how heterogeneity influences surface fluxes over the course of my previous research projects looking at CO₂ flux in polar oceans where there was patchy sea ice. I was excited to switch to forest research, where the flux signals are larger and ancillary measurements are easier (though still difficult) to acquire. I was intrigued by this project because of the sheer number and variety of measurements being thrown at the problem.

BAMS: What would you like readers to learn from this article?

Brian Butterworth (University of Wisconsin—Madison): I would like readers to learn that differing scales of surface features (like heterogeneity in vegetation) interact with scales of atmospheric motions in ways that are not always captured by weather models. I hope this study helps show that thoughtful experiment design can go a long way toward investigating some of the unknown and unresolved features in meteorological modeling. I also hope readers walk away knowing that the data collected in this project are freely available online. With nearly inexhaustible potential for interrogating exciting research questions, we hope other researchers will use them.

David Plummer (University of Wyoming): Boundary layer fluxes...
in our field with a comprehensive field study. The idea that our flux tower measurements are “missing” some components of the surface energy balance has been a problem for a long time. Several theories have come out having implications for how these measurements are used to benchmark weather and climate models and how we scale up flux measurements to regions. It only seemed natural to find a way to put them to the test, somewhere with a range of short and tall vegetation types, heterogeneous land surfaces, and varying atmospheric weather patterns.

**BAMS:** What surprised you the most about the work you document in this article?

**BB:** I was surprised by how variable the energy balance residual (nonclosure term) was across the domain, even in nearby towers. I also was surprised by how obvious the signature of mesoscale eddies was in the wind profiles on certain days—for example, when we saw oscillations of 30 minutes of strong updraft followed by 30 minutes of strong downdraft. While there were more questions raised than could be addressed in this one article, I was surprised at the enormous potential of this dataset for investigating them.

**James Mineau (University of Wisconsin—Madison):** This was my first time being a part of a research project. I was amazed by the collaboration and logistics that went into this comprehensive, densely instrumented study. Part of my duties as intern were to help introduce the incoming scientists to the local area and the domain. It was fascinating meeting all of the different researchers from different backgrounds, converging on this small town that I grew up in, and learning what they were here to research and how they got to this point in their careers. There were so many different organizations all with their own questions, but a shared goal of better understanding the heterogeneous landscape that we live in.

**LW:** I am still impressed by the sheer number of measurements we have collected in the three study months, especially during the intensive observation periods. I am currently working on setting up a simulation that is as realistic as possible. For this, I need a lot of different information, but no matter what I ask for, someone has measured it. This is exceptionally convenient and clearly not the case with every measurement campaign.

**BAMS:** What was the biggest challenge you encountered while doing this work?

**BB:** For me, the most difficult were the logistical challenges associated with having so many different teams and instruments come to the field. I helped coordinate lodging, instrument deliveries and installation, flight patterns, data collection schedules, etc., for the roughly 40 researchers who participated in field work.

**DP:** Coming into the project from the airborne measurements side, I’m always fascinated by the opportunity to be involved in something that provides an interesting new challenge . . . such as the planning of flight patterns that would safely allow the aircraft to fly as low as 300 feet above ground, all while coordinating with other researchers using weather balloons and instrumented drones in the same general area.

**AD:** Though I’ve worked in northern Wisconsin for two decades, I was not prepared for the level of intensity that would be required to permit and prep so many sites and set up so many instruments in so short a period. And I was pleasantly surprised by the level of dedication, professionalism, and perseverance that all groups made in making it happen. The mosquitoes and humidity were brutal sometimes, and occasionally a drone got lost or a tower was a bit askew, but the observations collected have all been fantastic and fascinating so far. I believe this experiment will be a landmark for this kind of work for some time.

**BAMS:** What’s next? How will you follow up?

**BB:** We will start incorporating in the analysis all the datasets that had longer spin-up periods, like connecting hyperspectral imagery maps of leaf characteristics to tower-measured primary production and connecting drone-based measurements of canopy structure and heat flux to eddy covariance EC tower fluxes. We will also run a machine-learning algorithm to map domain-wide fluxes with the goal of providing useful tools for incorporating the role of heterogeneity into meteorological models.

**TL:** More measurements! Conducting experiments similar to CHEESEHEAD in regions with surface characteristics different from those within the CHEESEHEAD domain will be critical for providing additional insights into land–atmosphere feedbacks and processes.

**DP:** Following CHEESEHEAD up with a project using a similar design methodology but for other land surface types or seasonal changes would be a fantastic research opportunity.
Monthly Weather Review

Monthly Weather Review (MWR) publishes research relevant to the analysis and prediction of observed atmospheric circulations and physics, including technique development, data assimilation, model validation, and relevant case studies. This research includes numerical and data assimilation techniques that apply to the atmosphere and/or ocean environments. MWR also addresses phenomena having seasonal and subseasonal time scales.

Open Access option for authors. All papers publicly accessible 12 months after publication.


Editorial Board

David M. Schultz, University of Manchester (Chief Editor)
Jeffrey Anderson, National Center for Atmospheric Research
Tommaso Benacchio, Politecnico di Milano
Kristen L. Corbosiero, University at Albany, State University of New York
Matthew D. Eastin, University of North Carolina at Charlotte
Clark Evans, University of Wisconsin-Milwaukee
Jidong Gao, NOAA/National Severe Storms Laboratory
Almut Gassmann, Institute for Atmospheric Physics, Kühlungsborn
Joshua P. Hacker, National Center for Atmospheric Research
Daniel Hodys, Naval Research Laboratory
Daryl Kleist, NOAA/NWS/NCEP/Environmental Modeling Center
Matthew R. Kumjian, The Pennsylvania State University
Ron McTaggart-Cowan, Environment and Climate Change Canada
Derek Posselt, NASA Jet Propulsion Laboratory, California Institute of Technology
Glen Romine, National Center for Atmospheric Research
Paul Roundy, University at Albany, State University of New York
Angela Rowe, University of Washington
Elizabeth Satterfield, Naval Research Laboratory
Michael Scheuerer, Norwegian Computing Center, NR
Russ S. Schumacher, Colorado State University
Stan Trier, National Center for Atmospheric Research
Christopher Weiss, Texas Tech University

Monthly Weather Review (MWR)
ISSN: 0027-0644; eISSN: 1520-0493