

Earth and Space Science



RESEARCH ARTICLE

10.1029/2021EA001842

Special Section:

Advances in scaling and modeling of land-atmosphere interactions

Key Points:

- Fusion of satellites with models for high space and time resolution land surface temperature needed for many surface-atmosphere studies
- Developed an approach that evaluates well across array of towers and aircraft observations from an intensive field experiment
- Additional downscaling with airborne hyperspectral imagery further refines the identification of hot spots as evaluated with drone observations

Correspondence to:

A. R. Desai,
desai@aos.wisc.edu

Citation:

Desai, A. R., Khan, A. M., Zheng, T., Paleri, S., Butterworth, B., Lee, T. R., et al. (2021). Multi-sensor approach for high space and time resolution land surface temperature. *Earth and Space Science*, 8, e2021EA001842. <https://doi.org/10.1029/2021EA001842>

Received 7 MAY 2021
Accepted 31 AUG 2021

Multi-Sensor Approach for High Space and Time Resolution Land Surface Temperature

Ankur R. Desai¹ , Anam M. Khan² , Ting Zheng³ , Sreenath Paleri¹ , Brian Butterworth¹ , Temple R. Lee^{4,5} , Joshua B. Fisher⁶ , Glynn Hulley⁷ , Tania Kleynhans⁸, Aaron Gerace⁸, Philip A. Townsend³ , Paul Stoy⁹ , and Stefan Metzger^{10,1}
¹Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI, USA, ²Nelson Institute for Environmental Studies, University of Wisconsin, Madison, WI, USA, ³Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, Madison, WI, USA, ⁴Cooperative Institute for Mesoscale Meteorological Studies, Norman, OK, USA, ⁵NOAA Air Resources Laboratory Atmospheric Turbulence and Diffusion Division, Oak Ridge, TN, USA, ⁶Schmid College of Science and Technology, Chapman University, Orange, CA, USA, ⁷Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, ⁸Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, USA, ⁹Department of Biological Systems Engineering, University of Wisconsin—Madison, Madison, WI, USA, ¹⁰Battelle, National Ecological Observatory Network, Boulder, CO, USA

Abstract Surface-atmosphere fluxes and their drivers vary across space and time. A growing area of interest is in downscaling, localizing, and/or resolving sub-grid scale energy, water, and carbon fluxes and drivers. Existing downscaling methods require inputs of land surface properties at relatively high spatial (e.g., sub-kilometer) and temporal (e.g., hourly) resolutions, but many observed land surface drivers are not continuously available at these resolutions. We evaluate an approach to overcome this challenge for land surface temperature (LST), a World Meteorological Organization Essential Climate Variable and a key driver for surface heat fluxes. The Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors (CHEESEHEAD19) field experiment provided a scalable testbed. We downscaled LST from satellites (GOES-16 and ECOSystem Spaceborne Thermal Radiometer Experiment on Space Station [ECOSTRESS]) with further refinement using airborne hyperspectral imagery. Temporally and spatially downscaled LST compared well to independent observations from a network of 20 micrometeorological towers and piloted aircrafts in addition to Landsat-based LST retrieval and drone-based LST observed at one tower site. The downscaled 50-m hourly LST showed good relationships with tower ($r^2 = 0.79$, RMSE = 3.5 K) and airborne ($r^2 = 0.75$, RMSE = 2.4 K) observations over space and time, with precision lower over wetlands and lakes, and some improvement for capturing spatio-temporal variation compared to a geostationary satellite. Further downscaling to 10 m using hyperspectral imagery resolved hot and cold spots across the landscape as evidenced by independent drone LST, with significant reduction in RMSE by 1.3 K. These results demonstrate a simple pathway for multi-sensor retrieval of high space and time resolution LST.

Plain Language Summary The temperature of the Earth's surface over land—land surface temperature (LST)—is an important variable to observe and forecast. Variation in LST over space and time at scales of meters and hours influence processes in the atmosphere, soils, vegetation, and water. For the worldwide coverage of LST, we rely on Earth-observing satellites. However, there are trade offs in how finely LST can be observed over space versus how often LST can be observed over time, given the characteristics of any one satellite's orbit, not to mention the obscuring effect of clouds. Therefore, methods are needed that enable data from multiple satellites as well as aircraft and towers if we want to observe LST at high space and time resolution. Here, we develop such an approach and test its accuracy over a test bed of extensive LST observations made by towers, drones, and aircraft during a field experiment in Northern Wisconsin USA.

1. Introduction

Land surface temperature (LST) is a World Meteorological Organization Essential Climate Variable that links the thermodynamics of earth's land surface with the dynamics of the overlying atmosphere (Dirmeyer

et al., 2012; Dolman et al., 2016). LST, equivalent to surface skin temperature, refers to the apparent temperature of an infinitesimally thin surface of ground (English, 2008). It is a consequence of the difference in the net radiative energy budget of the surface and rate of heat conduction into the ground. LST can vary greatly over short distances (Yi et al., 2020), as anyone who has walked across wet and dry sand on a beach during a sunny summer day can attest. For LST observation systems, then, the challenge becomes how to integrate that variation at space and time scales relevant to land-atmosphere interactions.

LST is most commonly measured based on principles related to radiative observations made across various wavelengths in the thermal infrared spectrum, given the tight relationship of electromagnetic blackbody radiation to temperature, as provided by the Planck function and in integrated form to the Stefan-Boltzmann relationship. The peak of earth's outgoing surface longwave radiation is in 8–12 micrometers and thermal infrared brightness temperatures reflect surface temperatures integrated over a few micrometers, making it a good proxy for LST (Hulley & Ghent, 2019). After the calculation of emissivity, these observations allow for the inversion of LST from longwave radiation measurements (Wang et al., 2014). On a fixed or moving platform, thermopile sensors facing earth can measure longwave radiation and be used to calculate *in situ* LST, after accounting for atmospheric correction. Typically, LST observations on a fixed grid are derived from thermal infrared brightness temperature or outgoing longwave radiation observations made by earth-observing satellites, in polar, irregular, or geostationary orbits (English, 2008; Li et al., 2013; Scarino et al., 2013). Orbits, costs, and logistics lead to trade offs retrieving high time frequency (typically from geostationary orbits) versus high spatial resolution (typical from polar or irregular orbits). Additionally, satellite LST is not easily retrieved in areas under heavy cloud cover.

Continuous high time and space resolution LST, including the diel cycle, is of high value for a number of scientific investigations (e.g., Kröniger et al., 2019). LST can vary by tens of degrees K over meters and change within seconds to hours, for example, due to shadows, wind, passing of clouds (Yi et al., 2020), or irrigation. These changes in LST then influence the heating of the soil, vegetation, and atmosphere over the course of the day (Dirmeyer et al., 2012; Taylor et al., 2012), and the dynamics that ensue as a result. In many land surface models, for example, those used in numerical weather prediction, LST is usually a derived value inferred from the modeled surface energy balance and soil physics, often averaged over an entire grid cell or a land cover tile, and not resolved at scales below hundreds of meters. Continuous LST over scales of meters and hours would provide a valuable benchmark to evaluating atmospheric surface layer and soil heat diffusion parameterizations, estimating turbulent heat fluxes (K. Xu et al., 2018), assimilation of LST for model grids (Bosilovich et al., 2007; Zheng et al., 2012), scaling of land-atmosphere fluxes and feedbacks (Metzger, 2018; T. Xu et al., 2018), and answering science questions related to fine-scale sub-kilometer space and sub-daily time heterogeneity of landscapes and habitats (Guillevic et al., 2019; Pincebourde and Salle., 2020). Biological organisms, in particular, are strongly influenced by small-scale microclimates and scaling these responses across regions is nonlinear (Bütikofer et al., 2020).

Given these needs, fusion approaches have been designed to combine multiple satellite data products and increase their joint space, time, and clear sky coverage (Anderson et al., 2021; Gao et al., 2012; Hu et al., 2020; Liu et al., 2006). However, current and upcoming generation satellites and computational capacity provide an even richer array of data fusion options (Freitas et al., 2013; Khan et al., 2021; Tomlinson et al., 2011). For example, NASA's ECOSystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS) is a thermal imager on the International Space Station (ISS) that, from this relatively low (~400 km) and fast precessing orbit, can image the globe at roughly 1–5 day repeat (at different hours of the day of every orbit) and at 70-m resolution (Fisher et al., 2020). Meanwhile, the latest NOAA Geostationary Operational Environmental Satellites (GOES-16 and GOES-17) image the Western Hemisphere at a nominal 15-min time-step with approximately 2 km resolution depending on view geometry. Fusion of these products have not been evaluated. High space and time resolution LST has been attempted in some locations (e.g., Sismanidis et al., 2016a, 2016b, 2018), but there is a need for greater evaluation across multiple approaches and sensors.

A number of remotely sensed features beyond thermal infrared also relate to LST and could improve down-scaling (Yue et al., 2020). For example, observations in visible and microwave wavelengths relate to processes such as vegetation activity and soil moisture, respectively, that in turn relate to fine-scale variation in LST. Hyperspectral remote sensing (*aka* imaging spectroscopy), in particular, may allow for fine-tuning

of LST by linking to surface mineralogy and crown-level foliar functional characteristics that affect foliar thermodynamics.

Prior studies often lacked a comprehensive spatial and temporal database of *in situ* LST at relevant space and time scales for evaluating LST fusion products and their uncertainty, critical for model assimilation (Bosilovich et al., 2007; Freitas et al., 2010). The recent Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors (CHEESEHEAD19) (Butterworth et al., 2021) field campaign included an array of towers, drones, and aircraft, in addition to remote-sensing thermal imagery from Landsat-8 (Gerace et al., 2020) that provides a comprehensive, open-access test bed for any fusion approach. Radiometric-derived LST over various landscapes is available over a 4-month period across a nearly 1,000 km² area of a heterogeneous, flat landscape of northern Wisconsin USA. Furthermore, visible and near infrared hyperspectral airborne imagery at 1 m resolution was flown in the domain several times, providing a second data source to evaluate alternative downscaling and fusion approaches based on surface cover characteristics rather than emissivity.

Here, we evaluate a novel high space (50 and 10 m) and time (hourly) resolution LST fusion approach using next generation thermal imagery. We ask: how reliably can we fuse high space and high temporal resolution satellites to generate continuous, cloud-free gridded LST? Further, hyper-resolution drone LST imagery at the submeter scale allows us to further evaluate downscaling of this gridded product to even smaller domains, necessary for some scientific applications (Pincebourde Salle., 2020). Thus, within a subset of our study area, we also test whether we can further downscale to higher resolution by connecting hyperspectral indices combined with the LST fusion. Finally, we discuss the implication of the work for advancing land-atmosphere interaction science.

2. Materials and Methods

Our general approach employs hierarchical fusion (Figure 1). As a prior, cloud-free, coarse-resolution (12 km) estimate of LST, we used data assimilation-constrained hourly LST from a set of three land surface models. These modeled LSTs are then fit on a pixel level to gap-fill geostationary satellite LST to generate gap-free medium resolution (1–2 km) hourly LST. Further spatial downscaling is accomplished using the suite of cloud-screened, quality-controlled high-resolution (50 m) LST and generating a regression surface that links the medium and high-resolution LST across all collected time points. The resulting high space and time resolution LST grids are then evaluated against a range of independent tower, aircraft, and satellite estimates of LST. Finally, an additional higher resolution downscaling to 10 m is conducted using hyperspectral imagery over an area where coincident submeter resolution drone LST was also measured.

2.1. Site Description

Analyses are centered on the observations collected during the CHEESEHEAD19 field campaign (Butterworth et al., 2021) conducted near Park Falls, Wisconsin, USA, in the central region of the North American continent from June to October 2019. CHEESEHEAD19 was an intensive surface-atmosphere field experiment investigating the role of surface spatial heterogeneity on atmospheric dynamics and the surface energy balance. As a result, a suite of observations was collected over a 10 × 10 km core domain and a 30 × 30 km extended domain, centered on the WLEF Park Falls Ameriflux very tall eddy covariance (US-PFa) tower, which is also an NOAA greenhouse gas (LEF) tall tower. Observations included 20 micrometeorology towers within the core domain, ground-based atmospheric profiling, drone and airborne remote sensing at various locations throughout, and more than 10,000 km of low-level meteorological aircraft observations in the extended domain. Upwelling and downwelling longwave radiation observations from towers, IR skin temperature retrieved from aircraft, and an independent satellite LST estimate from Landsat were used here to evaluate the LST product.

2.2. Input Data

All data products used for the generation of 50 m high and 10 m higher resolution LST were acquired from public open-access data repositories (Table 1). Each data product was extracted for all acquisitions from

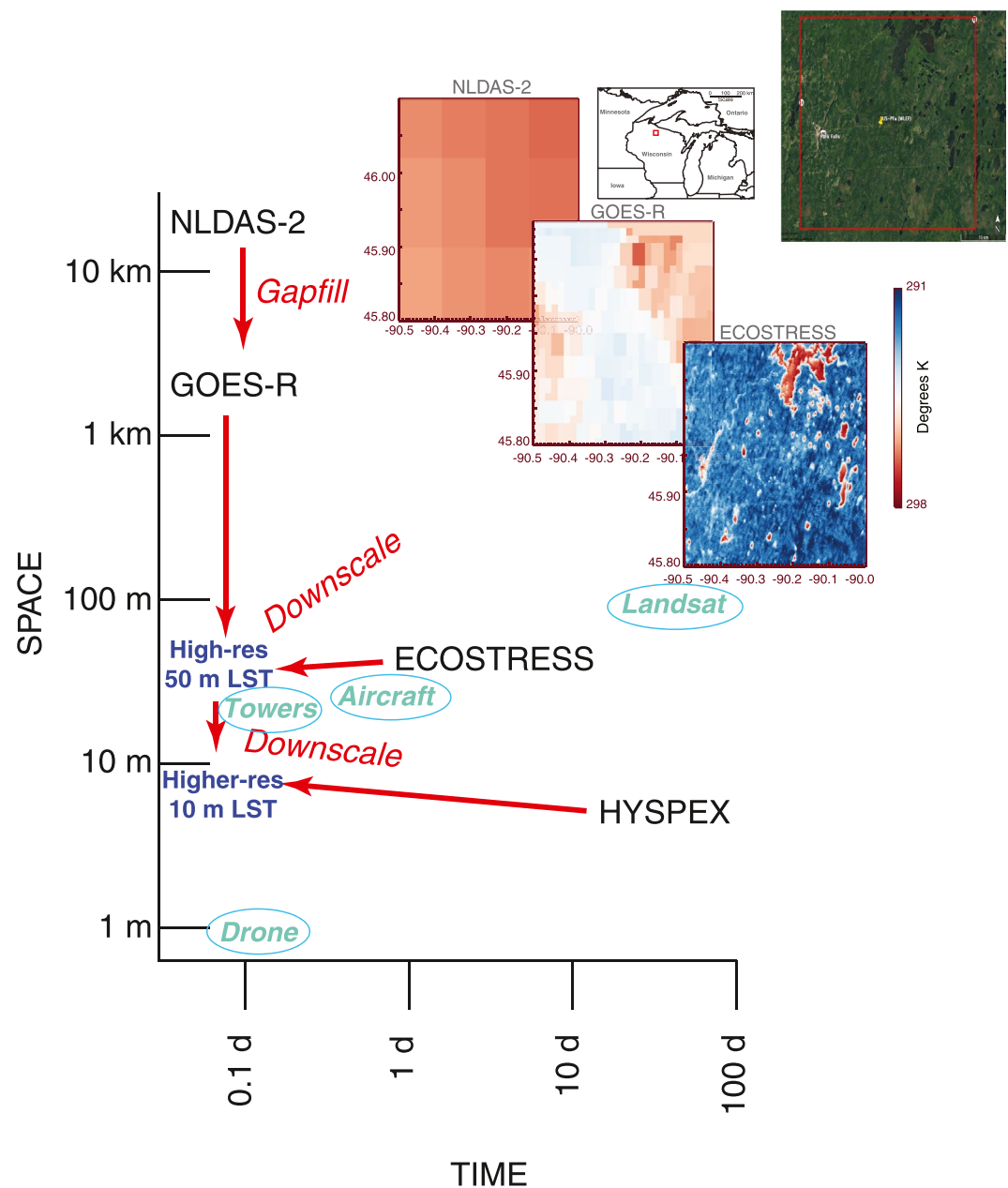


Figure 1. Stommel diagram schematic of space and time scale of input data products (black text), evaluation land surface temperature (LST) (cyan), high-resolution (50 m) and higher-resolution (10 m) downscaled LST (dark blue), and processes to create those (red arrows and text) over The Chequamegon Heterogeneous Ecosystem Energy-balance Study Enabled by a High-density Extensive Array of Detectors (CHEESEHEAD19) domain (map, upper right, and red box). Example input LST imagery is shown from August 7, 2019 0Z.

June 1 to October 31, 2019 and subset to a domain that encompassed the CHEESEHEAD19 extended domain (Figure 1). Descriptions of each data product are provided here.

For the prior modeled LST, we acquired LST from the National Land Data Assimilation System version 2 (NLDAS-2) (Xia et al., 2012). NLDAS is an observation reanalysis that constructs an optimal meteorological driver forcing based on gauge precipitation and bias-corrected shortwave radiation. This forcing is provided to a suite of land surface models, which output a common set of responses, including LST. NLDAS products are provided on a $\frac{1}{8}$ degree grid (approximately 12.5 km) across North America at hourly timestep. We

Table 1
Input Data Sources Used in Gap-Filling and Downscaling Land Surface Temperature

Product	Description	Spatial resolution	Temporal repeat frequency	URL
NLDAS-2	Data-assimilation model reanalysis LST	1/8 degree	hourly	https://ldas.gsfc.nasa.gov/nldas/
GOES-R	Geostationary satellite LST over Western Hemisphere	~2 km	15 min	https://www.goes-r.gov/products/baseline-LST.html
ECOSTRESS	Thermal imager on International Space Station	70 × 70 m	1–5 days; diurnal sampling	https://ecostress.jpl.nasa.gov/
UW HySpex	Visible to Shortwave IR airborne hyperspectral imager 400–2,500 nm	Varies, ~1 m	~Monthly	https://data.eol.ucar.edu/dataset/592.027

extracted LST for the three land surface models that are part of NLDAS and output surface skin temperature: Mosaic (Koster & Suarez, 1992), Noah-2.8 (Chen et al., 1996), and VIC (Liang et al., 1994). We calculated mean and variance moments on the modeled LST as a prior preparation.

NOAA's Geostationary Operational Environmental Satellites (GOES) are the primary U.S. operational geostationary weather satellites in orbit over the Western Hemisphere (Schmit et al., 2017). In recent years, LST has become a primary operational product of the GOES-R Advanced Baseline Imager (ABI) in the current generation GOES-16 and GOES-17 satellites (Yu et al., 2009). These outputs, at an approximately 2 km spatial resolution, are produced based on thermal channel split-window retrieval using the 11.2 and 12.3 μm channels with high surface emission and low atmospheric absorption. The algorithm also uses prescribed surface emissivity and an atmospheric radiative transfer model to produce an output at least once an hour for the Northern Hemisphere (more fully described at: <https://www.goes-r.gov/products/baseline-LST.html>). Target accuracy is 2.5 K and evaluations have shown it to be approaching 1.5 K (Yu et al., 2012).

ECOSTRESS is a thermal imager flown on the ISS (Fisher et al., 2020; Hulley et al., 2017). ECOSTRESS was launched in June 2018 and has been providing harmonized Level 2 70 × 70 m data products on surface temperature, evapotranspiration, water use efficiency, and drought stress since launch. We acquired the Level 2 LST and Emissivity product and the ECOSTRESS cloud cover product (described at https://lpdaac.usgs.gov/documents/423/ECO2_User_Guide_V1.pdf). LST is derived from a physically based Temperature and Emissivity Separation (TES) algorithm (Gillespie et al., 1998; Hulley & Hook, 2011). Atmospheric correction is performed using the RTTOV radiative transfer model (Matricardi, 2008; Saunders et al., 1999). Retrieval is based on thermal radiances in the 8.29, 8.78, 9.20, 10.49, and 12.09 μm bands. Validation accuracy was reported as 1.07 K (Hulley et al., 2021). QA flags were used to limit to best or nominal quality observations. The ECOSTRESS Level-2 CLD cloud-mask (<https://lpdaac.usgs.gov/products/eco2cldv001/>) was applied to mask any cloud-contaminated pixels. The ISS orbit is not sun-synchronous, so scenes are retrieved at different times of the day, with a repeat interval of 1–5 days depending on the location. A total of 118 full domain scenes were retrieved over the CHEESEHEAD19 domain during the study period spanning all hours of the day. Of these, 49 images were at least 50% cloud free. From that subset, 25 images (~weekly) were retained that were significantly ($p < 0.001$) correlated ($r > 0.3$) with the GOES imagery and had <50% cloud cover. This quality control filter was applied to remove those ECOSTRESS scenes that had significant noise, offsets, or missing regions.

The University of Wisconsin hyperspectral imager is a visible to near-infrared (400–2,500 nm) imaging spectrometer designed for airborne applications (HySpex, Norsk Elektro Optikk, Oslo, Norway). The HySpex consists of two boresighted imagers measuring a total of 474 narrow bands between 400 and 1,000 nm (3.26 nm spectral resolution) and 930–2,500 nm (5.45 nm spectral resolution). In CHEESEHEAD19, the HySpex was flown on a State of Wisconsin, Department of Transportation Cessna 210 at 1,400 m altitude above ground, allowing for a nominal 1 m pixel size over the core domain. The HySpex was flown multiple times over the study period (26 June through 30 August). The CHEESEHEAD19 study area is covered by 21 flight lines flown ± 2 h around solar noon. Here, observations from dates closest to the date of the drone overflight were used. Images were orthorectified (following Schläpfer & Richter, 2002) and atmospherically corrected (following Adler-Golden et al., 1999) to surface reflectance using LibRadTran (Emde et al., 2016) open-source code by Liu et al. (2019). Flight lines were subset to the regions of drone overflights. In our

Table 2
Evaluation Data Sources

Product	Description	Spatial resolution	Temporal repeat frequency	URL
University of Wyoming King Air (UWKA)	Upwelling infrared surface temperature	~10 m	Twice-daily over three 4-day periods	http://flights.uwyo.edu/projects/cheesehead19/
NCAR Integrated Surface Flux Station (ISFS)	Upwelling longwave radiation from 19 eddy covariance towers	~50 m	5-min average	https://data.eol.ucar.edu/dataset/592.025
Landsat split-window LST	Satellite land surface temperature	100 m	~16 days	https://doi.org/10.3390/rs12020224
NOAA UAS	Drone-based land surface temperature	Varies, ~1 m	Hourly in daytime over two 4-day periods	https://data.eol.ucar.edu/dataset/592.010

analyses, we used a normalized spectral index approach that reduced the need for additional processing to reduce bidirectional reflectance variation.

2.3. Evaluating LST Data

Several data products were used to evaluate the 50 and 10 m downscaled LST. These are noted in Table 2 and briefly described here. The University of Wyoming King Air (UWKA) is a meteorological research aircraft that flew linear transects in CHEESEHEAD19 focused on eddy covariance applications during three 4-day periods within the experiment window (July 9–13, 2019, August 20–23, 2019, and September 24–28, 2019). Flights were flown in mid-morning and mid-afternoon, usually 15 legs at 100 and 400 m altitude above ground spanning the 30 × 30 km extended domain, at approximately 90 m s⁻¹. UWKA included a downward-looking radiative thermometer (Heimann KT-19.85), which reports observed brightness temperature for the 9.5–11.5 μm IR spectrum with 0.5 K accuracy and 0.2 K RMSE over 1 s (~90 m at flight speed). This instrument reported temperature at 100-m flight altitude above ground, which was compared to our LST fusion to evaluate spatial variability. UWKA geospatial coordinates were used to average all 100-m above ground flight leg LST observations overlapping each pixel.

Twenty eddy covariance flux towers were located in the 10 × 10 km inner domain. These towers were located in a range of ecosystems, including mixed forests, evergreen forests, wetlands, and grass fields. Seventeen of these had four-component net radiation measurements (Hukseflux NR01) available, from which upwelling and downwelling longwave radiation were extracted to calculate LST. Following Malakar et al. (2018), we estimated surface emissivity at 10.6 and 11.3 μm based on the ASTER satellite global emissivity database, provided at 30 m resolution (Hulley et al., 2015). Surface emissivity was averaged over a 90 × 90 m box around the center coordinate of each tower. Hourly averaged LST estimates for each tower were then used to compare to LST from the hourly fusion product.

Landsat 8-based LST was also acquired for this domain. Here, we acquired an enhanced LST product from Landsat based on the two-channel split window algorithm from Gerace et al. (2020), an improvement over the operational single-channel algorithm. Given the high cloud cover of most scenes during the intermittently and anomalously rainy CHEESEHEAD19 campaign, we focused on a single scene collected on September 26, 2019 as an evaluation LST, whereas ECOSTRESS was used for training given its more frequent repeat coverage. Landsat LST thermal resolution is 100 m, but the output is at 30 m by cubic convolution to match the visible bands. For a reliable comparison of LST, we resampled both images to 100 m after reprojecting the LST image to the high-resolution 50 m LST grid.

The NOAA uncrewed airborne system drone is a DJI S-1000 that was outfitted with a downward-pointing FLIR Tau 2 infrared camera during CHEESEHEAD19 as well as iMet-XQ sensors to sample temperature, moisture, and pressure in situ. The infrared camera has a 7.5-mm lens, 336 × 256 pixel resolution, and view angle of 90° × 69° (Dumas et al., 2016, 2017; Lee et al., 2017, 2019). The DJI S-1000 was flown in July at a single eddy covariance flux tower site hourly throughout the day, over an area of approximately 500 × 500 m, which was a distance sufficient to cover a significant number of pixels. We focus on data obtained during

the flights on 12 July; flights on the other days with the DJI S-1000 during the July campaign were smaller in radius and thus less useful for downscaling.

2.4. High-Resolution (50 m) LST Fusion

We apply a fusion approach based on regression over multiple colocated images. The approach is of the same class as the STARFM fusion approach for downscaling of MODIS to Landsat resolution, based on pixel level and neighborhood correlation (Gao et al., 2006, 2015; Yoo et al., 2020). However, instead of sharpening one coarse image based on the relationship of overlapping fine (Landsat) and coarse (MODIS) resolution images and the corresponding increment, as is done for STARFM; here, by virtue of having a large number of high-resolution images and the need to capture the diel cycle of LST, we instead applied an approach more appropriate for capturing diel time variation. The basics are provided below and evaluated in the results.

The first step was to gap-fill cloud-covered LST data in GOES, as indicated by the GOES cloud flag. To do so, we used the NLDAS LST estimate from each of the three models. The average and standard deviation are used here as a prior estimate of LST. For each GOES pixel, the relevant NLDAS pixel is geolocated using a nearest neighbor approach. A linear regression debias is then applied to the hourly NLDAS LST for the same hour on each day when GOES LST was observed, so that each 12.5 km NLDAS pixel would have ~40 independent regressions against each ~2 km GOES pixel. Regression was performed using the fitexy routine in the IDL Astronomy Library (Landsman, 1993), allowing the slope and intercept to account for errors in both the predictor variable X (uncertainty of GOES, nominally set to 1.5 K) and response variable in Y (the standard deviation of NLDAS LST across the three models). For the 153-day period, each GOES pixel had 24 separate regressions (one for each hour of the day) applied to the matching NLDAS pixel. This enabled us to debias the mean and variance of LST over the season, and also correct for differences in the magnitude of the diel cycle. Missing values of LST in GOES were then replaced with these debiased NLDAS values. This approach assumes that any bias between NLDAS and GOES is independent of cloud cover.

Next, ECOSTRESS was used to downscale the gap-free 2 km GOES image to a standard 50 m grid, using a simple pixel-level linear fusion model (Yoo et al., 2020), after finding that non-linear and spatial models are not an improvement in our case. Both were first re-projected into a standard UTM grid with 50×50 m square pixels. For each ECOSTRESS pixel (for most points, up to 25 observations over the 153-day period, generally equally distributed across all hours of day and night), we extracted all nearest GOES observations matching in space and time, within the closest hour of the ECOSTRESS overpass. For spatial alignment, GOES was resampled to our target 50 m resolution using bilinear interpolation. No threshold was applied for maximum difference in temperature. The 95% of the difference in the two were within -5.0 – 3.8 K. Linear slope and intercept were then calculated for each again using fitexy with the documented uncertainty of 1 K for ECOSTRESS. Slopes outside of the 98% confidence interval, where 98% of calculated slopes fell within 0.9–1.4, were rejected to prevent unreasonable LST extrapolations. For the missing slope values, a neighborhood smoothing algorithm was applied from nearby pixels, and the intercept was recalculated based on the regression of slope to intercept ($r = -0.24$). Unlike traditional fusion approaches, which might only use one fine resolution scene and one coarse resolution image to downscale a subsequent or preceding coarse resolution image, here, we are using all ECOSTRESS images and all matching GOES images to develop a single fit. This linear fit is then applied to the GOES gap-filled imagery to downscale the image to hourly, 50 m resolution LST.

2.5. Case Study of Higher Resolution (10 m) Downscaling

To evaluate whether additional covariates from high-resolution optical and near-to-shortwave infrared imagery could further sharpen our hourly 50 m LST, we developed an approach based on predicting sub-pixel LST variation based on hyperspectral airborne remote sensing. In this approach, a regression of spectral indices to LST anomalies is developed based on drone LST observations and then applied to the 50 m LST image to downscale it to 10 m. For this downscaling test to 10 m resolution, the additional covariates were brought from 1 m HySpex imagery based on known relationships among optical indices and fine-scale surface temperature variations (Yang et al., 2017). Data from three HySpex flight acquisition scenes (June

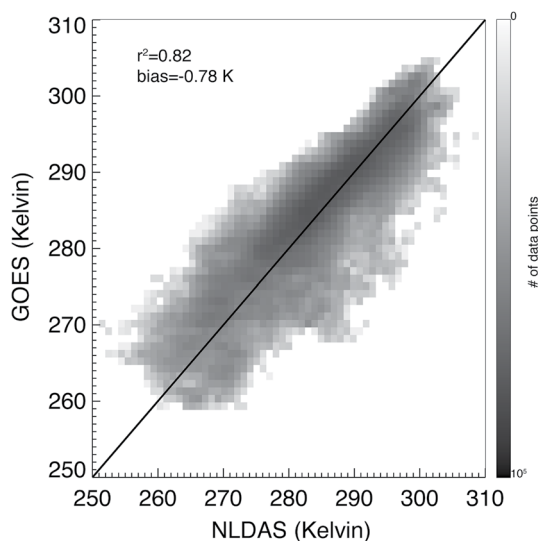


Figure 2. Strong correspondence of National Land Data Assimilation System (NLDAS)-modeled land surface temperature against observed geostationary satellite Geostationary Operational Environmental Satellites (GOES) temperature allowed for the filling of cloud gaps in GOES with NLDAS.

26, 2019, July 11, 2019, and August 8, 2019) were chosen to bracket the acquired drone LST image on July 12, 2019. The drone LST and HySpex imagery were upscaled to 10 m resolution using simple averaging and coaligned to a common grid. This had the benefit of increasing the signal-to-noise in the hyperspectral data. Following the approach by Dubois et al. (2018), analyses of the hyperspectral imagery utilized normalized difference spectral indices (NDSIs) with all two-band combinations of wavelengths:

$$\text{NDSI}(i,j) = (\text{Band}_i - \text{Band}_j) / (\text{Band}_i + \text{Band}_j) \quad (1)$$

Statistically, this enables the identification of key narrowband spectral features, while the use of ratios greatly decreases cross-track illumination effects related to sun-target-sensor geometry (i.e., bidirectional reflectance distribution function, BRDF). The downscaled 50 m LST from ECOSTRESS was subtracted from the upscaled 10 m drone LST to produce an LST anomaly map. This approach allows us in theory to build a more generalized model by assuming the higher resolution visible and near infrared reflectance provide information on how LST varies within the ECOSTRESS pixel due to variation in vegetation and surface properties. Each HySpex NDSI was separately regressed against the LST anomaly map. Band ratios with the largest average value of correlation r^2 across all three HySpex dates were then selected to develop a linear model to predict fine-scale LST from the anomaly map and three selected HySpex bands, after which multiple linear regression was used to develop coefficients that relate NDSI to LST anomalies.

We then constructed a higher-resolution 10 m LST by adding the ECOSTRESS-GOES based 50 m LST to anomalies predicted from applying the equation to the HySpex image closest in time (July 11, 2019). Using all of the 474 bands in HySpex allowed us to evaluate possible novel combinations of reflectance that could help explain variation in LST. The code that walks through the fusion methodology can be found at: <https://github.com/DesaiLab/LSTfusion>

3. Results

3.1. Cloud-Free Geostationary LST

After the diurnal regression step is applied, hourly NLDAS average LST corresponds well to retrieved GOES LST across the study domain (Figure 2). Overall 82% of the variation of GOES LST can be explained by NLDAS-modeled LST, though a small domain-wide cold bias of -0.78 K persists, with larger variance toward colder temperatures, potentially pointing to view-angle effects from shading or undetected clouds in GOES. The bias was more pronounced at night and near zero to slightly positive in daytime (range -1.87 – 0.7 K). Using the pixel-level, hour-specific regression, 60% of cloud-identified gaps were replaced with the modeled values. Missing observations were most prevalent late at night (9 UTC, 51% missing), during periods of fog or low-level stratus clouds, and a minimum in late afternoon (21 UTC, 25% missing), mostly during periods of extensive fair-weather cumulus cloud decks. Individual scenes had between 0% and 97% of pixels missing, averaging over 50% in early summer, during a particularly rainy period, to less than 40% during the normally drier autumn.

3.2. High-Resolution 50 m Fusion

Similar levels of correlation were found between the GOES gap-filled LST with retrieved ECOSTRESS LST (Figure 3, see Figure 1 for example of overlapping GOES and ECOSTRESS image), though with significant spatial variation. In general, within-pixel temporal correlation ($r = 0.59$ – 0.95) was stronger than across-pixel spatial correlation ($r = 0.32$ – 0.74). Low correlations were primarily found over water bodies, in particular the larger lake in the north of the domain, potentially from differences in retrieval algorithms or a documented cold bias on cooler surfaces in ECOSTRESS (Hulley et al., 2021).

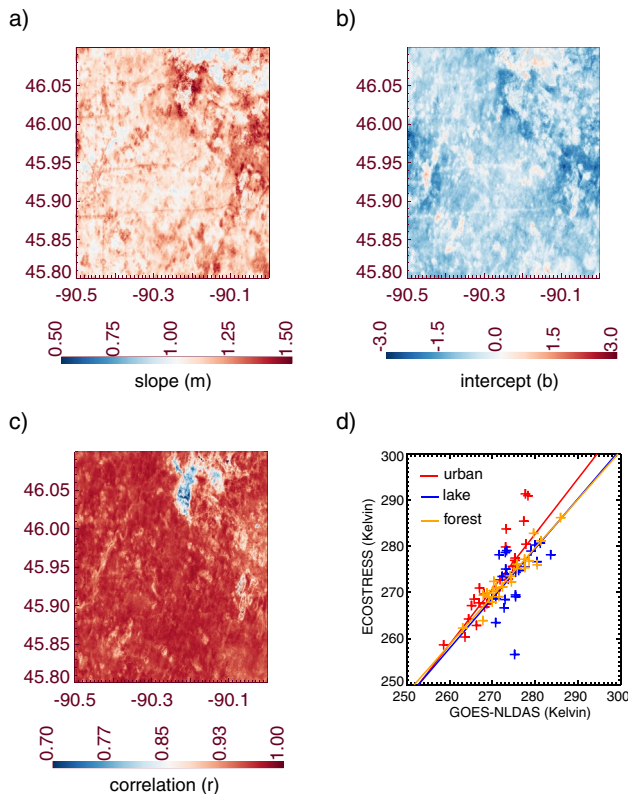


Figure 3. Pixel-by-pixel (a) linear slope, (b) intercept, (c) correlation, and (d) example regressions for three locations between 25 ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station images collected from January to October 2019 and gap-filled Geostationary Operational Environmental Satellites land surface temperature. Pixel-level temporal correlation ranges from 0.59 to 0.95 ($p < 0.01$) while individual image spatial correlation ranges from 0.32 to 0.74 ($p < 0.001$).

The regression of GOES to ECOSTRESS also varied in space, with median slope and intercept of 1.1 (interquartile range 1.06–1.17) and -0.87 K (interquartile range -0.60 to -1.18), respectively (Figures 3a and 3b). Particularly notable is the identification of urban areas (the City of Park Falls on western side) and highways. Slope and intercept were negatively correlated ($r = -0.24$). Locations with low slope (weaker diel and/or seasonal variation in ECOSTRESS compared to GOES) generally had higher (warmer) intercepts; for example, indicative of urban heat island or asphalt heat storage effects. Conversely, high slopes (stronger diel and/or seasonal variation) occurred in areas with lower (colder) intercepts, such as topographically low spots where cold-air pooling may depress mean temperature, including lakes, rivers, and bottomland forested areas.

These slopes and intercepts were applied to the GOES cloud-free LST to develop the downscaled high-resolution 50 m LST. Further evaluation of short time series of this product shows how the downscaled 50 m high-resolution LST better reflects differences in the diurnal cycle and means from the coarser resolution NLDAS or GOES, and closer to the tower-observed variations in LST, best resolved over wetlands (Figure 4a), but also reflecting good correspondence at resolving the warmer nighttime temperatures over forests (Figures 4a and 4b) and cooler daytime temperatures over lakes (Figure 4d).

3.3. Evaluation

We evaluate downscaled LST against several estimates of LST from tower, aircraft, and satellite. Spatial patterns were well captured in the high-resolution LST as compared to aircraft LST (Figure 5a, $r^2 = 0.75$) with small bias (-0.65 K) and an RMSE of 2.4 K. A larger >5 K bias is apparent in high LST locations, where the fusion product smooths out extremes given its linear averaging approach. In contrast, seasonal temporal variation biases were found to be more prevalent with colder temperatures, where the downscaled LST tended to underestimate the coldest temperatures observed by the towers (Figure 5b, $r^2 = 0.79$), though with better correlation.

No significant difference was found in the LST time series variation across land cover type, whether deciduous forest, evergreen forest, or wetland, with general RMSE of 3.5 K. Bias was larger than the airborne data at -2.6 K, especially later into the fall. Correlations within a land cover type were higher ($r^2 \sim 0.85$ – 0.88) than when pooled, as mean bias varied by land cover type. Wetlands had slightly larger bias (-3.2 K), RMSE (3.6 K), and lower correlation than the forested areas.

When the high-resolution 50 m LST was compared to independent satellite estimates of LST, a single 100 m Landsat LST scene generally revealed similar correspondence in the primary spatial pattern, but the overall correlation of the two products was much lower ($r^2 = 0.31$, Figure 6). The correlation was strongly influenced by the underestimation of higher LST range of the observation by our fusion product, over urban areas and a few forest clearings, implying an alternate weighting scheme may allow for better correspondence in those locations that are less prevalent in the area and thus under-represented in the calibration. As well, outside of those areas, variance of LST on this particular mid-day mid-summer scene is relatively small, within the RMSE shown for the high-resolution LST in the tower and aircraft comparison.

3.4. Higher-Resolution 10 m Downscaling

NDSI plots demonstrate a number of bands where visible and infrared band differences highly correlate with anomalies in sub-grid LST as observed by the drone (Figure 7). Most of the normalized differences with high correlation were on bands that were near each other, reflecting the role of specific spectral reflectance features of vegetation and soils that relate to LST variation. Here, we selected the top three

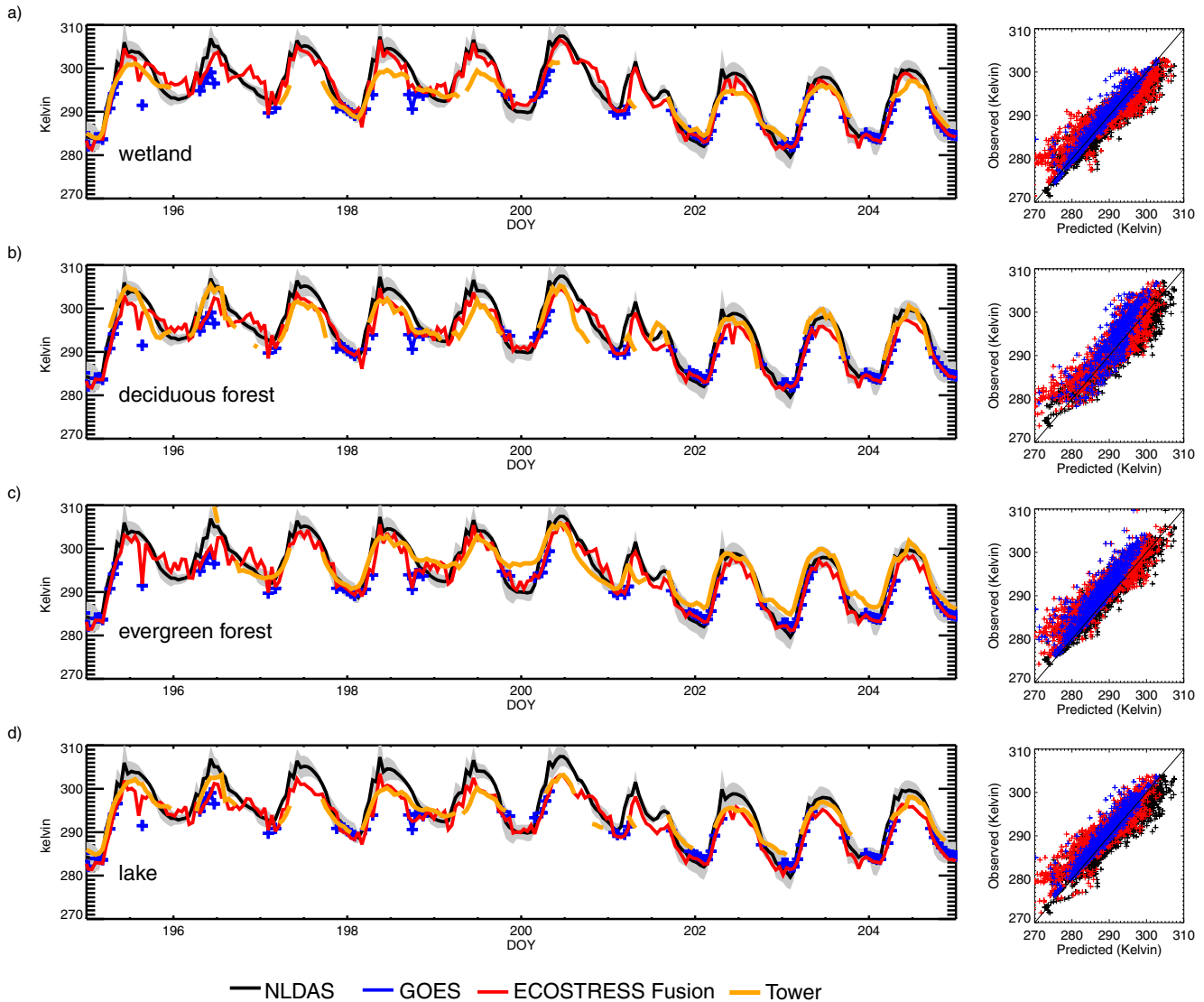


Figure 4. Comparison of land surface temperature diel cycle in (a) deciduous forest, (b) evergreen forest, (c) wetland, and (d) lake at four sites within ~2.5 km of each other from tower radiometric observations (orange), National Land Data Assimilation System (black line and gray shading representing spread in three models), Geostationary Operational Environmental Satellites (blue crosses, gaps indicate clouds), and the ECOSystem Spaceborne Thermal Radiometer Experiment on Space Station fusion product (red line) for mid-July, 2019.

consistent correlated NDSI band differences across the three flights. The three band pairs were: 1,982.7 nm and 1,470.5 nm in the shortwave infrared (NDSI_SWIR, $r = 0.380$), 709.1 and 760.2 nm in the red-edge spectral region (NDSI_EDGE, $r = 0.448$), and 651.6 and 504.7 nm at the red-green portions of the visible spectrum (NDSI_VIS, $r = 0.442$). In all three combinations, the second listed wavelength was subtracted from the first. The latter two are close to commonly used vegetation indices of NDVI (typically 630–690 nm and 760–900 nm) and the photochemical reflectance index (PRI, 531 and 571 nm), and correlations in those bands are not too far off from the selected bands (boxes in Figure 7), consistent to other studies linking LST and vegetation indices (e.g., Reynolds et al., 2008; Karnieli et al., 2010). The three bands are correlated with each other (r^2 range 0.30–0.61), with the NDSI_EDGE and NDSI_VIS most closely related. In addition, all three supported a more skillful model than any one or two of these bands, while additional bands did not appreciably change correlation or RMSE. Partial correlation coefficients show that each band contributed roughly equally in total correlation. With these three bands, a linear model was built to explain the sub-grid LST anomalies and applied to the downscaled LST_50m (Figure 8), expressed as:

$$\text{LST}_{10\text{m}} = 2.147 \times \text{NDSI_SWIR} + \text{NDSI_EDGE} + 5.143 \times \text{NDSI_VIS} + 4.566 + \text{LST}_{50\text{m}} \quad (2)$$

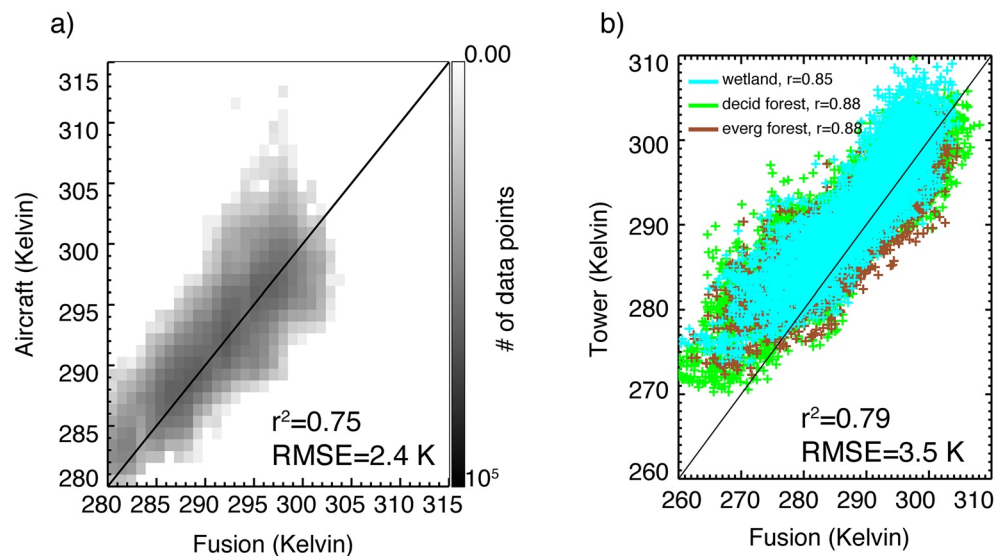


Figure 5. Comparison of fusion land surface temperature product against upwelling infrared temperature from (a) $\sim 10^5$ flight land surface temperature observations from the University of Wyoming King Air and (b) time series from 17 eddy covariance towers in deciduous forests (green), evergreen forests (brown), and wetlands or lakes (blue).

The resulting model produced a higher resolution 10 m LST map that was reasonably correlated to the drone imagery ($r^2 = 0.34$) and had significantly reduced bias compared to the high resolution 50 m LST from 2.35 K to near zero and lower bias-removed RMSE from 3.0 to 1.7 K. Most NDSI_SWIR values were positive (mean 0.12 ± 0.07), while NDSI_EDGE (-0.59 ± 0.10) and NDSI_VIS (-0.48 ± 0.09) were negative. Since all three coefficients were positive, the effect of positive SWIR was to increase LST, while for the mostly negative red-edge or visible reflectance was to decrease LST. The effect of these is to bring out key LST features, especially on the “hot-spot” side, such as a road and a larger open area, both of which were observed to have high LST but not well-detected in the original downsampled image. The higher NDSI_SWIR of these two features allowed this model to better capture its higher LST. The results paint a multistep pathway toward downscaling LST to meter scale resolution.

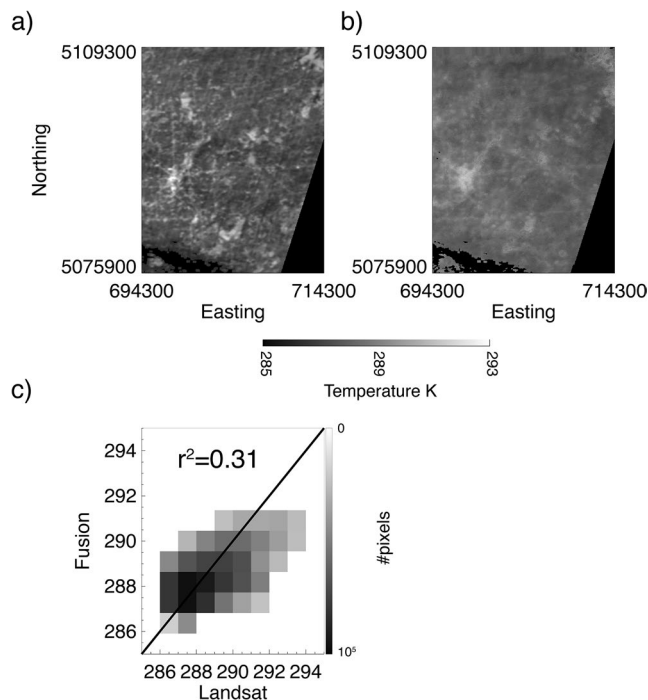


Figure 6. Comparison of (a) Landsat two-channel land surface temperature retrieval and (b) our 50 m fusion product reveals (c) relatively good correlation, though warmer areas in Landsat are underpredicted by the fusion.

4. Discussion

Land surface temperature exhibits high spatial and temporal variability. Depending on the application, capturing this variability can be essential for diagnosing land-atmosphere interactions, soil processes, and ecosystem thermal tolerances. Here, we demonstrated one approach to capture these scales of variations with multi-sensor fusion and find high skill in these when compared against independent LST observations. Both direct observations of LST and indirect observations of covariates provided information needed to downscale to hourly, 10 m resolution LST.

4.1. Challenges in LST Fusion

Our LST fusion approach performed well on evaluation, but several lingering uncertainties remain which require further investigation. The first involves the gap-filling of cloud cover. Previous satellite fusion investigations generally focused on clear-sky LST. The primary assumption made in our methodology is that the relatively strong linear relationship of pixel-level, hour-segregated NLDAS modeled LST, which does incorporate

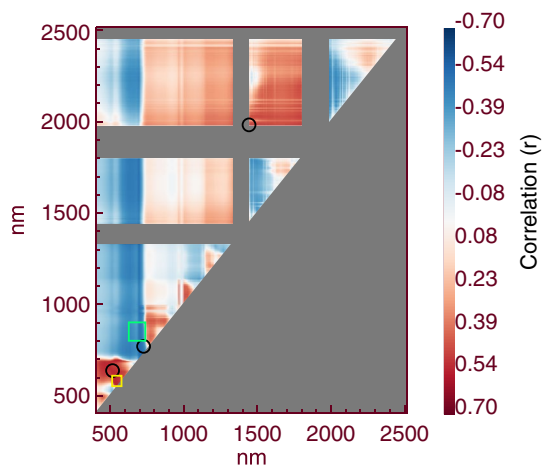


Figure 7. Normalized difference spectral index Pearson correlation (r) between band combinations from HySpex airborne hyperspectral imager and drone land surface temperature imagery. The top three most correlated band differences (1,470.466 and 1,982.657 nm, 709.077 and 760.173 nm, and 504.697 and 651.595 nm, noted in black circles) were used to construct a linear model for downscaling the fusion land surface temperature. Normalized difference vegetation indices region (green square) and photochemical reflectance index region (yellow box) noted for comparison.

the effect of clouds into its LST estimates (at least as reflected in the input model forcing), to the cloud-screened GOES is translatable to gap-filling. This approach assumes that LST during cloud cover is similar to LST during clear-sky conditions, given the same temperature for that time of day. Generally, the effect of clouds is to make LST cooler in daytime and warmer in nighttime compared to clear sky. An analysis of cloud cover (estimated as the ratio of observed shortwave radiation to potential maximum shortwave) versus difference in fusion to tower observed LST did not show any clear trend, suggesting this assumption is broadly reasonable.

Downscaling with ECOSTRESS and a linear model also brings additional uncertainty. The coverage of ECOSTRESS varies by the time of day and cloud cover, which means that each pixel had differing numbers of valid ECOSTRESS LST observations across the study period. Here, we assume no change in the seasonality of the relationship or temporal differences. Rather, we assume that what ECOSTRESS is mainly providing is differences in mean LST within the sub grid of a single GOES pixel (intercept) and changes in the diel amplitude (slope). However, this assumes that other biases are negligible, no changes occurred in land surface from disturbance or phenology, and seasonal variation in those two factors are zero. Subsetting by sub-season would be helpful here, but given the repeat interval and number of cloud-free images, statistical power would degrade noticeably. With a longer time period data set, additional subsetting may be warranted to evaluate such an approach. While some aspects

may have been better captured using a non-linear model, deviation from linear slopes across our ECOSTRESS and GOES pairs was rarely seen and initial tests with quadratic forms did not find improved fits. Data mining approaches, including data sharpening approaches, may improve performance (Gao et al., 2012).

Some of these performance issues show up when looking at the goodness of fit against towers and aircraft, and in the diel cycle plots. While the downscaling helps differentiate variation in LST by land cover type, it appears the methodology has challenges with a few land cover types. One is lakes, and especially lake-land edges, where pixel registration and gradients are missed leading to increased “noise” or blur in images around lakes. However, the visual inspection of geolocation errors did not find anything significantly skewed. The second is picking up cold LST values in autumn. The drone comparisons also suggest that the 70 m resolution of ECOSTRESS may still be challenging for picking up even finer-scale urban, road, or other hot spots on the landscape. Similar to seasonal subsetting, partitioning the regression by higher-resolution land cover observations may provide another approach to better reflect land cover variation.

View angle differences among the sensors may also contribute to differing error structures and biases that were not corrected in the provided Level 2 products used here (Anderson et al., 2021; Ermida et al., 2014; Gerace et al., 2020; Guillevic et al., 2013). Geostationary satellites in particular have strong angular effects as the sensor scans away from the central location, while ECOSTRESS has a $\pm 25^\circ$ acceptance swath, narrower than other polar orbiters. Surface skin temperature is also derived from different sets of wavelengths across the sensors, and biases from these may pose a challenge in addition to algorithmic differences in retrievals (Bosilovich et al., 2007). It is one reason we used mean bias removal in our regressions.

4.2. Mechanisms of LST Relationships to Visible to VSWIR Spectra

Though limited to a small number of images, our attempt to further downscale with visible to shortwave infrared hyperspectral imagery demonstrated improved ability to resolve fine-scale features such as roads and smaller wetlands observed in the drone imagery. The three band indices that contributed most to the NDSI regression represent key vegetation and soil features that link to LST variation. The strongest was in the shortwave IR, a region known to detect differences in soil thermal and moisture status. The other two in the visible and red-edge reflect signals of vegetation presence and photosynthetic activity, respectively. Actively photosynthesizing vegetation will have lower LST due to the cooling effect of concomitant transpiration

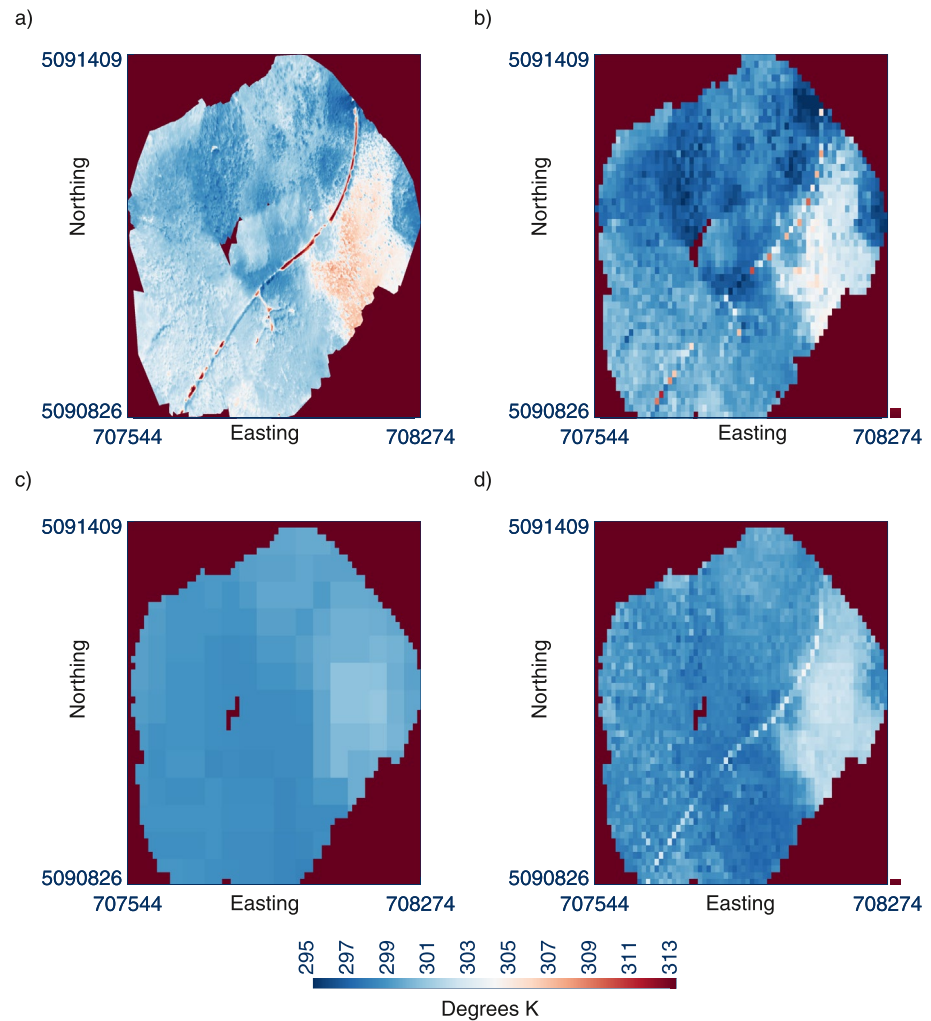


Figure 8. Comparison of (a) high-resolution (0.5 m) NOAA UAS drone land surface temperature on July 12, 2019 at 2230Z to (b) same image upscaled to 10m, (c) original fusion land surface temperature product (average of 22 and 23 Z), and (d) fusion product further downsampled with visible and near IR hyperspectral imagery collected on June 26, July 13, and Aug 6, 2019, demonstrating significant improvement in correlation ($r^2 = 0.14$ and $RMSE = 3.0$ K with 50 m and $r^2 = 0.34$ and $RMSE = 1.7$ K with 10 m imagery).

and given our formulation of NDSI, those areas had higher negative values in NDSI_EDGE and NDSI_VIS, which when combined with positive coefficients, led to lower LST over vegetated areas. The SWIR bands helped distinguish areas of exposed ground, and NDSI_SWIR was found to be most strongly positive over roads and open area. The broad areas of high correlation also partly overlap with commonly used band ratios, including NDVI and PRI, suggesting that broadband visible-IR remote sensing has strong potential for downscaling LST.

4.3. Comparison to Other Approaches

While several papers have assessed fusion approaches for gridded LST, literature on joint temporal and spatial LST downscaling is relatively limited, with primary applications over urban areas (e.g., Sismanidis et al., 2016a; Sismanidis et al., 2016b, 2018). Our results show that sub-daily temporal and sub-km spatial downscaling is possible while maintaining a similar level of uncertainty as previously published daily or less-frequent LST products (Freitas et al., 2010; Goettsche et al., 2013). Further, even without the additional spatialization from ECOSTRESS or HySpex, there is value in the greater use of geostationary satellite LST. Several of these satellites can now provide up to one-minute time resolution for target-mode operations,

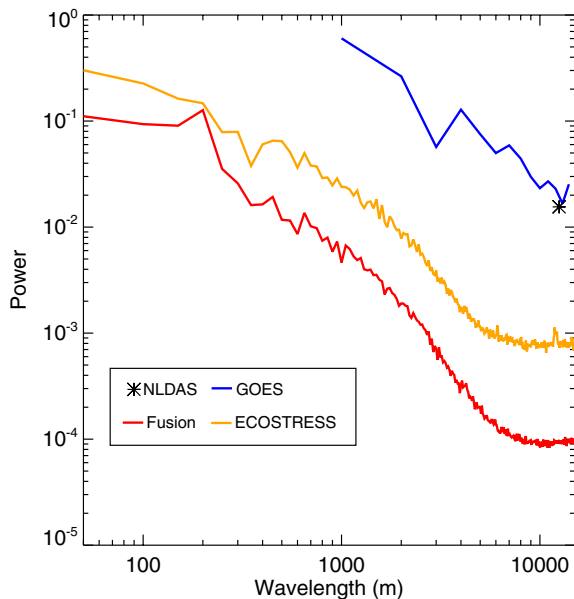


Figure 9. Radially integrated two-dimensional spatial power spectrum for a single clear sky day (June 16, 2019) compared among (a) GOES-NLDAS, (b) ECOsystem Spaceborne Thermal Radiometer Experiment on Space Station (ECOSTRESS), and (c) the fusion land surface temperature product. The fusion product shows better correspondence of spatial autocorrelation and structure to ECOSTRESS than Geostationary Operational Environmental Satellites (GOES).

covariance flux towers, flux footprint models, and input covariates (Metzger et al., 2013; Xu et al., 2017). For surface energy fluxes, such as sensible heat flux, LST is a key driver. Eddy fluxes of sensible heat during periods of high variability in wind direction reveal the presence of hot spots and hot moments of heat flux across space. The ERF methodology can identify those only if the input covariates are of sufficient spatial (decameter) and temporal (hourly or better) resolution to resolve those. While these flux hot spots

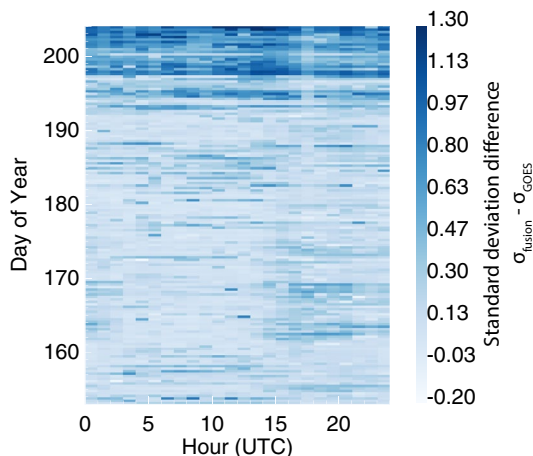


Figure 10. Difference in spatial standard deviation of land surface temperature (LST) between the 50 m fusion and Geostationary Operational Environmental Satellites as a function of hour of day (x-axis) and day of year (y-axis). Increasing heterogeneity in LST is found toward the autumn and afternoons in summer.

and fusion of these through a data assimilation approach would help develop global high-temporal resolution LST (Freitas et al., 2013; Xiao et al., 2021). Further work on using the higher frequency observations to reduce cloud coverage and increase the estimate of temperature variability could prove useful for developing more sub-daily LST-related products, including surface energy fluxes.

4.4. Applications of High Space and Time Resolution LST

There is a downside with temporal and spatial downscaling of LST, which is the increase in uncertainty as more products are fused and local calibrations fail to extrapolate well. The higher uncertainty does lead to the question of whether such an approach adds value. Beyond the aforementioned importance of fine space and time variation in LST for biological and geophysical processes, a number of studies suggests that higher resolution LST, even with greater uncertainty, aids in interpreting observations and testing hypotheses. We argue that the approach here demonstrates that relatively straightforward to implement linear fusion models can sufficiently capture much of the higher time and space variability in LST necessary for many science applications. Further, modifications of this approach, such as non-linear models, additional covariates, or seasonality (e.g., elevation for complex topography or vegetation phenology for crops) are relatively easy to incorporate.

As an example of the utility of high resolution LST, we present as an example its value to the Environmental Response Function (ERF) approach. ERF is a method to map surface-atmosphere fluxes of carbon, energy, and momentum across space and time from the fusion of eddy covariance flux towers, flux footprint models, and input covariates (Metzger et al., 2013; Xu et al., 2017). Previous ERF studies relied on linear regridding of coarser resolution LST products, decreasing the accuracy of hot spot localization (e.g., Xu et al., 2020). Thus, even at the acceptance of higher random uncertainty, a high space and time LST product is essential in this application. The variation in LST or difference of LST to air temperature is fit to an empirical model. Thus, it is the variation in LST that is guiding the methods, and accuracy is less important than spatial precision. In other cases, the magnitude of LST may be the driving factor, as is the case in models of evapotranspiration (Anderson et al., 2021; Guillevic et al., 2019) or atmospheric boundary-layer growth (Desai et al., 2006), in which case, the additional spatial information may be of less use, but the higher temporal information captures land-surface heat capacity and moisture holding impacts that influence the diel cycle of LST.

There are other cases where neither the variation nor magnitude matters, but rather the spatial structure. Consider Figure 9, where we depict the radially integrated spatial power spectrum of LST from GOES-NLDAS, ECOSTRESS, and the fused product. A number of fine scale modes of variation are present in the higher resolution products not found in GOES, which overestimates the autocorrelation. Similarly, when looking over time (Figure 10), the enhanced spatial resolution improves upon

GOES ability to detect increasing spatial variation of LST in autumn and during midday in summer. These patterns have been tied to generating heterogeneity in heat fluxes that promote mesoscale atmospheric circulations (Butterworth et al., 2021).

5. Conclusions

We demonstrated that a fusion of modeled land surface temperature with geostationary, irregular, and polar orbit observations and hyperspectral imagery provides a simple pathway for high space and time resolution LST for any region where those observations are available. LST estimates well captured many dynamics of spatial and temporal variation across a heterogeneous landscape of lakes, forests, wetlands, and urban areas in northern Wisconsin. Additional efforts should be placed on approaches to gap-filling for clouds, improvement of LST retrievals over water bodies and landscape transition edges, and multi-instrument evaluation. Our results suggest that continued effort to combine temporal and spatial estimates of LST can provide a fruitful path forward to better understand earth system processes, land surface data assimilation for modeling, and microclimate delineation.

Data Availability Statement

CHEESEHEAD19 observations are archived and described at the National Center for Atmospheric Research (NCAR) Earth Observing Lab (EOL) data repository: https://www.eol.ucar.edu/field_projects/cheesehead. Specific data include those from eddy covariance flux towers micrometeorology, archived at the Ameriflux website (US-PFa through US-PFt, <https://doi.org/10.17190/AMF/1717850>) and University of Wyoming aircraft (<https://doi.org/10.26023/5B70-4VP5-XY0V>), NOAA UAS drone LST (<https://doi.org/10.26023/F761-BMJ8-3E0C>), and UW HySpex hyperspectral imagery (<https://doi.org/10.26023/NMQG-NGWE-3400>). The output land surface temperature grids are available at the <https://doi.org/10.26023/5J4W-8XPH-250N>. ECOSTRESS observations can be obtained from: <https://ecostress.jpl.nasa.gov/data>. NLDAS model outputs are available at: <https://disc.sci.gsfc.nasa.gov/datasets?keywords=NLDAS>. GOES ABI LST data are at: <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C01521>. The LST fusion algorithm are available at: <https://doi.org/10.5281/zenodo.5279192>.

Acknowledgments

The authors thank the many data providers who helped extract observations including the National Center for Atmospheric Research (NCAR) Integrated Surface Flux Facility team, the University of Wyoming King Air team, and NOAA ATDD UAS team, including S. Oncley of NCAR and D. Plummer from U Wyoming. The CHEESEHEAD19 field study would not have been possible without the tremendous amount of work from dozens of colleagues. CHEESEHEAD19 was supported by NSF #1822420. The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under cooperative agreement by Battelle. This material is based in part upon work supported by the National Science Foundation through the NEON Program. GH contributed to this research at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. California Institute of Technology. Government sponsorship acknowledged. J.B.Fisher and G.Hulley were supported by NASA ECOSTRESS and PCS was supported by NSF #2034997.

References

- Adler-Golden, S. M., Matthew, M. W., Bernstein, L. S., Levine, R. Y., Berk, A., Richtsmeiera, S.C., et al. (1999). Atmospheric correction for shortwave spectral imagery based on MODTRAN4 International Society for Optics and Photonics. *Imaging Spectrometry V*, 3753, 61–69.
- Anderson, M., Yang, Y., Xue, J., Knipper, K. R., Yang, Y., Gao, F., et al. (2021). Interoperability of ECOSTRESS and Landsat for mapping evapotranspiration time series at sub-field scales. *Remote Sensing of Environment*, 252, 112189. <https://doi.org/10.1016/j.rse.2020.112189>
- Bosilovich, M. G., Radakovich, J. D., da Silva, A., Todling, R., & Verter, F. (2007). Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system. *Journal of the Meteorological Society of Japan*, 85A, 205–228. <https://doi.org/10.2151/jmsj.85a.205>
- Büttikofer, L., Anderson, K., Bebbler, D. P., Bennie, J. J., Early, R. I., & Maclean, I. M. D. (2020). The problem of scale in predicting biological responses to climate. *Global Change Biology*, 26, 6657–6666. <https://doi.org/10.1111/gcb.15358>
- Butterworth, B. J., Desai, A. R., Metzger, S., Townsend, P. A., Schwartz, M. D., Petty, G. W., et al., (2021). Connecting Land-Atmosphere Interactions to Surface Heterogeneity in CHEESEHEAD19, *Bulletin of the American Meteorological Society*, 102, E421–E445, <https://doi.org/10.1175/BAMS-D-19-0346.1>
- Chen, F., Mitchell, K., Schaake, J., Xue, Y., Pan, H.-L., Koren, V., et al. (1996). Modeling of land surface evaporation by four schemes and comparison with FIFE observations. *Journal of Geophysical Research*, 101(D3), 7251–7268. <https://doi.org/10.1029/95JD02165>
- Desai, A. R., Davis, K. J., Senff, C., Ismail, S., Browell, E. V., Stauffer, D. R., & Reen, B. P. (2006). A case study on the effects of heterogeneous soil moisture on mesoscale boundary layer structure in the southern Great Plains, USA. Part I: Simple prognostic model. *Boundary-Layer Meteorology*, 119(2), 195–238. <https://doi.org/10.1007/s10546-005-9024-6>
- Dirmeyer, P. A., Cash, B. A., Kinter, J. L., Jung, T., Marx, L., Stan, C., et al. (2012). Evidence for enhanced land-atmosphere feedback in a warming climate. *Journal of Hydrometeorology*, 13, 981–995. <https://doi.org/10.1175/jhm-d-11-0104.1>
- Dolman, A., Belward, A., Briggs, S., Dowell, M., Eggleston, S., Hill, K., et al. (2016). A post-Paris look at climate observations. *Nature Geoscience*, 9, 646. <https://doi.org/10.1038/ngeo2785>
- Dubois, S., Desai, A. R., Singh, A., Serbin, S., Goulden, M., Baldocchi, D., et al. (2018). Using imaging spectroscopy to detect variation in terrestrial ecosystem productivity across a water-stressed landscape. *Ecological Applications*, 28, 1313–1324. <https://doi.org/10.1002/eap.1733>
- Dumas, E. J., Lee, T. R., Buban, M., and Baker, B., (2016). Small unmanned aircraft system (sUAS) measurements during the 2016 Verifications of the Origins of Rotation in Tornadoes Experiment Southeast (VORTEX-SE), *NOAA technical memorandum*, OAR ARL-273.
- Dumas, E. J., Lee, T. R., Buban, M., and Baker, B., (2017). Small unmanned aircraft system (sUAS) measurements during the 2017 verifications of the origins of rotation in tornadoes experiment southeast (VORTEX-SE), *NOAA technical memorandum*, OAR ARL-274.

- Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., et al. (2016). The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geoscientific Model Development*, 9, 1647–1672. <https://doi.org/10.5194/gmd-9-1647-2016>
- English, S. J. (2008). The importance of accurate skin temperature in assimilation radiances from satellite sounding instruments. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 403–408. <https://doi.org/10.1109/TGRS.2007.902413>
- Ermida, S. L., Trigo, I. F., DaCamara, C. C., Göttsche, F. M., Olesen, F. S., & Hulley, G. (2014). Validation of remotely sensed surface temperature over an oakwood landscape-The problem of viewing and illumination geometries. *Remote Sensing of Environment*, 148, 16–27. <https://doi.org/10.1016/j.rse.2014.03.016>
- Fisher, J. B., Lee, B., Purdy, A. J., Halverson, G. H., Dohlen, M. B., Cawse-Nicholson, K., et al. (2020). ECOSTRESS: NASA's next generation Mission to measure evapotranspiration from the international Space Station. *Water Resources Research*, 56, e2019WR026058. <https://doi.org/10.1029/2019WR026058>
- Freitas, S. C., Trigo, I. F., Bioucas-Dias, J. M., & Goettsche, F.-M. (2010). Quantifying the uncertainty of land surface temperature retrievals from SEVIRI/Meteosat. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1). <https://doi.org/10.1109/TGRS.2009.2027697>
- Freitas, S. C., Trigo, I. F., Macedo, J., Barroso, C., Silva, R., & Perdigão, R. (2013). Land surface temperature from multiple geostationary satellites. *International Journal of Remote Sensing*, 34, 3051–3068. <https://doi.org/10.1080/01431161.2012.716925>
- Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., & Yang, Y. (2015). Fusing Landsat and MODIS Data for Vegetation Monitoring. *IEEE Geoscience and Remote Sensing Magazine*, 3, 47–60. <https://doi.org/10.1109/mgrs.2015.2434351>
- Gao, F., Kustas, W. P., & Anderson, M. C. (2012). A data mining approach for sharpening thermal satellite imagery over land. *Remote sensing*, 4, 3287–3319. <https://doi.org/10.3390/rs4113287>
- Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on Geoscience and Remote Sensing*, 44, 2207–2218. <https://doi.org/10.1109/tgrs.2006.872081>
- Gerace, A., Kleynhans, T., Eon, R., & Montanaro, M. (2020). Towards an Operational, Split Window-Derived Surface Temperature Product for the Thermal Infrared Sensors Onboard Landsat 8 and 9. *Remote Sensing*, 12, 224. <https://doi.org/10.3390/rs12020224>
- Gillespie, A., Rokugawa, S., Matsunaga, T., Cothorn, J. S., Hook, S., & Kahle, A. B. (1998). A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *IEEE Transactions on Geoscience and Remote Sensing*, 36, 1113–1126. <https://doi.org/10.1109/36.700995>
- Goettsche, F. M., Olesen, F. S., & Bork-Unkelbach, A. (2013). Validation of land surface temperature derived from MSG/SEVIRI with in situ measurements at Gobabeb, Namibia. *International Journal of Remote Sensing*, 34, 3069–3083. <https://doi.org/10.1080/01431161.2012.716539>
- Guillevic, P. C., Bork-Unkelbach, A., Goettsche, F. M., Hulley, G., Gastellu-Etchegorry, J.-P., Olesen, F. S., & Privette, J. L. (2013). Directional viewing effects on satellite land surface temperature products over sparse vegetation canopies—A multisensor analysis. *IEEE Geoscience and Remote Sensing Letters*, 10, 1464–1468. <https://doi.org/10.1109/LGRS.2013.2260319>
- Guillevic, P. C., Olioso, A., Hook, S. J., Fisher, J. B., Lagouarde, J. P., & Vermote, E. F. (2019). Impact of the revisit of thermal infrared remote sensing observations on evapotranspiration uncertainty—a sensitivity study using AmeriFlux data. *Remote Sens.*, 11, 573. <https://doi.org/10.3390/rs11050573>
- Hu, Y., Tani, R., Jiang, X., Li, Z. -L., Jiang, Y., & Liu, M. (2020). Spatial downscaling of land surface temperature based on surface energy balance, 2020 *IEEE international Geoscience and remote sensing symposium (IGARSS)* (pp. 4926–4929): USA. <https://doi.org/10.1109/IGARSS39084.2020.9323945>
- Hulley, G., Hook, S., Fisher, J., & Lee, C. (2017). ECOSTRESS, A NASA Earth-Ventures Instrument for studying links between the water cycle and plant health over the diurnal cycle, 2017. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 5494–5496, 8128248. <https://doi.org/10.1109/IGARSS>
- Hulley, G. C., & Ghent, D. (2019). *Taking the Temperature of the Earth: Steps towards integrated understanding of variability and change* (Vol. 2019, p. 256). Elsevier.
- Hulley, G. C., Goettsche, F., Rivera, G., Hook, S., Freepartner, R., Radocinski, R., et al. (2021). Validation and quality assessment of the ECOSTRESS level-2 land surface temperature and emissivity product. *IEEE Transactions on Geoscience and Remote Sensing*, 1–23. <https://doi.org/10.1109/tgrs.2021.3079879>
- Hulley, G. C., & Hook, S. J. (2011). Generating consistent land surface temperature and emissivity products between ASTER and MODIS data for earth science research. *IEEE Transactions on Geoscience and Remote Sensing*, 49, 1304–1315. <https://doi.org/10.1109/tgrs.2010.2063034>
- Hulley, G. C., Hook, S. J., Abbott, E., Malakar, N., Islam, T., & Abrams, M. (2015). The ASTER Global Emissivity Dataset (ASTER GED): Mapping Earth's emissivity at 100 meter spatial scale. *Geophysical Research Letters*, 42, 7966–7976. <https://doi.org/10.1002/2015GL065564>
- Karnieli, A., Agam, N., Pinker, R. T., Anderson, M., Imhoff, M. L., Gutman, G. G., et al. (2010). Use of NDVI and land surface temperature for drought assessment: Merits and limitations. *Journal of Climate*, 23(3), 618–633. <https://doi.org/10.1175/2009jcli2900.1>
- Khan, A. M., Stoy, P. C., Douglas, J. T., Anderson, M., Diak, G., Otkin, J. A., et al. (2021). Ongoing and emerging opportunities to improve environmental science using observations from the Advanced Baseline Imager on the Geostationary Operational Environmental Satellites. *Biogeosciences Discussions*, 18(13)4117–4141.
- Koster, R. D., & Suarez, M. J. (1992). Modeling the land surface boundary in climate models as a composite of independent vegetation stands. *Journal of Geophysical Research*, 97, 2697–2715. <https://doi.org/10.1029/91JD01696>
- Kröniger, K., Katul, G. G., De Roo, F., Brugger, P., & Mauder, M. (2019). Aerodynamic resistance parameterization for heterogeneous surfaces using a covariance function approach in spectral space. *Journal of the Atmospheric Sciences*, 76, 3191–3209. <https://doi.org/10.1175/jas-d-18-0150.1>
- Landsman, W. B. (1993). *Astronomical data analysis software and systems II, A.S.P. Conference Series*, In R. J. Hanisch, R. J. V. Brissenden, & J. Barnes (Eds.), (Vol. 52, p. 246).
- Lee, T. R., Buban, M., Dumas, E., & Baker, C. B. (2017). A new technique to estimate sensible heat fluxes around micrometeorological towers using small unmanned aircraft systems. *Journal of Atmospheric and Oceanic Technology*, 34, 2103–2112. <https://doi.org/10.1175/JTECH-D-17-0065.1>
- Lee, T. R., Buban, M., Dumas, E., & Baker, C. B. (2019). On the use of rotary-wing aircraft to sample near-surface thermodynamic fields: Results from recent field campaigns. *Sensors*, 19, 10. <https://doi.org/10.3390/s19010010>
- Li, Z., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., et al. (2013). Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, 131, 14–37. <https://doi.org/10.1016/j.rse.2012.12.008>
- Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99, 14415–14428. <https://doi.org/10.1029/94JD00483>

- Liu, N., Chlus, A., & Townsend, P. A. (2019). HyToolsPro: An open source package for pre-processing airborne hyperspectral images. In *2019 fall meeting AGU*, pp. GC51E–1120 <https://github.com/EnSpec/hypro>
- Liu, Y., Hiyama, T., & Yamaguchi, Y. (2006). Scaling of land surface temperature using satellite data: A case examination on ASTER and MODIS products over a heterogeneous terrain area. *Remote Sensing of Environment*, 105, 115–128. <https://doi.org/10.1016/j.rse.2006.06.012>
- Malakar, N. K., Hulley, G. C., Laraby, K., Co, Hook, S., & Schott, J. (2018). An operational land surface temperature product for landsat thermal data: Methodology and validation. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10), 5715–5735. <https://doi.org/10.1109/tgrs.2018.2824828>
- Matricardi, M., (2008). The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database. In *ECMWF research dept. Tech. Memo*.
- Metzger, S. (2018). Surface-atmosphere exchange in a box: Making the control volume a suitable representation for in-situ observations. *Agricultural and Forest Meteorology*, 255, 68–80. <https://doi.org/10.1016/j.agrformet.2017.08.037>
- Metzger, S., Junkermann, W., Mauder, M., Butterbach-Bahl, K., Trancón y Widemann, B., Neidl, F., et al. (2013). Spatially explicit regionalization of airborne flux measurements using environmental response functions. *Biogeosciences*, 10, 2193–2217. <https://doi.org/10.5194/bg-10-2193-2013>
- Pincebourde, S., & Salle, A. (2020). On the importance of getting fine-scale temperature records near any surface. *Global Change Biology*, 26, 6025–6027. <https://doi.org/10.1111/gcb.15210>
- Raynolds, M. K., Comiso, J. C., Walker, D. A., & Verbyla, D. (2008). Relationship between satellite-derived land surface temperatures, arctic vegetation types, and NDVI. *Remote Sensing of Environment*, 112, 1884–1894. <https://doi.org/10.1016/j.rse.2007.09.008>
- Saunders, R., Matricardi, M., & Brunel, P. (1999). An improved fast radiative transfer model for assimilation of satellite radiance observations. *Quarterly Journal of the Royal Meteorological Society*, 125, 1407–1425. <https://doi.org/10.1002/qj.1999.49712555615>
- Scarino, B., Minnis, P., Palikonda, R., Reichle, R. H., Morstad, D., Yost, C., et al. (2013). Retrieving clear-sky surface skin temperature for numerical weather prediction applications from geostationary satellite data. *Remote sensing*, 5, 342–366. <https://doi.org/10.3390/rs5010342>
- Schläpfer, D., & Richter, R. (2002). Geo-atmospheric processing of airborne imaging spectrometry data. Part 1: Parametric orthorectification. *International Journal of Remote Sensing*, 23, 2609–2630. <https://doi.org/10.1080/01431160110115825>
- Schmit, T. J., Griffith, P., Gunshor, M. M., Daniels, J. M., Goodman, S. J., & Lebair, W. J. (2017). A closer look at the ABI on the GOES-R series. *Bulletin of the American Meteorological Society*, 98, 681–698. <https://doi.org/10.1175/BAMS-D-15-00230.1>
- Sismanidis, P., Bechtel, B., Keramitsoglou, I., & Kiranoudis, C. T. (2018). Mapping the spatiotemporal dynamics of Europe's land surface temperatures. *IEEE Geoscience and Remote Sensing Letters*, 15, 202–206. <https://doi.org/10.1109/LGRS.2017.2779829>
- Sismanidis, P., Keramitsoglou, I., Bechtel, B., & Kiranoudis, C. (2016a). Assessing the capability of a downscaled urban land surface temperature time series to reproduce the spatiotemporal features of the original data. *Remote Sensing*, 8, 274. <https://doi.org/10.3390/rs8040274>
- Sismanidis, P., Keramitsoglou, I., Bechtel, B., & Kiranoudis, C. (2016b). Improving the downscaling of diurnal land surface temperatures using the annual cycle parameters as disaggregation kernels. *Remote Sensing*, 9, 23. <https://doi.org/10.3390/rs9010023>
- Taylor, C. M., de Jeu, R. A. M., Guichard, F., Harris, P. P., & Dorigo, W. A. (2012). Afternoon rain more likely over drier soils. *Nature*, 489, 423–426. <https://doi.org/10.1038/nature11377>
- Tomlinson, C. J., Chapman, L., Thornes, J. E., & Baker, C. (2011). Remote sensing land surface temperature for meteorology and climatology: A review. *Meteorological Applications*, 18, 296–306. <https://doi.org/10.1002/met.287>
- Wang, A., Barlage, M., Zeng, X., & Draper, C. S. (2014). Comparison of land skin temperature from a land model, remote sensing and in situ measurement. *Journal of Geophysical Research*, 119, 3093–3106. <https://doi.org/10.1002/2013JD021026>
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 NLDAS-2.1. Intercomparison and application of model products. *J. Geophys. Res.*, 117, D03109. <https://doi.org/10.1029/2011JD016048>
- Xiao, J., Fisher, J. B., Hashimoto, H., Ichii, K., & Parazoo, N. C. (2021). Emerging satellite observations for diurnal cycling of ecosystem processes. *Nature Plants*, 7, 877–887. <https://doi.org/10.1038/s41477-021-00952-8>
- Xu, K., Metzger, S., & Desai, A. R. (2017). Upscaling tower-observed turbulent exchange at fine spatio-temporal resolution using environmental response functions. *Agricultural and Forest Meteorology*, 232, 10–22. <https://doi.org/10.1016/j.agrformet.2016.07.019>
- Xu, K., Metzger, S., & Desai, A. R. (2018). Surface-atmosphere exchange in a box: Space-time resolved storage and net vertical fluxes from tower-based eddy covariance. *Agricultural and Forest Meteorology*, 255, 81–91. <https://doi.org/10.1016/j.agrformet.2017.10.011>
- Xu, K., Sühling, M., Metzger, S., Durden, D., & Desai, A. R. (2020). Can data mining help eddy-covariance see the landscape? A large-eddy simulation study. *Boundary-Layer Meteorology*, 176, 85–103. <https://doi.org/10.1007/s10546-020-00513-0>
- Xu, T., Bateni, S. M., Neale, C. M. U., Auligne, T., & Liu, S. (2018). Estimation of turbulent heat fluxes by assimilation of land surface temperature observations from GOES satellites into an ensemble Kalman smoother framework. *Journal of Geophysical Research: Atmospheres*, 123, 2409–2423. <https://doi.org/10.1002/2017JD027732>
- Yang, Y., Li, X., Pan Zhang, X. Y., Cao, C., & Cao, C. (2017). Downscaling land surface temperature in complex regions by using multiple scale factors with adaptive thresholds. *Sensors*, 17, 744. <https://doi.org/10.3390/s17040744>
- Yi, K., Smith, J., Jablonsk, A., Tatham, E., Scanlon, T., Lerda, M., et al. (2020). High heterogeneity in canopy temperature among co-occurring tree species in a temperate forest. *Journal of Geophysical Research-Biogeosciences*, 125, e2020JG005892. <https://doi.org/10.1029/2020JG005892>
- Yoo, C., Im, J., Park, S., & Cho, D. (2020). Spatial downscaling of MODIS land surface temperature: Recent research trends, challenges, and future directions. *Korean Journal of Remote Sensing*, 36, 609–626. <https://doi.org/10.7780/KJRS>
- Yu, Y., Tarpley, D., Privette, J. L., Flynn, L. E., Xu, H., Chen, M., et al. (2012). Validation of GOES-R satellite land surface temperature algorithm using SURFRAD ground measurements and statistical estimates of error properties. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 704–713. <https://doi.org/10.1109/TGRS.2011.2162338>
- Yu, Y., Tarpley, D., Privette, J. L., Goldberg, M. D., Rama Varma Raja, M. K., Vinnikov, K. Y., & Hui Xu, fm (2009). Developing algorithm for operational GOES-R land surface temperature product. *IEEE Transactions on Geoscience and Remote Sensing*, 47, 936–951. <https://doi.org/10.1109/TGRS.2008.2006180>

- Yue, J., Anderson, M. C., Gao, F., Hain, C., Sun, L., Yang, Y., et al. (2020). Sharpening ECOSTRESS and VIIRS land surface temperature using harmonized Landsat-Sentinel surface reflectances. *Remote Sensing of Environment*, 251, 112055. <https://doi.org/10.1016/j.rse.2020.112055>
- Zheng, W., Wei, H., Wang, Z., Zeng, X., Meng, J., Ek, M., et al. (2012). Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation. *Journal of Geophysical Research*, 117, D06117. <https://doi.org/10.1029/2011JD015901>