# End-to-End Reconstruction of High-Resolution Temperature Data Using Physics-Guided Deep Learning

**Shengjie Liu** [1] **Lu Zhang** [1] **Siqin Wang** [1]

## Abstract

High-resolution land surface temperature data with fine spatiotemporal granularity is essential for real-world applications. While satellites provide observations at 100 m every 16 days and coarser resolution hourly, these observations are incomplete due to cloud cover and long revisit times. Earth system models provide continuous hourly temperature data but at much coarser spatial resolution (0.1° to 0.25°). In this study, we present an end-to-end, physics-guided deep learning approach for temperature data reconstruction. The approach is a convolutional neural network that incorporates the annual temperature cycle and includes a linear term to amplify the coarse Earth system model temperatures using fine-scale satellite observations. We evaluate the approach using data from GOES-16 (2 km, hourly) and Landsat (100 m, every 16 days), demonstrating effective temperature reconstruction across selected areas. This simple yet effective approach, enabled by physics-guided deep learning, presents a promising direction for reconstructing temperature data under all weather conditions globally.

## 1. Introduction

Surface temperature is a critical physical property of the Earth's system and an important climate indicator (Hansen et al., 2010). Over land, land surface temperature (LST) can be heterogeneous due to complex surface characteristics, with urbanization further exacerbating these differences (Li et al., 2019). Earth system models are developed through synthesizing satellite and meteorological station observations, using physics-based energy balance modeling to simulate surface dynamics (Giorgi & Avissar, 1997). Currently, the best global simulations achieve a spatial resolution of 0.1° to 0.25°, or approximately 10 km to 25 km—much coarser than commonly used satellite data such as Landsat (100 m) and GOES-16 (2 km) (Irons et al., 2012).

High-resolution temperature data are important for many real-world applications (Liu et al., 2023; Li et al., 2023). However, the lack of high spatiotemporal resolution temperature data currently limits many applications, including disease modeling and city planning (Wang et al., 2019; Wimberly et al., 2021). For example, during extreme climate events, understanding neighborhood-level disparities requires timely temperature data. Yet Landsat—the most reliable satellite offering sub-kilometer resolution—only captures data every 16 days. Although geostationary satellites can provide hourly observations at 2 km resolution, cloud cover remains a major limitation, undermining the validity of these observations (Zhang & Du, 2022).

Integrating coarse-resolution temperature data from Earth system models with high-resolution, cloud-contaminated satellite observations presents a promising pathway toward achieving temperature data with high spatiotemporal resolution (Zhang et al., 2021; Wu et al., 2021; Liu et al., 2025). Several approaches have been proposed. For example, the reanalysis and thermal merging (RTM) method integrates reanalysis simulations and MODIS observations through the annual temperature cycle to capture the average trend, and employs random forests to model daily fluctuations (Zhang et al., 2021). Similarly, a more recent study proposed integrating the annual temperature cycle with Gaussian processes to account for daily variation (Liu et al., 2025). These methods can be categorized as two-stage models, which separately capture the annual trend (often through the annual temperature cycle) and daily fluctuations. However, they are typically hand-crafted and require a significant number of manually selected auxiliary variables to represent the Earth's surface. End-to-end training has not yet been achieved in temperature reconstruction tasks.

### 1.1. Deep learning for temperature data reconstruction

The usage of deep learning models presents an opportunity for a unified, end-to-end framework to reconstruct seamless surface temperature data (Rasp et al., 2018; Wegmann & Jaume-Santero, 2023). However, adapting deep learning

[1]University of Southern California. Correspondence to: Shengjie Liu <skrisliu@gmail.com>.

*ICML 2025 Workshop TerraBytes, $42^{nd}$ International Conference on Machine Learning*, Vancouver, Canada. 2025. Copyright 2025 by the author(s).

approaches for temperature reconstruction is challenging. A typical machine learning task requires pairs of training samples $\{x, y\}$, where the algorithm learns a function $f$ such that

$$y \sim f(x) \ , \tag{1}$$

where $x$ is the features and $y$ is the target. Typically in Earth observation, $x$ is multispectral imagery, and $y$ is the corresponding label. For temperature data reconstruction, two factors limit its direct adaptation. First, temperature reconstruction involves time-series data that varies over time at a given location, yet the features $x$ are often only available monthly or even annually (Li et al., 2023). Second, remote sensing features $x$ can only be collected when clouds are not present, whereas the primary interest often lies in reconstructing temperature under cloudy conditions. Another potential framework is superresolution, where $x$ is coarse-resolution data and $y$ the high-resolution data (Lloyd et al., 2021). However, integrating temporal dynamics remains difficult, and training still requires cloud-free pairs.

In this paper, we propose a different strategy: instead of directly predicting temperature, we predict the parameters of the annual temperature cycle to enable full temperature reconstruction. We introduce an additional linear module capture daily fluctuations and use convolutional layers on features representing Earth surface properties to capture spatiotemporal dependencies. The proposed approach allows for unified, end-to-end training.

## 2. Methodology

### 2.1. Overview: Physics-Guided Deep Learning With Time Consideration

In the classical machine learning paradigm, given training samples $\{x, y\}$, the goal is to learn a model $\mathcal{M}$ such that

$$y \sim \mathcal{M}(x). \tag{2}$$

Typically, $x$ is a set of features representing the Earth surface, which in vision tasks often correspond to images. This paradigm treats each sample independently, ignoring any temporal ordering. However, in temperature reconstruction tasks, time is a critical factor that must be explicitly incorporated:

$$y(t) \sim \mathcal{M}(x, t). \tag{3}$$

A naive approach is to include time $t$ as an additional input feature alongside $x$. Yet, this does not exploit the intrinsic temporal structure of the data, and due to the flexibility of neural networks, it risks producing unrealistic predictions that deviate from physical reality. To address this, we introduce the *Annual Temperature Cycle* (ATC) as a physical constraint within the network architecture. The ATC component explicitly models the overall seasonal trend, capturing significant time-series characteristics. To account for daily fluctuation, we introduce one additional linear term applied to the coarse-resolution temperature data from ERA5. Finally, the convolutional layers focus on learning the remaining spatiotemporal variations only. Specifically, the proposed model comprises three additive components, all embedded within a convolutional neural network:

$$y(t) \sim \mathcal{M}_{\mathrm{ATC}}(t \mid \phi_{\mathrm{ATC}}) + \mathcal{M}_{\rho}(t \mid \phi_{\rho}) + \mathcal{M}_{\mathrm{conv}}(\mathbf{X}_F \mid \phi_{\mathrm{conv}}), \tag{4}$$

where $t$ is time, $\mathbf{X}_F$ is features representing the Earth surface, and $\phi = \{\phi_{\mathrm{ATC}}, \phi_{\rho}, \phi_{\mathrm{conv}}\}$ denotes the learnable parameters of each component.

### 2.2. Problem Setup: Temperature Reconstruction

Let $\mathbf{X} \in \mathbb{R}^{H \times W \times C_T}$ denote the input temperature tensor, where $H$ and $W$ are the height and width of the tensor, $C_T$ is the number of time steps in the series, and some entries may be missing (NaN). We define a binary mask $\mathbf{M} \in \{0, 1\}^{H \times W \times C_T}$ indicating observed and missing values:

$$\mathbf{M}(i, j, t) = \begin{cases} 1, & \text{if } \mathbf{X}(i, j, t) \text{ is observed,} \\ 0, & \text{if } \mathbf{X}(i, j, t) \text{ is missing (NaN).} \end{cases}$$

We build a model $g_{\phi}$ for temperature reconstruction, which is to map $\mathbf{X}$ to its estimate $\widehat{\mathbf{X}}$, composed of three additive modules:

$$\begin{aligned} \widehat{\mathbf{X}} = g_{\phi}(\mathbf{X}) = \mathcal{M}_{\mathrm{ATC}}(\mathbf{X} \mid \phi_{\mathrm{ATC}}) \\ + \mathcal{M}_{\rho}(\mathbf{X} \mid \phi_{\rho}) \\ + \mathcal{M}_{\mathrm{conv}}(\mathbf{X}_F \mid \phi_{\mathrm{conv}}), \end{aligned} \tag{5}$$

where:

- $\mathcal{M}_{\mathrm{ATC}} : \mathbb{R}^{H \times W \times C_T} \to \mathbb{R}^{H \times W \times C_T}$ is a pixel-wise overall temperature trend over time, and we use the annual temperature cycle for modeling, with parameters $\phi_{\mathrm{ATC}}$. Each pixel $(i, j)$ has its own learned parameters,
- $\mathcal{M}_{\rho} : \mathbb{R}^{H \times W \times C_T} \to \mathbb{R}^{H \times W \times C_T}$ is a pixel-wise module to capture daily fluctuation. We use a linear term on the coarse-resolution Earth system model land surface temperature data, parameterized by $\phi_{\rho}$,
- $\mathcal{M}_{\mathrm{conv}} : \mathbb{R}^{H \times W \times C_F} \to \mathbb{R}^{H \times W \times C_T}$ is a global spatiotemporal module implemented via convolutional layers $\phi_{\mathrm{conv}}$, which replaces the spatial filtering process in commonly used two-stage approaches and is the key to end-to-end training in temperature reconstruction.

### 2.3. Specific Model Components

**Annual Temperature Cycle** The ATC module models the overall temporal temperature trend at each pixel $(i, j)$ using a cosine function:

$$m_{\mathrm{ATC}}(i, j, t \mid \phi_{\mathrm{ATC}}) = a_{i,j} + b_{i,j} \cos\left(\frac{2\pi t}{T} + \varphi_{i,j}\right),$$

where $a_{i,j,c}$, $b_{i,j,c}$, and $\varphi_{i,j,c}$ are learned parameters corresponding to the annual mean temperature, amplitude, and phase shift respectively, and $T$ is the period (e.g., 365 days).

**Linear Module $\rho$ to Capture Daily Fluctuation** This module models the daily fluctuation that can be reflected from the coarse-resolution Earth system model temperature data $\mathbf{T}_c(i,j,t)$:

$$m_\rho(i,j,t \mid \phi_\rho) = w_{i,j} \cdot \mathbf{T}_c(i,j,t),$$

where $w_{i,j}$ is pixel-wise learnable weights. We can obtain the coarse-resolution temperature data resampled to the target resolution $\mathbf{T}_c(i,j,t)$ via

$$\mathbf{T}_c(i,j,t) = \mathrm{Resample}\big(\widetilde{\mathbf{T}}_c(t)\big) \quad \text{at pixel } (i,j),$$

where $\widetilde{\mathbf{T}}_c$ is the Earth system model temperature at its native resolution and $\mathrm{Resample}(\cdot)$ denotes an interpolation operator mapping coarse grid data to the fine grid indexed by $(i,j,t)$.

**Convolutional Layers** The $\mathcal{M}_{\mathrm{conv}}$ module functions as a global spatiotemporal filtering, capturing complex spatial patterns and nonlinearities that complement the ATC and coarse-resolution temperature data from Earth system model. The input of this module is an additional tensor $\mathbf{X}_F \in \mathbb{R}^{H\times W\times C_F}$, representing Earth surface features such as spectral reflectance or deep representations. $\mathbf{X}_F$ differs in channel size $C_F$ from the temperature input $\mathbf{X}$ and is exclusively used as input to the convolutional module. A series of convolutional operations parameterized by $\phi_{\mathrm{conv}}$ on $\mathbf{X}_F$ to produce spatially refined corrections that enhance the temperature reconstruction, so that the final prediction is consistent with the high-resolution cloudless observations. This separation allows the model to leverage rich Earth surface information without interfering with the temporal modules, improving overall accuracy and reliability.

Specifically, in the experiments, we designed the convolutional layers with four residual blocks, each consisting of two 3×3 convolutional layers and two batch normalization layers, each followed by a ReLU activation. We first upscale the input features representing Earth's surface (five spectral bands in the experiments) to 16 channels, then progressively to 64, 128, and finally to 365 or 366 channels.

### 2.4. Loss Function

To handle missing data, the reconstruction loss is computed only over observed values, using the mask $\mathbf{M}$:

$$\mathcal{L}_{\mathrm{rec}}(\phi) = \frac{1}{\sum_{i,j,t}\mathbf{M}(i,j,t)} \left\| \mathbf{M} \odot \left(\widehat{\mathbf{X}} - \mathbf{X}\right) \right\|_1,$$

where $\odot$ denotes element-wise multiplication and $\|\cdot\|_1$ is the element-wise L1 norm (sum of absolute values).

## 3. Results and Analysis

### 3.1. Datasets

**GOES-16 Geostationary Satellite Data** GOES-16, the first satellite in NOAA's GOES-R series, was launched in November 2016 and enabled 2 km resolution temperature monitoring every 5 minutes for the first time (Beale et al., 2019). The resulting land surface temperature (LST) data are produced hourly, with an accuracy of approximately 2.5 K when surface emissivity is known and proper atmospheric correction is applied, and around 5 K otherwise (Schmit et al., 2018). This study uses all GOES-16 data from 2022 over a 700×550 pixel region centered on New York City, spanning from Québec City to North Carolina and west to Detroit. We use five GOES-16 spectral bands centered at 0.47, 0.64, 0.86, 1.61, and 2.24 µm as input features.

**Landsat Data** We select a region centering Altadena in Los Angeles from February 2024 to January 2025, covering the recent Eaton Fire that has destroyed more than 9,000 structures. The Landsat surface temperature data is derived using a single channel algorithm, with 60% of the observations within 2 K accuracy (Laraby & Schott, 2018). We used the annual mean spectral reflectance from the first seven spectral bands to represent the Earth surface.

### 3.2. Experimental Setup

We conducted the experiments on PyTorch. On the GOES-16 data, we used the Adam optimizer with a learning rate of 0.1 to training the network for 200 epochs. For the Landsat data, we first train the ATC parameters for 300 epochs, then we freeze the ATC parameters and train the convolutional layers for 200 epochs, using the Adam optimizer with a learning rate of 0.1. For both datasets, we reserved 20% of the valid observed data for testing.

### 3.3. Results on GOES-16 Data

Figure 1 presents the reconstruction results for seven selected days, covering a wide range of temperature values across all four seasons, and compares three methods: the ATC model alone, a naïve CNN, and the proposed physics-guided CNN. Among these, the physics-guided CNN consistently demonstrates the best visual performance. For instance, in Figure 1e, comparison with the near-complete ground truth reveals that the ATC model tends to overestimate temperatures in the northern region, while the naïve CNN produces overly smoothed patterns—primarily due to its heavy reliance on spatiotemporal dependencies—and fails to capture the temperature distribution over the Great Lakes. In contrast, the physics-guided CNN generates reconstructions that most closely align with the ground truth.

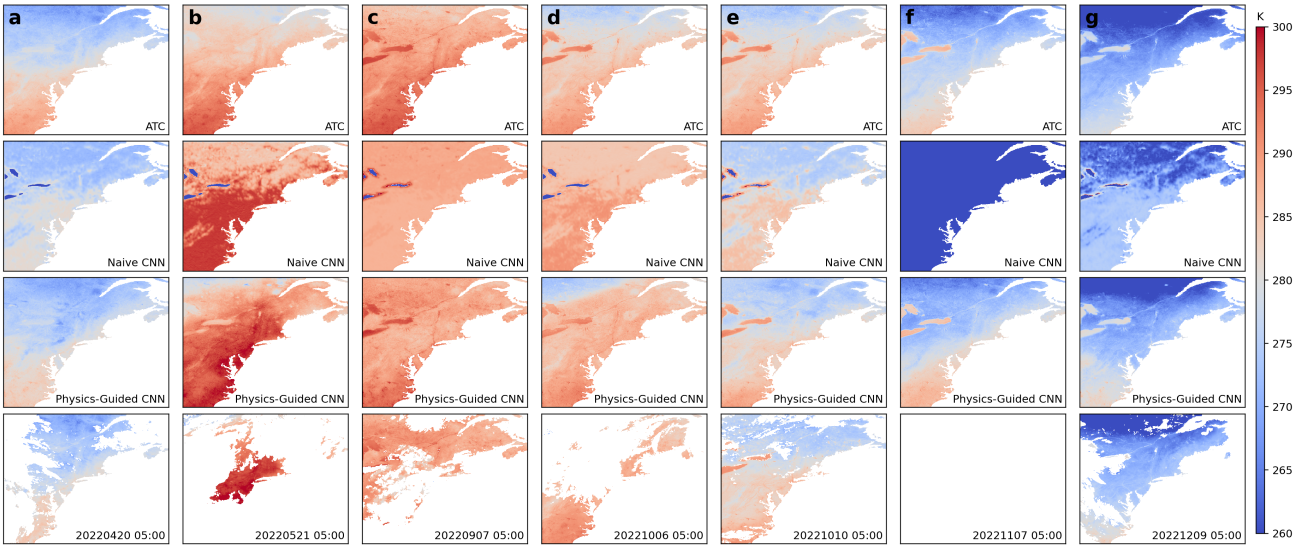


*Figure 1.* Reconstruction results on GOES-16 geostationary satellite data

*Table 1.* Result comparison on the geostationary satellite dataset.

| | Training Data | | Test Data | |
|---|---|---|---|---|
| | MAE (K) | RMSE (K) | MAE (K) | RMSE (K) |
| ATC | 3.63 | 4.91 | 3.75 | 5.02 |
| Naïve CNN | 3.67 | 15.85 | 3.68 | 15.91 |
| Physics-Guided CNN | 1.77 | 2.64 | 1.80 | 2.67 |

*Table 2.* Result comparison on the Landsat dataset.

| | Training Data | | Test Data | |
|---|---|---|---|---|
| | MAE (K) | RMSE (K) | MAE (K) | RMSE (K) |
| ATC | 7.78 | 13.61 | 8.55 | 14.18 |
| Naïve CNN | 2.65 | 14.45 | 2.64 | 14.36 |
| Physics-Guided CNN | 2.06 | 6.40 | 2.13 | 6.42 |

On days without valid observations (e.g., Figure 1f), the naïve CNN fails entirely to reconstruct the temperature field, as this becomes an extrapolation task—a scenario in which conventional deep learning models that rely heavily on statistical patterns often fall short. In contrast, physics-informed approaches, including both the ATC model and the physics-guided CNN, successfully produce plausible reconstructions. In other cases, the physics-guided CNN consistently outperforms the ATC model, visually and through quantitative evaluation metrics (Table 1), owing to its ability to capture spatiotemporal variations more effectively through convolutional layers.

### 3.4. Results on Landsat Data

The results on the Landsat dataset are presented in Table 2. The ATC model shows the weakest performance, likely due to the dataset's sparse coverage and pronounced spatiotemporal heterogeneity. In contrast, the physics-guided CNN consistently achieves the best results, demonstrating strong predictive capability even under challenging conditions.

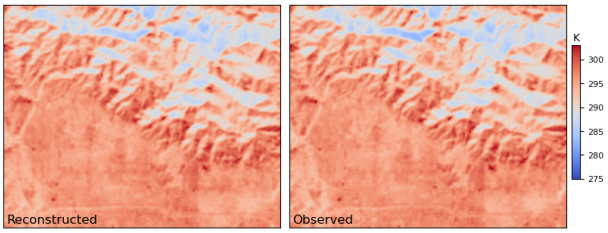


*Figure 2.* Reconstruction results on Landsat data

## 4. Conclusion

In this study, we presented a physics-guided, end-to-end deep learning framework that unifies surface temperature reconstruction into a single vision-based model, achieving accuracy of 2–3 K, comparable to satellite observations across two datasets of varying resolution. Central to our approach is the use of Earth surface property representations. In this work, we used annual mean spectral reflectance. Recent advances in deep representation learning and Earth foundation models are showing promise for capturing more accurate surface characteristics (Marsocci et al., 2024; Zhu et al., 2024). Incorporating such deep features can further enhance the model's ability to capture fine-scale temperature daily fluctuations and spatiotemporal dynamics, and will be a valuable direction for future research.

## Impact Statement

This paper presents a novel framework and preliminary work for reconstructing seamless temperature data to support diverse real-world applications.

## Acknowledgment

S.L. was supported by a USC Dana and David Dornsife College of Letters, Arts and Sciences / Graduate School Fellowship at the University of Southern California. L.Z. was supported by the National Institutes of Health (NIH) under grant numbers P20HL176204, P30ES007048 and R01ES031590. The views and conclusions expressed in this study are those of the authors and do not necessarily represent the official policies or endorsements of the NIH. The authors sincerely appreciate the open data policies and practices of the National Oceanic and Atmospheric Administration (NOAA), the National Aeronautics and Space Administration (NASA), and the European Centre for Medium-Range Weather Forecasts (ECMWF), without which studies like this one would not have been possible.

## References

Beale, C., Norouzi, H., Sharifnezhadazizi, Z., Bah, A. R., Yu, P., Yu, Y., Blake, R., Vaculik, A., and Gonzalez-Cruz, J. Comparison of diurnal variation of land surface temperature from goes-16 abi and modis instruments. *IEEE Geoscience and Remote Sensing Letters*, 17(4):572–576, 2019.

Giorgi, F. and Avissar, R. Representation of heterogeneity effects in earth system modeling: Experience from land surface modeling. *Reviews of Geophysics*, 35(4):413–437, 1997.

Hansen, J., Ruedy, R., Sato, M., and Lo, K. Global surface temperature change. *Reviews of Geophysics*, 48(4), 2010.

Irons, J. R., Dwyer, J. L., and Barsi, J. A. The next landsat satellite: The landsat data continuity mission. *Remote Sensing of Environment*, 122:11–21, 2012.

Laraby, K. G. and Schott, J. R. Uncertainty estimation method and landsat 7 global validation for the landsat surface temperature product. *Remote Sensing of Environment*, 216:472–481, 2018.

Li, W., Ni, L., Li, Z.-l., Duan, S.-B., and Wu, H. Evaluation of machine learning algorithms in spatial downscaling of modis land surface temperature. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7):2299–2307, 2019.

Li, Z.-L., Wu, H., Duan, S.-B., Zhao, W., Ren, H., Liu, X., Leng, P., Tang, R., Ye, X., Zhu, J., et al. Satellite remote sensing of global land surface temperature: Definition, methods, products, and applications. *Reviews of Geophysics*, 61(1), 2023.

Liu, S., Wu, A.-M., and Ho, H. C. Spatial variability of diurnal temperature range and its associations with local climate zone, neighborhood environment and mortality in los angeles. *Urban Climate*, 49:101526, 2023.

Liu, S., Wang, S., and Zhang, L. Daily land surface temperature reconstruction in landsat cross-track areas using deep ensemble learning with uncertainty quantification. *arXiv preprint arXiv:2502.14433*, 2025.

Lloyd, D. T., Abela, A., Farrugia, R. A., Galea, A., and Valentino, G. Optically enhanced super-resolution of sea surface temperature using deep learning. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–14, 2021.

Marsocci, V., Jia, Y., Bellier, G. L., Kerekes, D., Zeng, L., Hafner, S., Gerard, S., Brune, E., Yadav, R., Shibli, A., et al. Pangaea: A global and inclusive benchmark for geospatial foundation models. *arXiv preprint arXiv:2412.04204*, 2024.

Rasp, S., Pritchard, M. S., and Gentine, P. Deep learning to represent subgrid processes in climate models. *Proceedings of the National Academy of Sciences*, 115(39): 9684–9689, 2018.

Schmit, T. J., Lindstrom, S. S., Gerth, J. J., and Gunshor, M. M. Applications of the 16 spectral bands on the advanced baseline imager (abi). *Journal of Operational Meteorology*, 6(4):33–46, 2018. doi: 10.15191/nwajom.2018.0604. URL https://doi.org/10.15191/nwajom.2018.0604.

Wang, C., Wang, Z.-H., Wang, C., and Myint, S. W. Environmental cooling provided by urban trees under extreme heat and cold waves in us cities. *Remote Sensing of Environment*, 227:28–43, 2019.

Wegmann, M. and Jaume-Santero, F. Artificial intelligence achieves easy-to-adapt nonlinear global temperature reconstructions using minimal local data. *Communications Earth & Environment*, 4(1):217, 2023.

Wimberly, M. C., de Beurs, K. M., Loboda, T. V., and Pan, W. K. Satellite observations and malaria: new opportunities for research and applications. *Trends in Parasitology*, 37(6):525–537, 2021.

Wu, P., Yin, Z., Zeng, C., Duan, S.-B., Göttsche, F.-M., Ma, X., Li, X., Yang, H., and Shen, H. Spatially continuous and high-resolution land surface temperature product generation: A review of reconstruction and spatiotemporal fusion techniques. *IEEE Geoscience and Remote Sensing Magazine*, 9(3):112–137, 2021.

Zhang, X., Zhou, J., Liang, S., and Wang, D. A practical reanalysis data and thermal infrared remote sensing data merging (rtm) method for reconstruction of a 1-km all-weather land surface temperature. *Remote Sensing of Environment*, 260:112437, 2021.

Zhang, Z. and Du, Q. Hourly mapping of surface air temperature by blending geostationary datasets from the two-satellite system of goes-r series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 183:111–128, 2022.

Zhu, X. X., Xiong, Z., Wang, Y., Stewart, A. J., Heidler, K., Wang, Y., Yuan, Z., Dujardin, T., Xu, Q., and Shi, Y. On the foundations of earth and climate foundation models. *arXiv preprint arXiv:2405.04285*, 2024.