

A NOVEL AND PRACTICAL APPROACH TO GENERATE ALL-WEATHER 30-METER LAND SURFACE TEMPERATURE DATA

Shengjie Liu¹, Lu Zhang², Siqin Wang¹

¹Spatial Sciences Institute, Dornsife College of Letters, Arts and Sciences,
University of Southern California, Los Angeles, CA, USA

²Division of Biostatistics, Department of Population and Public Health Sciences,
Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

ABSTRACT

Daily temperature data at high spatial resolution are fundamental for many applications but not yet available due to clouds and satellite's long revisit time. Here, we propose a novel and practical approach to generate daily 30-meter land surface temperature (LST) in complex urban areas, using New York City as case study. In the proposed approach, we first use enhanced annual temperature cycles to capture the overall temporal trend from Landsat valid-observed data. We then use Gaussian processes to link the residuals from valid-observed pixels to cloud-covered pixels based on the similarity of surface property. The LST reconstruction process is solved through deep ensemble learning, which provides robust estimations and offers prediction intervals. Experimental results show that the proposed approach can successfully generate LST under clear-sky and heavily cloud-covered situations (RMSE = 0.75-0.78 K). Further, the reconstructed LST achieved indistinguishable performance, compared with valid-observed LST, in estimating near-surface air temperature (RMSE = 1.42-2.96 K). The proposed approach takes one step forward in providing daily temperature data at 30-meter resolution.

Index Terms— land surface temperature, air temperature, temperature data reconstruction, annual temperature cycle, Gaussian processes, ensemble learning

1. INTRODUCTION

As an important climate indicator [1], temperature is a critical variable in many applications [2]. Having accurate temperature data is important for understanding and addressing global issues, such as climate and health in the form of cardiorespiratory diseases and mortality [3, 4]. There are two aspects of temperature: temporal changes and spatial variations [5]. In the temporal domain, it is critical to have daily data to capture the norm and extreme of climate events [2]. In the spatial domain, understanding temperature changes over space allows

us to understand its impacts on different neighborhoods and populations, a fundamental environmental justice issue [6].

However, no single data source is capable to monitor temperature in a temporally and spatially continuous fashion [2, 7, 8]. There are two current popular data sources: near-surface air temperature from weather stations, and land surface temperature (LST) data from satellites [8]. Weather stations are expensive and impossible to set up in every city location. Satellite data are subject to clouds and long revisit time. Given the challenge and significance, generating temperature data at high spatiotemporal resolution has been an important task in remote sensing [2, 7, 8].

Many methods have been developed to generate daily LST products [8]. For example, the annual temperature cycle (ATC) is ideal to model daily mean temperature over time [9]. Several studies have adopted and improved upon the ATC model to generate seamless LST data [10, 11]. Most of these methods are applied to MODIS data at 1 km resolution. At this resolution, it is impossible to distinguish temperature variations within a city (Fig. 1). In Fig. 1, we compare the MODIS and Landsat LST over New York City's zip codes. As many zip codes are smaller than 1 km \times 1 km, MODIS is incapable to distinguish temperature differences between zip codes, not to mention the variations within them. Landsat, on the other hand, captures great temperature variations within the city (Fig. 1). Additionally, some coastal areas such as Manhattan have no valid data from MODIS due to the coarse-resolution water mask. Having temperature data finer than zip codes is critical for many applications, because many socioeconomic data are aggregated at the zip code or finer level. For example, New York City has the daily syndromic surveillance data collected and released to the public at zip code, and London has many socioeconomic statistics aggregated at an equivalent scale (Middle Super Output Area) [4, 12]

While there is a debate on the best use case for LST, air temperature and other types of temperature, given the current sensor technology, seamless data are mostly originated from satellite observations that are fundamentally brightness temperature [13]. Most air temperature and other forms of tem-

Supported by USC Dornsife College / Graduate School Fellowship and NIH (P30ES007048) through USC SCEHSC. Email: liusheng@usc.edu

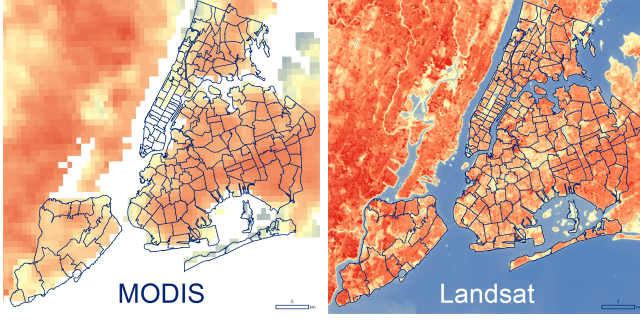


Fig. 1. MODIS and Landsat spatial resolution comparison over New York City's zip codes. Landsat can capture spatial variations that MODIS cannot.

perature data are generated based on satellite observations. The heat exchange between land surface and near-surface air means that LST is bound to link with air temperature (mitigated by some factors including ground cover) [14]. Thus, having seamless LST data at high spatiotemporal resolution means it is one step away from having other forms of temperature data [15, 16, 17]. As only a very limited number of in-situ LST sites exist in remote areas, compared to the in-situ weather stations that are installed in cities, and the fact that air temperature is commonly used in many applications, we can use the ability to generate air temperature to indirectly validate the estimation of all-weather LST [11, 15].

In this study, we propose a novel and practical approach for Landsat LST reconstruction that consists of enhanced ATC models and Gaussian processes. This approach can provide robust estimations and is among the first to offer prediction intervals (through deep ensemble learning). The proposed approach achieved successful LST reconstruction under clear-sky and heavily cloudy situations. The reconstructed LST also shows indistinguishable performance in estimating near-surface air temperature, compared to valid-observed LST, indirectly demonstrating its effectiveness.

2. DATA AND METHOD

2.1. Study Area and Data

We select New York City (NYC) as the study area (Fig. 2). Having a population density of 11,232/km² (28,872/km² in Manhattan), NYC is the most populous city in the United States. NYC adopts open data policy, with many socioeconomic and health statistics aggregated to the zip code level freely available to the public, including the syndromic surveillance system [4, 18]. Having daily temperature data at high spatiotemporal resolution is therefore critical for analyzing temperature-related events. NYC also happens to be situated within two Landsat paths (Fig. 2), doubling data frequency.

With both Landsat 8 and 9 in operation, we successfully obtained 91 Landsat scenes from different dates in 2023. We

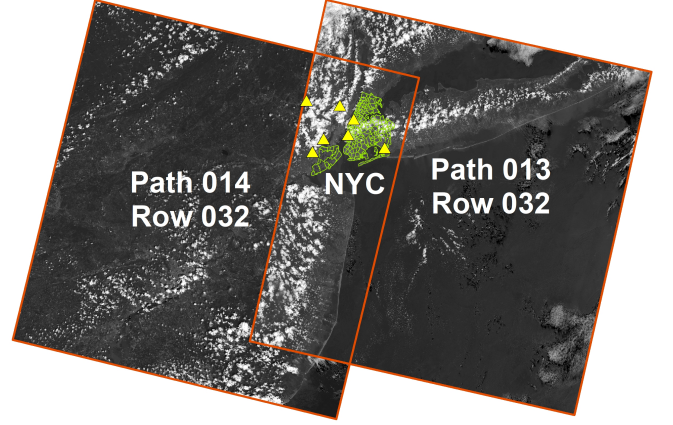


Fig. 2. New York City overlapped with zip code boundaries and seven weather stations (marked with triangle Δ), completely situated within two Landsat paths.

use the thermal band, Band 10, from the level 2 products. We mark all pixels labeled as clear from the cloud mask as clear. The cloud mask is not perfect as noted in data description. We further visually inspected all scenes by splitting them into nine subsets per scene. We manually mark all pixels within a subset as cloudy if it appears cloudy but the cloud mask fails to capture, and mark all pixels within a subset as clear if it appears clear but some areas are incorrectly labeled as not clear (especially at the land-water boundary).

2.2. Method: Enhanced ATC and Gaussian Processes

The reconstructed LST (\mathcal{T}_s^R) is generated through enhanced ATC and Gaussian processes (GP):

$$\mathcal{T}_s^R(d) = \mathcal{T}_s^{ATC}(d) + \mathcal{T}_s^{GP}(d), \quad (1)$$

where d is day of year, $s = [x, y]$ is pixel index, $\mathcal{T}_s^{ATC}(d)$ is the ATC-reconstructed temperature as a function of day of year, and $\mathcal{T}_s^{GP}(d)$ is the temperature residual estimated from GP. This equation can be written out as

$$\mathcal{T}_s^R(d) = C_s + A_s \cos(2\pi/365(d - \phi_s)) + b_s \mathcal{T}^c(d) + \mathcal{GP}_d(s, \mathbf{s}_d), \quad (2)$$

where $C_s + A_s \cos(2\pi/365(d - \phi_s))$ is the classic ATC with three parameters $\{C_s, A_s, \phi_s\}$, b_s is to capture daily fluctuation (we obtain the ERA5 mean skin temperature as \mathcal{T}^c), and $\mathcal{GP}_d(s, \mathbf{s}_d)$ is a daily-specific GP model to link temperature residuals from valid-observed pixels (\mathbf{s}_d) to any cloud-covered pixel (s). We assume pixels with similar surface properties have similar temperature residuals on the same day. We use their blue, green, red and near-infrared mean spectral reflectance in 2022-2024 from Sentinel-2 to account for similarity of surface properties.

2.3. Deep Ensemble Learning

We use deep ensemble learning to find the solutions of the model to achieve robust estimations and obtain prediction intervals. In enhanced ATC, we minimize the L1 loss, which is equivalent to maximizing a Laplace likelihood and can be written out as (with the location index s omitted)

$$p(\mathcal{T} | \omega) \propto \exp\left(\frac{-|\mathcal{T} - \hat{\mathcal{T}}^{ATC}|}{\beta}\right), \quad (3)$$

where $\omega = \{C, A, \phi, b\}$ is the parameters of enhanced ATC to calculate $\hat{\mathcal{T}}^{ATC}$, \mathcal{T} is the valid-observed training data, and β is a scale factor of the Laplace distribution. We obtain 200 snapshots per 2 epochs after the training loss is flat and then calculate the prediction mean and 95% intervals.

2.4. Gaussian Processes To Link Temperature Residuals

On the same day, we assume the LST residuals ($\mathcal{T}^\epsilon = \mathcal{T} - \mathcal{T}^{ATC}$) are related to certain surface property that can be modeled through Gaussian processes, with mean function $m(s)$ and covariance function $k(\cdot, \cdot)$ as (with d omitted):

$$\mathcal{T}^\epsilon(s) \sim \mathcal{GP}(m(s), k(\cdot, \cdot)), \quad (4)$$

where $k(s, s')$ models the similarity between two pixels (s, s'). We use the mean spectral reflectance from 2022-2024 as the features to model their surface similarity. The GP model can be solved through maximizing the log-likelihood function and inherently offers prediction mean and intervals [19]. We use GPyTorch to accelerate computation [20].

The final prediction is the combination of prediction mean from ATC and GP. Based on the law of total variance, we can simply add the respective prediction intervals from ATC and GP to achieve the final prediction intervals.

2.5. Indirect Validation with Air Temperature

For indirect validation, we test the reconstructed LST's performance in estimating air temperature over seven stations (Fig. 2). This is an increasingly-used validation strategy in areas without in-situ LST sites [2, 11, 15]. We use a simple linear regression model that considers only LST, NDVI, elevation, solar radiation factor, and solar zenith angle:

$$\mathcal{T}_a = c + \mathcal{T}_{lst} \times \alpha_1 + V_{ndvi} \times \alpha_2 + V_{elv} \times \alpha_3 + V_{srf} \times \alpha_4 + V_{sza} \times \alpha_5. \quad (5)$$

3. RESULTS AND ANALYSIS

3.1. Results on All-Weather LST Reconstruction

We first show the scatter plots under mostly clear, partially cloudy, and heavily cloudy situations, corresponding to cloud cover $<20\%$, $20\%-80\%$, and $>80\%$ (Fig. 3). The validation is

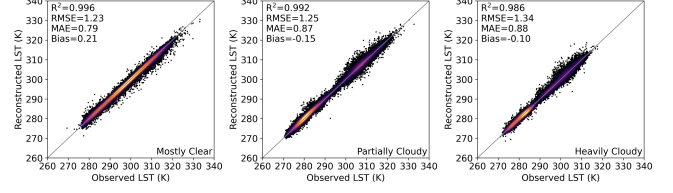


Fig. 3. Reconstruction comparison over mostly clear situations ($<20\%$ clouds), partially cloudy situations ($20\%-80\%$ clouds), and heavily cloudy situations ($>80\%$ clouds).

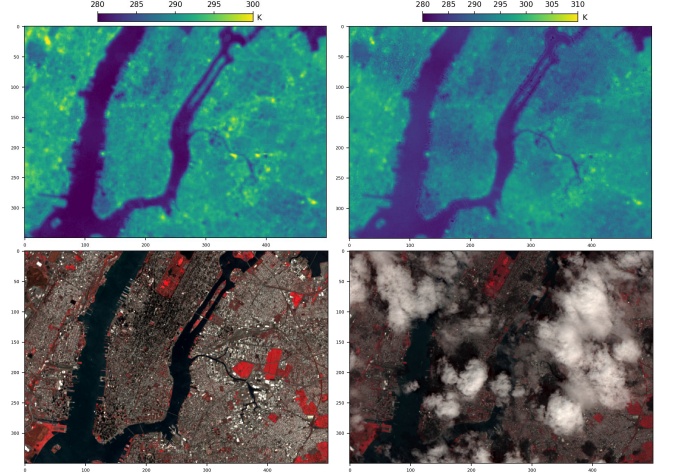


Fig. 4. Reconstructed LST images under clear (*left*) and cloudy (*right*) situations.

conducted over 20% hold-out test data. The reconstruction results are very similar under the three situations (RMSE=1.23-1.34 K), with less cloud coverage slightly better performance. Regardless of clear or heavily cloudy situations, the proposed approach successfully generated accurate LST estimations.

We show the zoom-in view of the reconstructed LST images under clear and cloudy situations in Fig. 4. The cloudy reconstructed LST image shows some pattern different from the clear one. Cloud-covered pixels, including cloud shadow, show relatively lower LST values. In partially cloudy situations, even though some pixels are valid-observed, they are different from clear-sky situations and should be treated differently. In terms of solar energy balance, under partially cloudy situations, the Earth surface receives less energy than clear situations (depending on cloud coverage, thickness, opacity etc.) [21]. With cloud coverage going higher, the LST surface is closer to overcast situations than clear-sky situations. We are only able to observe them due to gaps between clouds, but over the next hour with the clouds shifting, the same location is likely blocked by clouds and becomes non-observable. The different LST patterns are solid evidence that the proposed approach can generate accurate LST in all-weather situations (Fig. 4). This all-weather capability comes from (1) amplifying the coarse-resolution ERA5 tem-

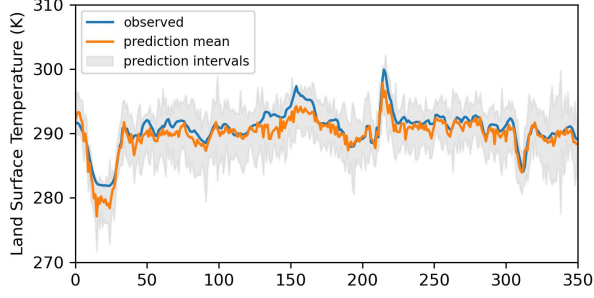


Fig. 5. Demo showing prediction mean and intervals

Table 1. Ablation analysis. ATC is baseline. DEL denotes using deep ensemble learning. GP denotes Gaussian processes.

		ATC	ATC _{DEL}	ATC+GP	ATC _{DEL} +GP
Clear Sky (0% clouds)	R^2	0.891	0.961	0.892	0.974
	RMSE (K)	2.13	1.67	1.53	0.75
	MAE (K)	1.72	1.4	1.17	0.59
Heavily Cloudy (>80% clouds)	R^2	0.835	0.876	0.927	0.981
	RMSE (K)	2.55	2.16	1.55	0.78
	MAE (K)	1.94	1.62	1.18	0.56
	Bias (K)	0.13	0.13	0.03	0.03

perature at higher resolution using valid-observed Landsat pixels and (2) the surface of temperature residuals modeled through Gaussian processes. The proposed approach can also generate prediction intervals, as shown along the x-axis 300 from the clear situation in Fig. 5.

3.2. Indirect Validation via Estimating Air Temperature

We use the capability to estimate air temperature as indirect validation. This is an increasingly-used validation approach for areas without in-situ LST sites [11, 22], such as NYC. It also makes more sense for the current study because (1) it allows us to gain insights of the results within the study area, and (2) near-surface air temperature, which is often generated from seamless LST, is more commonly used in real-world applications [3, 13]. We use simple linear regression models to estimate air temperature from LST, separately using valid-observed and reconstructed LST. If the reconstructed LST shows similar performance with the valid-observed LST, it means they have good agreements with each other, indirectly validating the results. As shown in Fig. 6, the estimation through reconstructed LST (RMSE=1.65-2.73 K) achieved comparable performance in estimating air temperature compared with valid-observed LST (RMSE=1.42-2.96 K).

3.3. Ablation Analysis

We show ablation analysis in Table 1. The experiments were conducted separately based on cloud coverage (0 clouds, over 80% clouds). Deep ensemble learning greatly increases accuracy, in both clear-sky and cloudy situations, by reducing

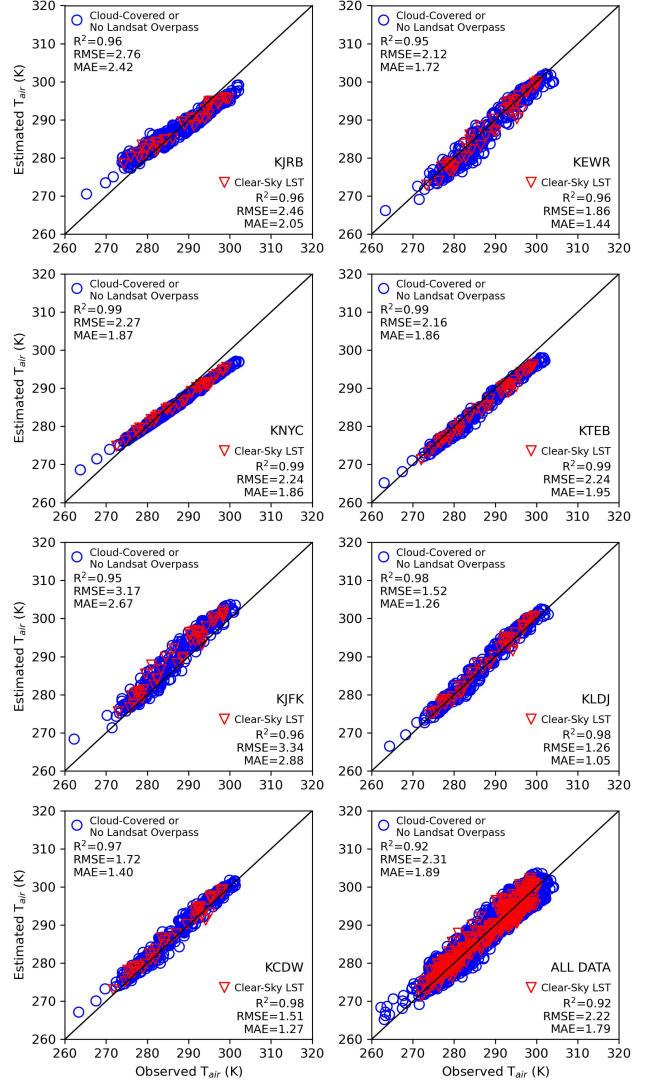


Fig. 6. Results on estimating air temperature through valid-observed and reconstructed LST (indirect validation).

RMSE from 2.13 to 1.67 K (clear sky) and from 2.55 to 2.16 K (heavily cloudy). The usage of GP further reduces RMSE and greatly reduces bias.

4. CONCLUSION

We propose a novel and practical approach to generate daily 30-meter LST data. The proposed approach amplifies ERA5 signals to high spatial resolution through enhanced ATC based on valid-observed Landsat pixels and links temperature residuals from valid-observed to cloud-covered pixels, thereby achieving all-weather LST generation. The method is enhanced via deep ensemble learning to offer robust results and prediction intervals. Experimental results show successful daily LST data generation under all-weather conditions.

5. REFERENCES

- [1] Frank J Wentz and Matthias Schabel, “Precise climate monitoring using complementary satellite data sets,” *Nature*, vol. 403, no. 6768, pp. 414–416, 2000.
- [2] Penghai Wu, Zhixiang Yin, Chao Zeng, Si-Bo Duan, Frank-Michael Göttsche, Xiaoshuang Ma, Xinghua Li, Hui Yang, and Huanfeng Shen, “Spatially continuous and high-resolution land surface temperature product generation: A review of reconstruction and spatiotemporal fusion techniques,” *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 3, pp. 112–137, 2021.
- [3] Antonio Gasparrini, Yuming Guo, Masahiro Hashizume, Eric Lavigne, Antonella Zanobetti, Joel Schwartz, Aurelio Tobias, Shilu Tong, Joacim Rocklöv, Bertil Forsberg, et al., “Mortality risk attributable to high and low ambient temperature: a multicountry observational study,” *Lancet*, vol. 386, no. 9991, pp. 369–375, 2015.
- [4] Shengjie Liu and Hung Chak Ho, “Effects of socioeconomic status and greenspace on respiratory emergency department visits under short-term temperature variations: An age-stratified case time-series study,” *Soc. Sci. Med.*, vol. 343, pp. 116613, 2024.
- [5] James Hansen, Makiko Sato, Reto Ruedy, Ken Lo, David W Lea, and Martin Medina-Elizade, “Global temperature change,” *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 39, pp. 14288–14293, 2006.
- [6] Shengjie Liu and Emily Smith-Greenaway, “Racial and ethnic minorities disproportionately exposed to extreme daily temperature variation in the united states,” *PNAS Nexus*, vol. 3, no. 5, pp. pgae176, 2024.
- [7] Qihao Weng, “Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends,” *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 4, pp. 335–344, 2009.
- [8] Aolin Jia, Shunlin Liang, Dongdong Wang, Kanishka Mallick, Shugui Zhou, Tian Hu, and Shuo Xu, “Advances in methodology and generation of all-weather land surface temperature products from polar-orbiting and geostationary satellites: A comprehensive review,” *IEEE Geosci. Remote Sens. Mag.*, 2024.
- [9] Alexander R Stine, P Huybers, and Inez Y Fung, “Changes in the phase of the annual cycle of surface temperature,” *Nature*, vol. 457, no. 7228, pp. 435–440, 2009.
- [10] Haiping Xia, Yunhao Chen, Adu Gong, Kangning Li, Long Liang, and Zhen Guo, “Modeling daily temperatures via a phenology-based annual temperature cycle model,” *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 14, pp. 6219–6229, 2021.
- [11] Yujia Yang, Wei Zhao, Yanqing Yang, Mengjiao Xu, Hamza Mukhtar, Ghania Tauqir, and Paolo Tarolli, “An annual temperature cycle feature constrained method for generating modis daytime all-weather land surface temperature,” *IEEE Trans. Geosci. Remote Sens.*, 2024.
- [12] Malcolm N Mistry and Antonio Gasparrini, “Real-time forecast of temperature-related excess mortality at small-area level: towards an operational framework,” *Environ. Res. Health*, vol. 2, no. 3, pp. 035011, 2024.
- [13] Jalonne L White-Newsome, Shannon J Brines, Daniel G Brown, J Timothy Dvonch, Carina J Gronlund, Kai Zhang, Evan M Oswald, and Marie S O’Neill, “Validating satellite-derived land surface temperature with in situ measurements: a public health perspective,” *Environ. Health Perspect.*, vol. 121, no. 8, pp. 925–931, 2013.
- [14] Mikhail Ivanovich Budyko, “The heat balance of the earth’s surface,” *Sov. Geogr.*, vol. 2, no. 4, pp. 3–13, 1961.
- [15] A Benali, AC Carvalho, JP Nunes, Nuno Carvalhais, and A Santos, “Estimating air surface temperature in portugal using modis 1st data,” *Remote Sens. Environ.*, vol. 124, pp. 108–121, 2012.
- [16] Josh Hooker, Gregory Duveiller, and Alessandro Cescatti, “A global dataset of air temperature derived from satellite remote sensing and weather stations,” *Sci. Data*, vol. 5, no. 1, pp. 1–11, 2018.
- [17] Zander S Venter, Oscar Brousse, Igor Esau, and Fred Meier, “Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data,” *Remote Sens. Environ.*, vol. 242, pp. 111791, 2020.
- [18] Richard Heffernan, F Mostashari, D Das, M Besculides, C Rodriguez, J Greenko, L Steiner-Sichel, S Balter, A Karpati, P Thomas, et al., “System descriptions: New york city syndromic surveillance systems,” *Morb. Mortal. Wkly. Rep.*, vol. 53, no. suppl, pp. 23–27, 2004.
- [19] Shengjie Liu and Lu Zhang, “Deep feature gaussian processes for single-scene aerosol optical depth reconstruction,” *IEEE Geosci. Remote Sens. Lett.*, 2024.
- [20] Jacob Gardner, Geoff Pleiss, Kilian Q Weinberger, David Bindel, and Andrew G Wilson, “Gpytorch: Blackbox matrix-matrix gaussian process inference with gpu acceleration,” *Adv. Neural Inf. Process.*, vol. 31, 2018.
- [21] Robert E Dickinson, “Land surface processes and climate—surface albedos and energy balance,” in *Adv. Geophys.*, vol. 25, pp. 305–353. Elsevier, 1983.
- [22] Wei Zhao and Si-Bo Duan, “Reconstruction of daytime land surface temperatures under cloud-covered conditions using integrated modis/terra land products and msg geostationary satellite data,” *Remote Sens. Environ.*, vol. 247, pp. 111931, 2020.