

Received 14 July 2022, accepted 3 August 2022, date of publication 16 August 2022, date of current version 29 August 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3198673



RESEARCH ARTICLE

Multisensor Machine Learning to Retrieve High Spatiotemporal Resolution Land Surface Temperature

KATE DUFFY^{ID}, THOMAS J. VANDAL, AND RAMAKRISHNA R. NEMANI

BAER Institute, Mountain View, CA 94043, USA
NASA Ames Research Center, Mountain View, CA 94043, USA

Corresponding author: Kate Duffy (duffy.m.kate@gmail.com)

This work was supported by the NASA Earth eXchange (NEX).

ABSTRACT Climate change is making heat waves more frequent, long-lasting, and severe. While multiple satellite types provide data to monitor surface temperature, geostationary (GEO) sensors provide near-continuous, continental-scale observations which can better capture the diurnal variability of land surface temperature (LST) than intermittent observations from low-earth orbit (LEO) sensors. However, standard products from GEO satellites are available at coarsened spatial and temporal resolutions compared to the native sensor resolution. Using datasets from the NASA Earth Exchange, we leveraged co-located, co-temporal observations from LEO and GEO satellites to learn a data-driven mapping using a convolutional neural network. The resulting NASA Earth eXchange Artificial Intelligence LST (NEXAI-LST) achieved a mean absolute error of 1.73 K relative to the target LEO product and improves on both spatial and temporal resolution [2 km, 10 minute] compared to the GEO full disk standard product [10 km, hourly]. In validation against measurements from a ground-based sensor network, NEXAI-LST achieves similar or better fit than both LEO and GEO standard products, while depending none of the prior knowledge of land surface and atmospheric states required by physical-statistical models. Further, application of the model to unseen LEO and GEO satellites demonstrates robust generalization of the model across spatial region, time of day, and sensor. In support of NASA's open-source science initiative, we make our NEXAI-LST product, model, and codes available to facilitate data exploration and further studies.

INDEX TERMS Datasets, deep learning, emulation, remote sensing, land surface temperature.

I. INTRODUCTION

As the world warms, patterns of extreme heat events are intensifying around the globe. In addition to raising average temperatures, greenhouse gases have been associated with strengthening the patterns of atmospheric circulation associated with heat waves [1], [2]. These patterns suggest potential for increasingly common, severe, and long-lasting heat events. In the United States, extreme heat is a leading cause of weather-related deaths [3], with economically disadvantaged census tracts and people of color bearing disproportionate exposure to heat in cities [4]. Heat events exacerbate cardiovascular and respiratory illnesses [5], [6],

The associate editor coordinating the review of this manuscript and approving it for publication was Stefania Bonafoni^{ID}.

impact birth outcomes [7], and burden health and emergency services. Surface urban heat islands (SUHI) describe the effect of elevated land surface temperature (LST) in urban areas compared to surrounding rural areas [8], [9]. While monitoring capabilities on the ground are limited, satellite data can be used to identify SUHI with coverage all over the world. SUHI mapping enables assessment of trends over time as well as near real-time monitoring of conditions that are relevant to human health and comfort. Recent heat events in the Pacific Northwest (2021), Western United States (2020), and Australia (2019) have drawn attention to the deadly lack of resilience and the need for data to increase preparedness for climate extremes.

From a climate perspective, LST is critical for the study of land surface processes and to constrain the surface energy

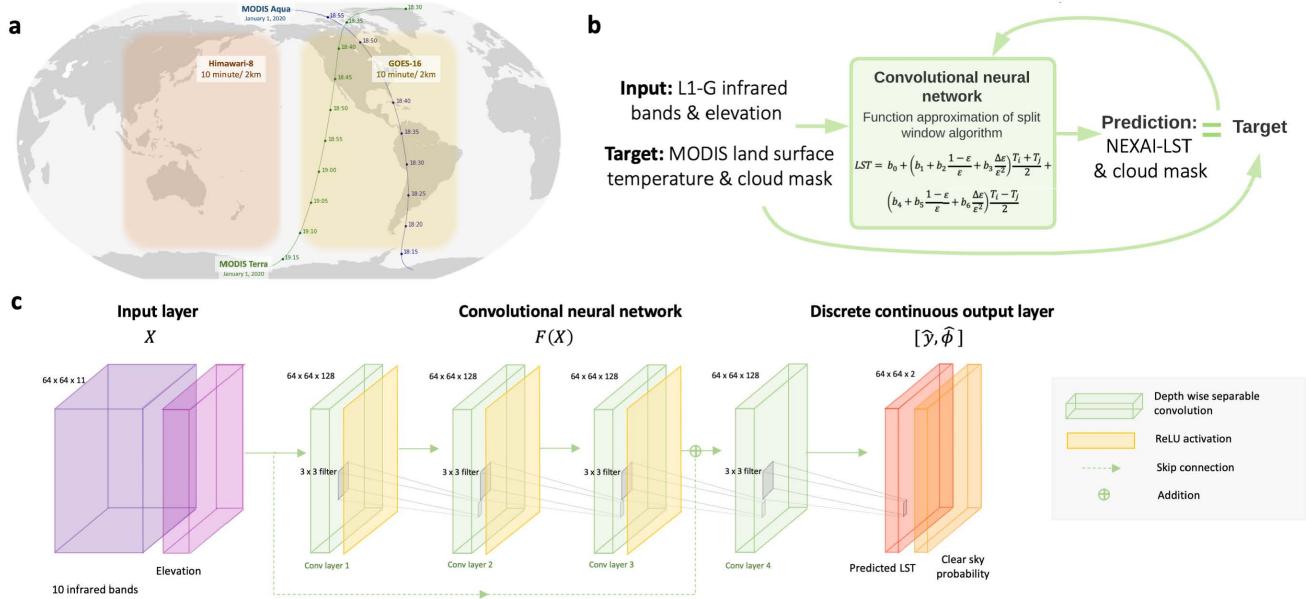


FIGURE 1. a, Geostationary (GEO) satellites like GOES-16 and Himawari-8 provide observations with unprecedented spatial and temporal consistency over their full disk views. Low earth orbiting (LEO) satellites produce a multitude of higher-level products that coincide with GEO data several times per day at each location. b, Our proposed framework leverages co-located, co-temporal observations from LEO and GEO satellites to build models for data-driven algorithm emulation. c, We trained a convolutional neural network to predict NEXAI-LST and classify clear sky pixels based on GEO infrared bands and elevation.

budget and model parameters. LST provides information about energy partitioning into sensible and latent heat fluxes that makes it a key parameter to study land-atmosphere interactions. LST helps us to understand the effects of climate on vegetation [10], hydrology [11], [12], and ecosystems [13] due to its close association with evapotranspiration. As surface temperatures can change over the course of minutes, land surfaces exhibit diurnal temperature fluctuations which are far more information-rich than individual observations. Diurnal cycles are primarily a lagged response to incoming and outgoing radiation, but there is significant spatial and seasonal heterogeneity in the timing of peak temperature and the range between daily minimum and maximum [9], [14]. The diurnal range of LST is a key, though relatively poorly observed, indicator of climate change [15], [16].

Various approaches have been used to observe land surface temperature, which remains a challenging task due to its ill-posed nature and the heterogeneity of land surface types [17]. Ground based observations of LST are sparse and not generally adequate to study spatial variation. Observations from low earth orbit (LEO) satellites such as NASA's Moderate Resolution Imaging Spectroradiometer (MODIS) have accurate and comprehensively validated products. MODIS data products boast global coverage, relatively long record lengths, and high spatial resolution. However, LEO sensors like MODIS have limited temporal resolution, reflecting the intermittent revisits permitted by LEO orbit patterns. MODIS flies onboard Terra and Aqua, two companion satellites with offset 12-hour revisit times.

The infrequency of observations makes it challenging for LEO satellites to capture information about diurnal variation.

Traditionally deployed for weather observations, geostationary (GEO) satellites have emerging applications for science due to their spatial and spectral similarities to LEO sensors, including the requisite thermal bands to retrieve land surface temperature. By orbiting at a high altitude of 35,785 km, GEO sensors have an identical orbital period to the Earth and remain in fixed positions as viewed from the Earth's surface. As a result, GEO sensors like the Geostationary Operational Environmental Satellites (GOES)-16 and -17 can provide full disk observations every 10 minutes or more. GEO observations resolve diel surface changes and have increased robustness to cloud cover [18].

Motivated by the complementary attributes of separate satellite datasets, we introduce a framework for cross-sensor emulation, in which a model is trained on overlapping LEO-GEO observations. We propose NEXAI-LST (NASA Earth eXchnage Artificial Intelligence Land Surface Temperature), a convolutional neural network that learns a mapping between GOES-16's infrared bands and MODIS Terra LST to accurately predict geostationary LST with a built-in clear sky classifier (Figure 1). Key contributions of this work include:

- 1) Prediction of a high quality, spatially and temporally consistent LST at a higher resolution (10 minute/2 km) than the GOES-16 standard full disk product (1 hour/10 km) (Figure 2).
- 2) Elimination of the need for ancillary datasets in our machine learning-based approach.

TABLE 1. Comparison of NEXAI-LST and MODIS LST using metrics of absolute error. NEXAI-LST derived from GOES-16 and Himawari-8 are spatially and temporally coincident with MODIS Terra and Aqua, resulting in four sensor combinations. While GOES-16/Terra pairs were used for training, all data from Himawari-8 and Aqua were reserved for validation. Metrics of error on LST prediction include root mean square error (RMSE), mean absolute error (MAE) and bias. These statistics and clear sky classification accuracy are presented with 95% confidence intervals.

		RMSE (K)	90% CI	MAE (K)	90% CI	Bias (K)	90% CI	Clear sky accuracy (%)	90% CI
G16 NEXAI-LST	Terra	2.409	[2.112, 2.705]	1.733	[1.483, 1.983]	0.176	[-0.126, 0.478]	94.6	[93.2, 95.9]
	Aqua	2.612	[2.287, 2.938]	1.908	[1.639, 2.177]	0.629	[0.304, 0.953]	94.5	[93.1, 95.9]
H8 NEXAI-LST	Terra	3.077	[2.784, 3.369]	2.144	[1.955, 2.334]	-0.572	[-0.881, -0.264]	93.2	[91.0, 95.4]
	Aqua	2.718	[2.49, 2.946]	2.019	[1.843, 2.196]	0.069	[-0.204, 0.341]	93.1	[90.8, 95.5]

- 3) Demonstration of the generalizability of the model across unseen times of day and spatial regions using holdout LEO and GEO sensors.
- 4) Greater absolute and relative fit to in-situ measurements from the SURFRAD (SURface RADiation Budget) Network, compared to two standard LST products.

Our experimental results suggest that the learned representation is robust in generalization to unseen satellite sensors. Further, multiple metrics of fit to ground observations suggest superior performance in comparison to the National Oceanic and Atmospheric Administration (NOAA) GOES-16 standard product and the MODIS Daily LST product. This performance is achieved without prior knowledge of current land and atmosphere states, in contrast with traditional physical-statistical models. To the best of our knowledge, there is no other framework to learn data-driven land surface temperature models that leverages overlapping observations from multiple sensors. To promote future research and development, we have made the NEXAI-LST product, model, and codes publicly available for download at <https://data.nas.nasa.gov/geonex/geonexdata/ML/nexai-lst/> and on GitHub at <https://github.com/KateDuffy/LEO-GEO-landsurfacetemp>.

II. RELATED WORK

Traditional remote sensing methods for LST retrieval use thermal infrared measurements supplemented by data from weather models and other satellite sensors. The thermal infrared channels of satellite sensors receive a signal that is determined by the Earth's surface temperature, surface emissivity/reflectivity, atmospheric effects, and solar radiation. Land surface temperature is extracted from this signal using methods from two main families: single infrared channel and split infrared channel. The single channel equation uses modeling of radiation scattering and absorption through the atmosphere and requires extensive ancillary information about current atmospheric profiles. Split channel approaches use differential absorption in two channels to partially account for atmospheric and surface effects. The MODIS LST retrieval algorithm is based on the split channel approach using MODIS bands 31 and 32 (central wavelengths = 11 and 12 μm) [19]. Coefficients of the generalized split window algorithm depend on viewing zenith angle, atmospheric surface temperature and water vapor, and are

estimated from regression analysis of a radiative transfer model. The MODIS LST algorithm is reduced to a lookup table in the operational implementation. The lookup table is organized as a database by water vapor saturation, air temperature, and surface emissivity.

There are several challenges that make LST retrieval a difficult problem. Atmospheric water vapor is the main atmospheric contribution to the thermal infrared signal that reaches satellites. As a result of water vapor, the actual surface temperature is generally higher than the brightness temperature measured at the satellite. The relationship between radiance and temperature is nonlinear, making linear models, like the single and split channel methods, imprecise, especially for hot and wet atmospheric conditions [20]. This systematic error increases as a function of column water vapor, such that including water vapor information improves LST accuracy [21]. However, the need for prior knowledge of land and atmospheric states leads to the propagation of error into LST products. LST products are also subject to strong directional effects due to structure of surfaces, including trees, topography, and buildings. Thus, differences in view angle can introduce significant difference between products.

Deep learning has begun to demonstrate promise in traditionally challenging problems in the Earth sciences, from precipitation nowcasting [22] to climate downscaling [23], [24]. Convolutional neural networks (CNN) can automatically extract spatial features, and it has been suggested that the ability to apply nonlinear reasoning using spatial context is behind CNN's power in the geosciences [25]. The general function approximation ability of neural networks, including CNNs, has been applied to the emulation of physical models in scientific domains ranging from turbulent flow [26] to atmospheric radiative transfer modeling [27], [28], astrophysics, climate modeling, biogeochemistry, high energy density physics, and more [29]. In the context of remote sensing, machine learning has been applied to generate spatial datasets like tree cover maps [30], synthetic sensor spectra [31], and poverty maps [32].

In land surface temperature retrieval, machine learning has been used to address several longstanding challenges. A method using knowledge-driven deep learning has been developed to deduce LST retrieval mechanisms and reduce the need for acquiring prior knowledge [33]. This paper

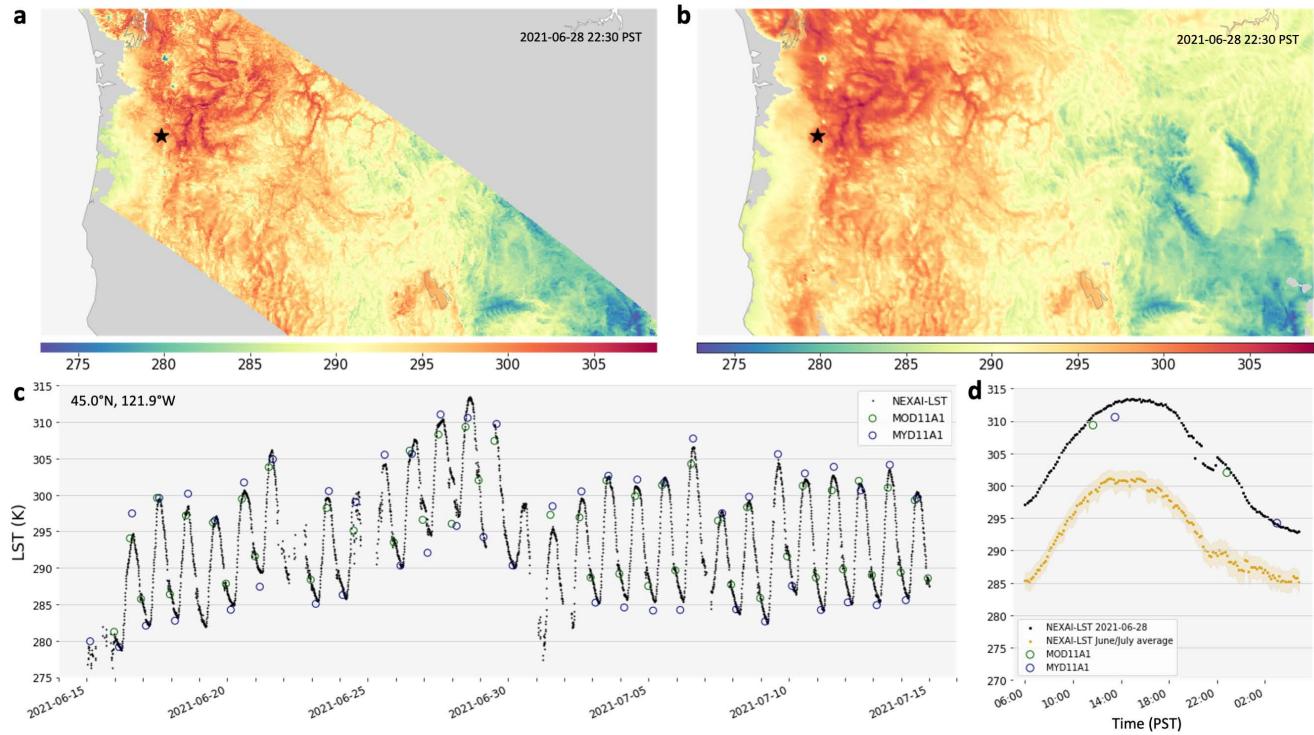


FIGURE 2. Summer 2021 brought a persistent, unprecedented heatwave to the Pacific Northwest. Land surface temperature across the region was produced by MODIS Terra (MOD11A1) (a) and predicted from GOES-16 by the ML model (b) at 22:30 PST on June 28, 2021. c, MODIS Terra and Aqua products, plotted with hollow green and blue markers, exhibit agreement with the model inferences but capture less detail about the temporal profile of LST. d, 10-minute LST on June 28, 2021 exceeds the June/July average for 2021.

demonstrated that in most cases the thermal infrared bands and water vapor data available from geostationary sensors are sufficient for retrieval requirements. Machine learning has also been used to fuse high-resolution optical imagery with thermal infrared data to estimate sub-pixel LST [34], to reconstruct missing data in LST [35], [36] and to retrieve emissivity and LST with better accuracy than the standard model [37].

III. MATERIALS AND METHODS

A. STUDY AREA AND DATASETS

A group of geostationary satellites including (GOES)-16 and -17 (NASA/NOAA) and Himawari-8 and -9 (Japanese Meteorological Agency) provides near-global coverage at a high temporal resolution. The Advanced Baseline Imager (ABI) on board GOES-16 and the Advanced Himawari Imager (AHI) on board Himawari-8 are similar sensors positioned on nearly opposite sides of the earth. At 75.2°W, GOES-16 observes the Americas, while Himawari-8, positioned at 140.7°E, covers east Asia and the western Pacific Ocean. ABI and AHI each scan the full disk, an area approximately 120° by 120°, every 10 minutes. The sensors have similar spectral range and 2 km spatial resolution in the infrared bands [38]. The Level 1 - Geostationary (L1-G) data were generated on a common grid by the Geostationary-NASA Earth eXchange (GeoNEX)

project [39]. The common geographic projection spans from 60°S to 60°N and from 180°W to 180°E.

MODIS is a flagship sensor on board NASA's Terra and Aqua satellites. Together, Terra and Aqua observe every location on Earth each 1-2 days, making observations in 36 discrete bands at 1 km or better resolution. We obtained MODIS Land Surface Temperature Daily L3 Global 1km (Terra: MOD11A and Aqua: MYD11A1) for the years 2019 and 2020.

The SURFRAD network was established in 1993 with the objective of providing accurate, continuous, long-term measurements of the surface radiation budget in the United States [40]. Seven stations covering diverse climates provide 1-minute measurements, including upwelling and downwelling infrared radiation. While measurements from networks like SURFRAD are too sparse to systematically validate remote sensing products, they provide high-quality, independent measurements that complement scientific efforts at comparing and benchmarking various LST products. We accessed SURFRAD measurements of upwelling and downwelling longwave flux corresponding to satellite observation times for the year 2020 at <http://gml.noaa.gov/grad/surfrad/>. SURFRAD LST was retrieved from the upwelling longwave flux measurement F_{LW}^{\uparrow} , downwelling longwave flux measurement F_{LW}^{\downarrow} , the Stefan-Boltzmann constant σ and the broadband longwave surface emissivity ϵ_{sf} .

TABLE 2. Summary statistics of the NEXAI-LST distributions and the coincident MODIS LST distributions. Distribution mean, variance, skew, and kurtosis are provided with 95% confidence intervals.

		Mean (K)	90% CI	Variance (K^2)	90% CI	Skew	90% CI	Kurtosis (%)	90% CI
GOES-16	Terra NEXAI	292.996 292.801	[290.969, 295.022] [290.777, 294.824]	99.698 93.161	[81.424, 117.971] [74.746, 111.576]	0.218 0.158	[0.123, 0.312] [0.061, 0.255]	-0.252 -0.354	[-0.385, -0.118] [-0.499, -0.209]
	Aqua NEXAI	293.536 292.953	[291.448, 295.624] [290.915, 294.991]	124.71 112.47	[103.536, 145.884] [91.938, 133.001]	0.301 0.192	[0.207, 0.396] [0.092, 0.291]	-0.429 -0.559	[-0.576, -0.282] [-0.705, -0.414]
	Terra NEXAI	293.751 294.507	[291.748, 295.755] [292.519, 296.494]	65.382 64.994	[46.102, 84.662] [45.17, 84.819]	0.120 0.051	[-0.014, 0.254] [-0.097, 0.199]	0.079 0.086	[-0.131, 0.288] [-0.164, 0.335]
	Aqua NEXAI	294.359 294.438	[292.231, 296.488] [292.316, 296.559]	81.66 73.17	[59.525, 103.796] [52.4, 93.94]	0.095 0.072	[-0.053, 0.242] [-0.094, 0.239]	-0.175 -0.083	[-0.379, 0.029] [-0.418, 0.251]

We assumed a value of 0.97 for the broadband longwave surface emissivity following the recommendation of previous studies [41], [42].

$$LST = \left(\frac{F_{LW}^{\uparrow} - (1 - \epsilon_{sfc})F_{LW}^{\downarrow}}{\epsilon_{sfc}\sigma} \right)^{\frac{1}{4}} \quad (1)$$

1) RATIONALE FOR THE CHOICE OF INPUT FEATURES

The predictors, which consist of 10 GOES-16 ABI infrared bands and 1 layer of elevation information, were selected based on prior indications of relevance to the task. Atmospheric water vapor (bands 7-10) is the main contributor to the attenuated thermal infrared signal that reaches satellite sensors [20], [21]. The cloud properties wavelength (band 11) is used in the MODIS LST algorithm, while the accuracy of retrievals also depends on ozone (band 12) [43]. Band 13 provides longwave infrared information that aids in identification of clouds and other atmospheric features. Bands 14 and 15 convey information about surface longwave radiation and are used with the split window technique in the GOES-R ABI LST standard product. In the case of GOES-16 data, the 10 infrared bands were also chosen for their availability at all times of day. The remaining 6 GOES-16 bands are visible or near-infrared and are available only during daytime. Elevation was also included as a covariate. Elevation influences the thermal environment by a negative correlation with LST [44].

2) CONSTRUCTION OF TRAINING DATABASE

We used geolocation information and observation times to match MODIS Terra/Aqua Daily LST with L1-G products generated by GeoNEX. MODIS LST was projected to the 2 km L1-G grid in Python. MODIS observation times were converted from local solar time to the Universal Coordinated Time system used by L1-G. Then, the nearest 10-minute geostationary observations were selected from the NEX database to obtain quasi-simultaneous pairs observed no more than 5 minutes apart. Both daytime and nighttime observations were used as infrared observations are available day and night. Tile matching resulted in over 30,000 MODIS Terra tiles paired with GOES-16 L1-G data from 2019. For validation and generalization experiments, pairs were also

constructed for the year 2020 with Terra/GOES-16, Aqua/GOES-16, Terra/Himawari-8, and Aqua/Himawari-8. Elevation data was prepared as an ancillary data source from the Shuttle Radar Topography Mission global digital elevation model (SRTM30) by projecting elevation to the 2 km L1-G grid.

B. MODEL ARCHITECTURE AND TRAINING

We present NEXAI-LST, a deep learning model for retrieval of land surface temperature using remotely sensed infrared bands and elevation data. We adopt a convolutional neural network architecture that is suitable to various image processing tasks and has been used extensively in modeling land surface temperature due to its capacity for leveraging spatial features [35], [36], [45]. The detailed architecture is shown in Figure 1c.

The convolutional neural network, $F(x)$, maps geostationary data, x , to land surface temperature, y , and cloud mask m . The 11 input channels are composed of 10 geostationary infrared bands and 1 layer of elevation information. The model consists of four convolutional layers, each with 128 filters of size 3 by 3, enclosed by one skip connection, the presence of which reformulates the task of the network as learning a residual function [46], or the difference between the output mapping $F(x)$ and input x . The first three convolutional layers are followed by Rectified Linear Unit (ReLU) activations, and the last by an identity mapping. The model excludes pooling layers, which is consistent with the task's dependence on relatively local information and helps to maintain temporal consistency in local features in the presence of global changes in the images. The relatively wide, shallow network was selected as a result of superior performance in a grid search evaluating various hyperparameters. While deeper model designs have become common in recent years, studies have found deep networks that cannot be realized by shallow networks [47], as well as wide networks which cannot be realized by any narrow network [48].

Following a previously published approach to handle discrete-continuous distributions in geoscience data, the model was conditioned to predict LST as one channel and clear sky probability, \hat{p} , as a second channel [23]. The corresponding terms of the loss function used least square errors

(L2) and binary cross entropy loss to compute the distance between the model output and the targets. For the network $F(x)$ with weights W , the loss function was written as following for $\{m_i\}_{i=1}^N$ the classification label, $\{p_i\}_{i=1}^N$ the predicted clear sky probability, $\{y_i\}_{i=1}^N$ land surface temperature label, $\{\hat{y}_i\}_{i=1}^N$ the predicted land surface temperature, N the number of pixels in the patch, and D the number of clear sky pixels in the patch:

$$\begin{aligned} L(W) = & \frac{1}{N} \sum_{i=1}^N [m_i \log(\hat{p}_i + (1 - m_i) \log(1 - \hat{p}_i))] \\ & + \frac{1}{D} \sum_{i, m_i > 0} \|y_i - \hat{y}_i\|^2 \end{aligned} \quad (2)$$

Data was divided into training (Terra/GOES-16 pairs from 2019) and validation (all pairs from 2020) sets. As training examples, 64 by 64 pixel patches were extracted from images with no overlap. Patches with no valid LST pixels (i.e. all cloud/water) were discarded with 66.67% probability. LST and geostationary bands were normalized to mean 0 standard deviation 1. During training, patches were rotated, flipped horizontally, and flipped vertically with 50% probability to expand the dataset. Gradient descent was handled using Adam optimization with $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1e - 7$ and a learning rate $1e - 4$ for 300k iterations on one NVIDIA V100 GPU on Pleiades, NASA Ames's high performance computing system.

IV. EXPERIMENTS

A. LST PREDICTION FROM GOES-16

After training the model on coincident GOES-16 data and MODIS Terra LST for the year 2019, we found that the model predicted LST for the year 2020 with a mean absolute error (MAE) of 1.733 K, 95% CI [1.483, 1.983] and root mean squared error (RMSE) of 2.409 K, 95% CI [2.112, 2.705] using equal spatial representation sampling (Table 1). A small positive bias of 0.176 K, 95% CI [-0.126, 0.478], relative to Terra LST contributed to the overall error. These metrics are relative to MODIS LST, which has been validated within $+/- 1$ K in clear sky conditions over various temperature ranges [49], [50], [51]. LST should only be retrieved under clear sky conditions to avoid conflating surface temperature with cloud top temperature. On the task of discriminating between clear sky pixels and cloudy/water pixels, the model achieved 94.6% accuracy, 95% CI [93.2, 95.9].

The model was trained on one full year of data, comprised of both night and day observations and a full cycle of seasons. Different seasons of the year are associated with dry and wet atmospheres that can affect retrieval errors. We evaluated the performance of LST retrieval and clear sky prediction over different seasons and found that MAE varied by 0.5 K over the course of the year, with higher error in spring and summer than in winter and fall.

A quantile-quantile (Q-Q) plot with a 45° reference line provides a direct pixel-wise comparison of the distributions

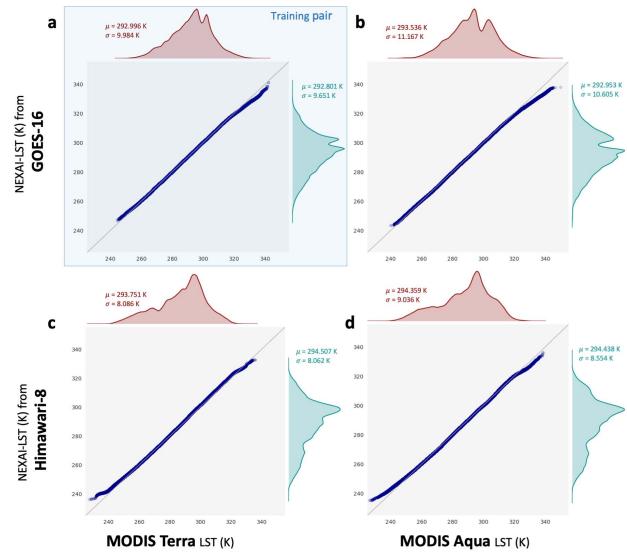


FIGURE 3. Quantile-quantile plots are presented for the two sensors used to train the model (GOES-16 and Terra; a), a holdout LEO satellite (GOES-16 and Aqua; b), a holdout GEO satellite (Himawari-8 and Terra; c), and two holdout satellites (Himawari-8 and Aqua; d).

with quantiles of the reference distribution (Terra LST) on the x-axis and quantiles of the NEXAI-LST distribution (from GOES-16) on the y-axis. (Figure 3). Points on the Q-Q plots formed a relatively straight line but show a somewhat lighter tail in the NEXAI-LST distribution than the reference distribution. This indicates that compared to the Terra LST distribution, the NEXAI-LST predictions had less data in the distribution extremes. However, NEXAI-LST closely reproduced the complex, multimodal, overall shape of the LST distribution. The distributions of Terra LST ($\mu = 292.996$ K, $\sigma^2 = 99.698$ K²) and NEXAI-LST from GOES-16 ($\mu = 292.801$ K, $\sigma^2 = 93.161$ K²) were statistically similar for corresponding times and locations (Table 2). The ML model also approximated the higher order statistics of the LST distribution. The ability of a model to capture observed skewness and kurtosis is scientifically important as these are potentially nonstationary aspects of climatology that are of interest for climate adaptation. Previous studies have indicated that the negative skewness of temperature is decreasing under climate change, resulting in fewer cold extremes and more hot extremes [52], [53], [54].

Patterns of bias in 10 quantiles of the distribution (Figure 4) indicate modest (<1K) biases between NEXAI-LST (GOES-16) and Terra LST when excluding the top decile. The ML model tends to overestimate temperature in the lowest (coolest) quantiles of the distribution while underestimating temperature in the uppermost (warmest) quantiles. This pattern is consistent with a thinner-tailed predicted distribution than the reference distributions from MODIS Terra. However, a strong linear relation ($R^2 = 0.977$) indicated that the regression model, which was trained on datasets from 2019, is well-fitted to the new observations.

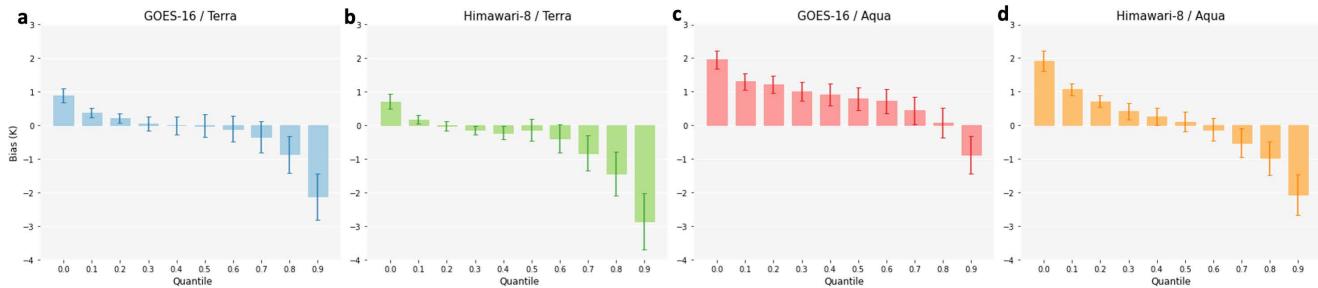


FIGURE 4. Biases are presented in ten quantiles that span mean and extreme land surface temperatures. For GOES-16 NEXAI-LST and MODIS Terra (a), Himawari-8 NEXAI-LST and MODIS Terra (b), GOES-16 NEXAI-LST and MODIS Aqua (c), and Himawari-8 NEXAI-LST and MODIS Aqua (d). Across the training and holdout sensors, biases are less than 1.5K in magnitude in all quantiles except for the top and bottom deciles. In all four applications of the model, the pattern of biases is consistent with thinner-tailed predicted than actual distributions.

TABLE 3. Summary of the ABI infrared bands, which each have a best spatial resolution of 2 km.

ABI Band	Central wavelength (μm)	Nickname
7	3.9	Shortwave window
8	6.2	Upper-level water vapor
9	6.9	Midlevel water vapor
10	7.3	Lower-level water vapor
11	8.4	Cloud-top phase
12	9.6	Ozone
13	10.3	"Clean" longwave window
14	11.2	Longwave window
15	12.3	"Dirty" longwave window
16	13.3	CO_2 longwave

B. GENERALIZATION TO HOLDOUT SENSORS

We used LST observations from Terra's companion satellite, Aqua, to test generalization of the model to times of day outside of the training dataset. Terra crosses the equator in a descending orbit (north to south) at 10:30am local time and an ascending orbit (south to north) at 10:30pm local time. Aqua crosses the equator in an ascending orbit at 1:30pm local time and a descending orbit at 1:30am local time. Thus, Aqua provides LST retrievals at times of day unseen in the training set. We also applied the trained model to predict land surface temperature from the infrared bands of the Advanced Himawari Imager on board the geostationary satellite Himawari-8. Himawari-8 observes Asia-Pacific every 10 minutes with similar bands to GOES-16's Advanced Baseline Imager.

Statistical performance was slightly poorer on the holdout sensors than the training sensors, with MAE in the range of 1.908 to 2.144 K and RMSE in the range of 2.612 to 3.077 K. NEXAI-LST values from both GOES-16 and Himawari-8 had a positive bias when compared to coincident MODIS Aqua LST, while NEXAI-LST values from Himawari-8 had a negative bias in reference to Terra LST. Accuracy of clear sky discrimination held constant on the holdout sensors at 93 to 94%.

Q-Q plots involving the two holdout sensors indicated similar distributions between MODIS LST and NEXAI-LST in terms of mean, variance, skew, and kurtosis. As seen for

the training sensor set, the NEXAI-LST distributions were slightly lighter-tailed than the MODIS distributions. R^2 values in the range of 0.940 to 0.979 indicated a strong linear association held between the labels and predictions regardless of distributional shifts in the holdout data with respect to the training data.

Patterns of bias in 10 quantiles of the distribution reflected qualitatively similar patterns of error across the combinations of holdout sensors. In general, bias became monotonically less positive between the lowest quantile and some middle quantile with near-zero bias, then monotonically more negative up to the highest quantile. The location of the transition point from positive to negative bias varies across combinations of sensors, which may be attributed in part to the difference in the Aqua and Terra distributions. As a rule, the biases were most pronounced in the tails of the distribution.

C. SPATIAL PATTERNS OF ERROR

The spatial patterns of error suggest that the sign as well as magnitude of NEXAI-LST errors exhibited clustering in space which may be related to spatial variation in land surface cover, atmospheric conditions, or other factors (Figure 5). In the region observed by GOES-16, higher magnitude errors clustered on the west coasts of North and South America. In the Himawari-8 region, the greatest magnitude of errors appeared in Central Asia, where a large (5 to 6 K) negative bias predominates. Elsewhere in the Himawari-8 region, biases were small and nearly uniformly positive.

The relatively few observations per geographic location in the tropics reflected the rarity of clear sky observations from MODIS in these regions. The robustness of geostationary sensors to cloud cover provides a motivation to utilize GEO sensors in the tropics in particular [18]. However, their underrepresentation in the training dataset does not appear to have resulted in poorer performance in terms of absolute error or bias. Rather, high errors and negative bias clustered in arid, high elevation regions including the Tibetan Plateau and Andes mountains, and the Rockies region in the western United States. Poorer generalization to the Tibetan Plateau region may be attributable to very low atmospheric

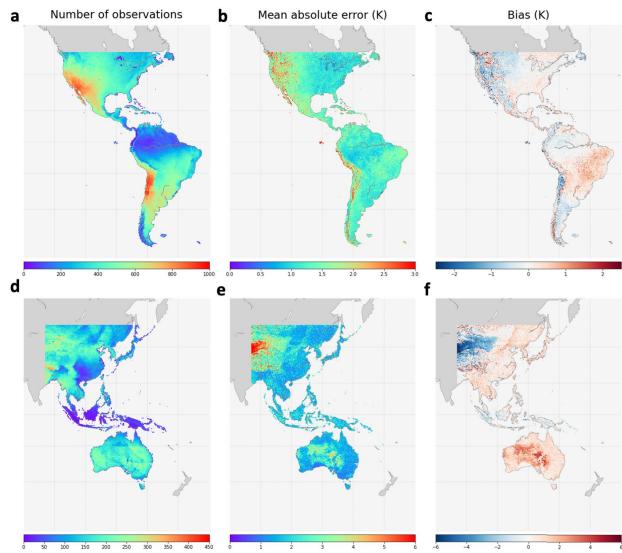


FIGURE 5. NEXAI-LST predicted from GOES-16 (a-c) and Himawari-8 (d-f) are evaluated by comparison to the MODIS Terra and Aqua LST products (predicted - MODIS). a,d, The number of matching LEO/GEO pixels shows significant spatial variation in the number of clear sky observations. For GOES-16 predictions, mean absolute error (b) is greatest in the arid, high elevation western regions, where bias is largely negative (c). For Himawari-8 predictions, a large negative bias also appears in the Tibetan plateau (e-f).

water vapor content over this arid, high elevation region [55]. As the main determinant of atmospheric transmittance, column water vapor has strong effects on retrieval error [56]. This error demonstrates the limited generalization of the ML model to physical conditions not represented in the training dataset, or could be attributable to exaggerated differences in pixel size and viewing angle caused by topography.

D. VALIDATION WITH GROUND MEASUREMENTS

We compared GOES-16 NEXAI-LST inferences and LST products from traditional physical-statistical models with ground-based measurements of LST from the seven SURFRAD stations. For benchmarking, NEXAI-LST was extracted at times corresponding to MODIS Terra/Aqua LST (Table 4 Panel a) and NOAA GOES-16 standard full-disk LST (Table 4 Panel b). The NEXAI-LST product exhibited greater agreement with SURFRAD measurements than either product in terms of both absolute and relative fit. Relatively poor performance at station DRA in Desert Rock, Nevada appeared across methods. Higher error at DRA likely occurred due to heterogeneous surroundings or a negative dewpoint bias associated with low moisture over the site [42].

In comparison to the MODIS LST products, NEXAI-LST has a lower MAE at 5 out of 7 stations, a lower RMSE at 6 out of 7 stations, and a higher R^2 value with SURFRAD for all 7 stations. The ML model achieves a reduction of the RMSE by nearly 0.5 K compared to the MODIS LST product used for training. Where viewing angle differences and atmospheric path length differences may have affected

the training pairs, these validation results suggest that the model has learned to overcome these biases and noise.

In comparison to the NOAA standard product, NEXAI-LST has a higher R^2 value with SURFRAD for 5 out of the 7 stations. In a mean across measurement stations, NEXAI-LST also has lower MAE and RMSE with respect to SURFRAD. The NOAA LST product is derived using a split window technique from the same Level 1 data used to predict NEXAI-LST. However, the split window algorithm additionally uses ancillary information including atmospheric total precipitable water, land surface emissivity, and snow/ice mask, and a cloud mask. While results based on ground measurements with limited spatial representativeness should not be over interpreted, they suggest that our model may perform as well as or better than the standard model while requiring no ancillary data as inputs.

V. OBSERVING HEAT EVENTS USING NEXAI-LST

Over the summer of 2021, heat waves in the United States Pacific Northwest produced multiple record-setting temperatures affecting ill-adapted population hubs. We applied our model, which can generate full disk NEXAI-LST with a frequency up to every 10 minutes, to predict LST for the region over the period spanning from June 15 to July 15, 2021. In Figure 2a-b, MODIS Terra LST and NEXAI-LST are presented side by side for observations taking place at 20:30 PST on July 28, 2021. A time series of NEXAI-LST at a 10-minute time step demonstrates the temporal profile of diurnal temperature variation over the course of days (Figure 2c). MODIS Terra and Aqua overpasses, plotted in green and blue, exhibit agreement with the model inferences. At the same time, their temporal sampling is insufficient to describe key characteristics of LST such as the intensity and timing of peak temperature, the minimum temperature, and the time-integrated exceedance of temperature thresholds. Terra and Aqua products are not guaranteed to capture the daily minima and maxima that are necessary to establish diurnal temperature range, including on the hotter-than-average June 28, 2021 2d.

Temporal consistency, or the uniformity of predictions for similar conditions close together in time, is a key indicator of model reliability. Achieving temporal consistency can be a challenge in ML approaches, such as this one, which process images frame by frame. However, the correlation of the NEXAI-LST predictions over time demonstrates the temporal stability of this model.

VI. DISCUSSION

The GOES-R standard land surface temperature product aims to meet an accuracy goal of 2.5 K conditional with 1) known emissivity, 2) known atmospheric correction and 3) 80% channel correction, and 5 K otherwise [57]. NEXAI-LST, which did not take direct advantage of any of these conditions, meets the 2.5 K goal for the set of training sensors and exceeds the 5 K goal for all combinations of sensors, while improving on both spatial and temporal resolution at the full disk scale. However, our model does depend on the

TABLE 4. Comparison of GOES-16 NEXAI-LST and coincident LST from physical-statistical models (MODIS Terra/Aqua, panel a; GOES-16 standard product, panel b) against ground station observations from the SURFRAD network from the year 2020.

(a) NEXAI-LST and MODIS Terra/Aqua LST						
Station	MAE		RMSE		R^2	
	NEXAI	MODIS	NEXAI	MODIS	NEXAI	MODIS
DRA (n = 982)	4.414	5.079	5.343	5.949	0.981	0.971
BON (n = 428)	2.232	2.608	3.005	3.638	0.941	0.914
GWN (n = 489)	2.548	2.959	3.177	3.752	0.958	0.921
PSU (n = 353)	1.981	1.965	2.529	2.770	0.967	0.950
SXF (n = 482)	2.190	2.432	2.821	3.588	0.974	0.948
TBL (n = 617)	2.818	2.457	3.642	3.396	0.980	0.970
FPK (n = 511)	1.812	2.366	2.474	3.203	0.982	0.964
Average	2.571	2.838	3.284	3.757	0.969	0.948

(b) NEXAI-LST and NOAA GOES-16 standard LST product						
Station	MAE		RMSE		R^2	
	NEXAI	NOAA	NEXAI	NOAA	NEXAI	NOAA
DRA (n=5574)	4.155	4.131	4.839	4.978	0.987	0.986
BON (n=2773)	1.787	1.638	2.420	2.418	0.963	0.956
GWN (n=2157)	2.011	2.003	2.505	2.432	0.968	0.964
PSU (n=2313)	1.773	2.491	2.308	3.151	0.968	0.940
SXF (n=3244)	1.920	1.593	2.432	2.143	0.979	0.980
TBL (n=3357)	2.076	2.059	2.544	2.642	0.988	0.983
FPK (n=2981)	1.365	1.243	1.823	1.711	0.988	0.989
Average	2.155	2.165	2.696	2.782	0.977	0.971

availability of training data taking advantage of these ancillary datasets.

Variance between MODIS LST and NEXAI-LST can be attributed to sources including: imperfect approximation of model physics by the network, spatial and temporal shifts between LEO and GEO data, differences in viewing and illumination geometry, error in the GEO radiance values, and error in the LEO data product. It should be noted that MODIS LST products are commonly validated for intermediate values in the distribution: Coll *et al.* found a bias of +0.1 K and standard deviation of 0.6 K for a range of surface temperatures between 298.15 and 305.15 K [49], while Wan *et al.* found better than 1 K accuracy in the range from 263 to 300 K [50]. According to our assessment, these ranges of intermediate temperature represent only limited portions (approximately 30% and approximately 80%, respectively) of the observed LST distribution. Neither study evaluated performance on the warmest 10% of observations. Evaluating the bias of our predictions relative to MODIS LST across multiple quantiles of the MODIS LST distribution indicated that NEXAI-LST is less biased for the middle of the distribution but has higher error in the warm and cool tails of the distribution, which have less well-defined accuracy characteristics. The proposed model was trained with the objective of minimizing global error between outputs and labels, so understanding the model's performance in rare or complex conditions is challenging. Training with importance weighting might improve performance for rare cold and hot values in the future. Other ML-based studies of LST prediction have attained MAE of 2.85 K on day ahead

prediction [58] and MAE of 0.16–0.26 K using other weather variables as predictors [59] over limited geographic regions.

This modeling approach, which was demonstrated for two geostationary sensors, has potential application to further geostationary sensors, which could result in nearly full global coverage. In contrast to hand-tuned physics-based models, which result in spatially and temporally coarsened data, our ML model can produce LST at spatiotemporal resolutions constrained only by the attributes of Level 1 data (2 km/10 minutes). Sub-hourly geostationary observations result in more cloud-free observations [18], which, together with their unprecedented spatial and temporal consistency, can help reconstruct cloudy images and answer questions about temperature variation over fine timescales. Remote sensing of surface urban heat islands [9], [60] has an increasingly important role in providing data to inform climate adaptation policy and extreme heat response.

VII. CONCLUSION

In this study we used deep learning to train a data-driven model for MODIS-like land surface temperature for application to geostationary sensors. We found that the ML model could predict spatially and temporally consistent LST at a finer resolution than the standard GOES-R product. Further, we explored the generalizability of the approach across unseen times of day, spatial regions, and satellite sensors. In the generalization studies, we found that a model trained on one LEO-GEO pair (Terra and GOES-16) can credibly transfer to holdout LEO and GEO satellites based on statistical performance, but with some limitations in application to out-of-sample conditions. In validation against independent, ground-based measurements, we found that our model performs as well or better than the standard GOES-16 and MODIS LST product, while not depending on prior information about land surface and atmosphere states. These experiments demonstrate the capability of deep learning models to approximate complex physics-based functions by learning from huge, real-world datasets. Our LEO-GEO approach is complementary to physics-based modeling, as it draws obliquely upon the MODIS LST algorithm to generate NEXAI-LST, a product of improved spatial and temporal resolution.

REFERENCES

- [1] G. A. Meehl and C. Tebaldi, “More intense, more frequent, and longer lasting heat waves in the 21st century,” *Science*, vol. 305, no. 5686, pp. 994–997, Aug. 2004.
- [2] S. E. Perkins-Kirkpatrick and S. C. Lewis, “Increasing trends in regional heatwaves,” *Nature Commun.*, vol. 11, no. 1, pp. 1–8, Dec. 2020.
- [3] Centers for Disease Control and Prevention, “Heat-related deaths—United States, 1999–2003,” *Morbidity Mortality Weekly Rep.*, vol. 55, no. 29, pp. 796–798, 2006.
- [4] A. Hsu, G. Sheriff, T. Chakraborty, and D. Manya, “Disproportionate exposure to urban heat island intensity across major US cities,” *Nature Commun.*, vol. 12, no. 1, pp. 1–11, Dec. 2021.
- [5] T. Li, J. Ban, R. M. Horton, D. A. Bader, G. Huang, Q. Sun, and P. L. Kinney, “Heat-related mortality projections for cardiovascular and respiratory disease under the changing climate in Beijing, China,” *Sci. Rep.*, vol. 5, no. 1, pp. 1–8, Sep. 2015.

- [6] C. K. Uejio, J. D. Tamerius, J. Vredenburg, G. Asaeda, D. A. Isaacs, J. Braun, A. Quinn, and J. P. Freese, "Summer indoor heat exposure and respiratory and cardiovascular distress calls in New York City, NY, U.S.," *Indoor Air*, vol. 26, no. 4, pp. 594–604, Aug. 2016.
- [7] A. Barreca and J. Schaller, "The impact of high ambient temperatures on delivery timing and gestational lengths," *Nature Climate Change*, vol. 10, no. 1, pp. 77–82, Jan. 2020.
- [8] Y.-C. Chen, H.-W. Chiu, Y.-F. Su, Y.-C. Wu, and K.-S. Cheng, "Does urbanization increase diurnal land surface temperature variation? Evidence and implications," *Landscape Urban Planning*, vol. 157, pp. 247–258, Jan. 2017.
- [9] Y. Chang, J. Xiao, X. Li, S. Frolking, D. Zhou, A. Schneider, Q. Weng, P. Yu, X. Wang, X. Li, S. Liu, and Y. Wu, "Exploring diurnal cycles of surface urban heat island intensity in Boston with land surface temperature data derived from GOES-R geostationary satellites," *Sci. Total Environ.*, vol. 763, Apr. 2021, Art. no. 144224.
- [10] R. R. Nemani and S. W. Running, "Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data," *J. Appl. Meteorol.*, vol. 28, no. 4, pp. 276–284, Apr. 1989.
- [11] M. Zink, J. Mai, M. Cuntz, and L. Samaniego, "Conditioning a hydrologic model using patterns of remotely sensed land surface temperature," *Water Resour. Res.*, vol. 54, no. 4, pp. 2976–2998, Apr. 2018.
- [12] H. L. Shah, T. Zhou, M. Huang, and V. Mishra, "Strong influence of irrigation on water budget and land surface temperature in Indian subcontinental river basins," *J. Geophys. Res., Atmos.*, vol. 124, no. 3, pp. 1449–1462, Feb. 2019.
- [13] D. Sims, A. Rahman, V. Cordova, B. Elmasri, D. Baldocchi, P. Bolstad, L. Flanagan, A. Goldstein, D. Hollinger, and L. Misson, "A new model of gross primary productivity for North American ecosystems based solely on the enhanced vegetation index and land surface temperature from MODIS," *Remote Sens. Environ.*, vol. 112, no. 4, pp. 1633–1646, Apr. 2008.
- [14] Y. Yan, K. Mao, J. Shi, S. Piao, X. Shen, J. Dozier, Y. Liu, H.-L. Ren, and Q. Bao, "Driving forces of land surface temperature anomalous changes in North America in 2002–2018," *Sci. Rep.*, vol. 10, no. 1, pp. 1–13, Dec. 2020.
- [15] K. Braganza, D. J. Karoly, and J. M. Arblaster, "Diurnal temperature range as an index of global climate change during the twentieth century," *Geophys. Res. Lett.*, vol. 31, no. 13, Jul. 2004, Art. no. L13217.
- [16] G. Wang and M. E. Dillon, "Recent geographic convergence in diurnal and annual temperature cycling flattens global thermal profiles," *Nature Climate Change*, vol. 4, no. 11, pp. 988–992, Nov. 2014.
- [17] Z.-L. Li, B.-H. Tang, H. Wu, H. Ren, G. Yan, Z. Wan, I. F. Trigo, and J. A. Sobrino, "Satellite-derived land surface temperature: Current status and perspectives," *Remote Sens. Environ.*, vol. 131, pp. 14–37, Apr. 2013.
- [18] H. Hashimoto, W. Wang, J. L. Dungan, S. Li, A. R. Michaelis, H. Takenaka, A. Higuchi, R. B. Myneni, and R. R. Nemani, "New generation geostationary satellite observations support seasonality in greenness of the Amazon evergreen forests," *Nature Commun.*, vol. 12, no. 1, pp. 1–11, Dec. 2021.
- [19] Z. Wan, "New refinements and validation of the collection-6 MODIS land-surface temperature/emissivity product," *Remote Sens. Environ.*, vol. 140, pp. 36–45, Jan. 2014.
- [20] B.-H. Tang, "Nonlinear split-window algorithms for estimating land and sea surface temperatures from simulated Chinese Gaofen-5 satellite data," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6280–6289, Nov. 2018.
- [21] J. A. Sobrino, Z.-L. Li, and M. P. Stoll, "Impact of the atmospheric transmittance and total water vapor content in the algorithms for estimating satellite sea surface temperatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 31, no. 5, pp. 946–952, Sep. 1993.
- [22] X. Shi, Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–11.
- [23] T. Vandal, E. Kodra, J. Dy, S. Ganguly, R. Nemani, and A. R. Ganguly, "Quantifying uncertainty in discrete-continuous and skewed data with Bayesian deep learning," in *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Jul. 2018, pp. 2377–2386.
- [24] Y. Liu, A. R. Ganguly, and J. Dy, "Climate downscaling using YNet: A deep convolutional network with skip connections and fusion," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 3145–3153.
- [25] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and M. Prabhat, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, no. 7743, pp. 195–204, Feb. 2019.
- [26] R. Wang, K. Kashinath, M. Mustafa, A. Albert, and R. Yu, "Towards physics-informed deep learning for turbulent flow prediction," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 1457–1466.
- [27] R. Lagerquist, D. Turner, I. Ebert-Uphoff, J. Stewart, and V. Hagerty, "Using deep learning to emulate and accelerate a radiative-transfer model," *J. Atmos. Ocean. Technol.*, vol. 38, no. 10, pp. 1673–1696, Jul. 2021.
- [28] K. Duffy, T. J. Vandal, W. Wang, R. R. Nemani, and A. R. Ganguly, "A framework for deep learning emulation of numerical models with a case study in satellite remote sensing," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 5, 2022, doi: [10.1109/TNNLS.2022.3169958](https://doi.org/10.1109/TNNLS.2022.3169958).
- [29] M. F. Kasim, D. Watson-Parris, L. Deaconu, S. Oliver, P. Hatfield, D. H. Froula, G. Gregori, M. Jarvis, S. Khatiwala, J. Korenaga, J. Topp-Mugglestone, E. Viezzier, and S. M. Vinko, "Building high accuracy emulators for scientific simulations with deep neural architecture search," *Mach. Learn., Sci. Technol.*, vol. 3, no. 1, Mar. 2022, Art. no. 015013.
- [30] S. Basu, S. Ganguly, R. R. Nemani, S. Mukhopadhyay, G. Zhang, C. Milesi, A. Michaelis, P. Votava, R. Dubayah, L. Duncanson, B. Cook, Y. Yu, S. Saatchi, R. DiBiano, M. Karki, E. Boyda, U. Kumar, and S. Li, "A semiautomated probabilistic framework for tree-cover delineation from 1-m NAIP imagery using a high-performance computing architecture," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5690–5708, Oct. 2015.
- [31] T. J. Vandal, D. McDuff, W. Wang, K. Duffy, A. Michaelis, and R. R. Nemani, "Spectral synthesis for geostationary satellite-to-satellite translation," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, pp. 1–11, 2021.
- [32] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, no. 6301, pp. 790–794, 2016.
- [33] H. Wang, K. Mao, Z. Yuan, J. Shi, M. Cao, Z. Qin, S. Duan, and B. Tang, "A method for land surface temperature retrieval based on model-data-knowledge-driven and deep learning," *Remote Sens. Environ.*, vol. 265, Nov. 2021, Art. no. 112665.
- [34] G. Yang, R. Pu, W. Huang, J. Wang, and C. Zhao, "A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 4, pp. 2170–2178, Apr. 2010.
- [35] P. Wu, Z. Yin, H. Yang, Y. Wu, and X. Ma, "Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network," *Remote Sens.*, vol. 11, no. 3, p. 300, Feb. 2019.
- [36] M. Chen, B. H. Newell, Z. Sun, C. A. Corr, and W. Gao, "Reconstruct missing pixels of Landsat land surface temperature product using a CNN with partial convolution," *Proc. SPIE*, vol. 11139, Sep. 2019, Art. no. 111390E.
- [37] K. Mao, J. Shi, H. Tang, Z.-L. Li, X. Wang, and K.-S. Chen, "A neural network technique for separating land surface emissivity and temperature from ASTER imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 1, pp. 200–208, Jan. 2008.
- [38] T. Chang and X. J. Xiong, "Assessment of GOES-16/ABI middle wave infrared band using references of Himawari-8/AHI and Aqua/MODIS," *Proc. SPIE*, vol. 11127, Sep. 2019, Art. no. 111270T.
- [39] W. Wang, S. Li, H. Hashimoto, H. Takenaka, A. Higuchi, S. Kalluri, and R. Nemani, "An introduction to the geostationary-NASA earth exchange (GeoNEX) products: 1. Top-of-atmosphere reflectance and brightness temperature," *Remote Sens.*, vol. 12, no. 8, p. 1267, Apr. 2020.
- [40] J. A. Augustine, J. J. DeLuisi, and C. N. Long, "SURFRAD—A national surface radiation budget network for atmospheric research," *Bull. Amer. Meteorol. Soc.*, vol. 81, no. 10, pp. 2341–2358, 2000.
- [41] K. Wang, "Estimation of surface long wave radiation and broadband emissivity using moderate resolution imaging spectroradiometer (MODIS) land surface temperature/emissivity products," *J. Geophys. Res., Atmos.*, vol. 110, no. D11, 2005, Art. no. D11109.
- [42] A. K. Heidinger, I. Laszlo, C. C. Molling, and D. Tarpley, "Using SURFRAD to verify the NOAA single-channel land surface temperature algorithm," *J. Atmos. Ocean. Technol.*, vol. 30, no. 12, pp. 2868–2884, Dec. 2013.

- [43] G. C. Hulley, S. J. Hook, and C. Hughes, "MODIS MOD21 land surface temperature and emissivity algorithm theoretical basis document," Jet Propuls. Lab., Nat. Aeronaut. Space Admin., Pasadena, CA, USA, JPL Publication 12-17, 2012.
- [44] X. Peng, W. Wu, Y. Zheng, J. Sun, T. Hu, and P. Wang, "Correlation analysis of land surface temperature and topographic elements in Hangzhou, China," *Sci. Rep.*, vol. 10, no. 1, pp. 1–16, Dec. 2020.
- [45] H. Jia, D. Yang, W. Deng, Q. Wei, and W. Jiang, "Predicting land surface temperature with geographically weighed regression and deep learning," *Wiley Interdiscipl. Rev., Data Mining Knowl. Discovery*, vol. 11, no. 1, Jan. 2021, Art. no. e1396.
- [46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [47] Y. Bengio and O. Delalleau, "On the expressive power of deep architectures," in *Proc. Int. Conf. Algorithmic Learn. Theory*. Berlin, Germany: Springer, 2011, pp. 18–36.
- [48] Z. Lu, H. Pu, F. Wang, Z. Hu, and L. Wang, "The expressive power of neural networks: A view from the width," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 1–9.
- [49] C. Coll, V. Caselles, J. M. Galve, E. Valor, R. Niclos, J. M. Sánchez, and R. Rivas, "Ground measurements for the validation of land surface temperatures derived from AATSR and MODIS data," *Remote Sens. Environ.*, vol. 97, no. 3, pp. 288–300, Aug. 2005.
- [50] Z. Wan, Y. Zhang, Q. Zhang, and Z.-L. Li, "Validation of the land-surface temperature products retrieved from Terra moderate resolution imaging spectroradiometer data," *Remote Sens. Environ.*, vol. 83, nos. 1–2, pp. 163–180, Nov. 2002.
- [51] Z. Wan, "New refinements and validation of the MODIS land-surface temperature/emissivity products," *Remote Sens. Environ.*, vol. 112, no. 1, pp. 59–74, 2008.
- [52] K. A. McKinnon, A. Rhines, M. P. Tingley, and P. Huybers, "The changing shape of Northern hemisphere summer temperature distributions," *J. Geophys. Res., Atmos.*, vol. 121, no. 15, pp. 8849–8868, Aug. 2016.
- [53] A. Rhines, K. A. McKinnon, M. P. Tingley, and P. Huybers, "Seasonally resolved distributional trends of North American temperatures show contraction of winter variability," *J. Climate*, vol. 30, no. 3, pp. 1139–1157, Feb. 2017.
- [54] B. H. Samset, C. W. Stjern, M. T. Lund, C. W. Mohr, M. Sand, and A. S. Daloz, "How daily temperature and precipitation distributions evolve with global surface temperature," *Earth's Future*, vol. 7, no. 12, pp. 1323–1336, Dec. 2019.
- [55] X. Guan, L. Yang, Y. Zhang, and J. Li, "Spatial distribution, temporal variation, and transport characteristics of atmospheric water vapor over Central Asia and the arid region of China," *Global Planet. Change*, vol. 172, pp. 159–178, Jan. 2019.
- [56] F. Wang, Z. Qin, C. Song, L. Tu, A. Karnieli, and S. Zhao, "An improved mono-window algorithm for land surface temperature retrieval from Landsat 8 thermal infrared sensor data," *Remote Sens.*, vol. 7, no. 4, pp. 4268–4289, Apr. 2015.
- [57] P. Yu, Y. Yu, Y. Rao, Y. Liu, and H. Wang, "Status of the GOES-R land surface temperature product," in *Proc. AGU Fall Meeting*. Washington, DC, USA: AGU, 2018, p. 1.
- [58] S. Kartal and A. Sekertekin, "Prediction of MODIS land surface temperature using new hybrid models based on spatial interpolation techniques and deep learning models," *Environ. Sci. Pollut. Res.*, pp. 1–20, May 2022.
- [59] R. Maddu, A. R. Vanga, J. K. Sajja, G. Basha, and R. Shaik, "Prediction of land surface temperature of major coastal cities of India using bidirectional LSTM neural networks," *J. Water Climate Change*, vol. 12, no. 8, pp. 3801–3819, Dec. 2021.
- [60] D. Zhou, S. Zhao, S. Liu, L. Zhang, and C. Zhu, "Surface urban heat island in China's 32 major cities: Spatial patterns and drivers," *Remote Sens. Environ.*, vol. 152, pp. 51–61, Sep. 2014.



KATE DUFFY received the B.Sc. degree in environmental engineering from Cornell University, Ithaca, NY, USA, in 2015, and the M.S. degree in civil engineering and the Ph.D. degree in interdisciplinary engineering from Northeastern University, Boston, MA, USA, in 2019 and 2021, respectively.

She is currently a Research Scientist with the NASA Ames Research Center, Mountain View, CA, USA, and the Bay Area Environmental Research Institute, Moffett Field, CA, USA. Her research interests include climate change, remote sensing, and applied machine learning.



THOMAS J. VANDAL received the B.S. degree in mathematics from the University of Maryland, College Park, MD, USA, in 2012, and the Ph.D. degree in interdisciplinary engineering from Northeastern University, Boston, MA, USA, in 2018.

He is currently a Research Scientist with the NASA Ames Research Center, Mountain View, CA, USA, and the Bay Area Environmental Research Institute, Moffett Field, CA, USA. His research interests include machine learning and Earth sciences.



RAMAKRISHNA R. NEMANI received the B.S. degree from Andhra Pradesh Agricultural University, Bapatla, India, in 1979, the M.S. degree from Punjab Agricultural University, Ludhiana, India, in 1982, and the Ph.D. degree from the University of Montana, Missoula, MT, USA, in 1987.

He is currently a Senior Earth Scientist with the Ecological Forecasting Laboratory, NASA Ames Research Center, Moffett Field, CA, USA. His research interests include ecological forecasting and collaborative computing in the Earth sciences.