Improving Land Surface Temperature Estimation in Cloud Cover Scenarios using Graph-Based Propagation

Iain Rolland[1], Sivasakthy Selvakumaran[1], Shaikh Fairul Edros Ahmad Shaikh[2], Perrine Hamel[2], and Andrea Marinoni[3]

[1]University of Cambridge
[2]Nanyang Technological University
[3]The Arctic University of Norway

April 16, 2024

Abstract

Land surface temperature (LST) serves as an important climate variable which is relevant to a number of studies related to energy and water exchanges, vegetation growth and urban heat island effects. Although LST can be derived from satellite observations, these approaches rely on cloud-free acquisitions. This represents a significant obstacle in regions which are prone to cloud cover.

In this paper, a graph-based propagation method, referred to as GraphProp, is introduced. This method can accurately obtain LST values which would otherwise have been missing due to cloud cover. To validate this approach, a series of experiments are presented using synthetically-obscured Landsat acquisitions. The validation takes place over scenarios ranging from between 10% and 90% cloud cover across three urban locations. In presented experiments, GraphProp recovers missing LST values with a mean absolute error of less than 1.1C, 1.0C and 1.8C in 90% cloud cover scenarios across the studied locations respectively.

Hosted file

`AGU_SuppInfo_Word_Mar_23.docx` available at https://authorea.com/users/762233/articles/739023-improving-land-surface-temperature-estimation-in-cloud-cover-scenarios-using-graph-based-propagation

Improving Land Surface Temperature Estimation in Cloud Cover Scenarios using Graph-Based Propagation

Iain Rolland[1], Sivasakthy Selvakumaran[1], Shaikh Fairul Edros Ahmad Shaikh[2], Perrine Hamel[2], and Andrea Marinoni[1,3]

[1]University of Cambridge, Department of Engineering, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom
[2]Nanyang Technological University, Asian School of the Environment, 50 Nanyang Avenue, Singapore, 639798
[3]UiT the Arctic University of Norway, Department of Physics and Technology, P.O. box 6050 Langnes, NO-9037, Tromsø, Norway

Key Points:

- Gaps in satellite-derived land surface temperature (LST) measurements caused due to clouds can be tackled using graph-based propagation
- The proposed approach, GraphProp, recovers missing LST values more accurately than existing tensor completion methods from literature
- The presented results show the approach to be robust even in highly challenging settings including up to 90% cloud cover

Corresponding author: Iain Rolland, `imr27@cam.ac.uk`

Abstract

Land surface temperature (LST) serves as an important climate variable which is relevant to a number of studies related to energy and water exchanges, vegetation growth and urban heat island effects. Although LST can be derived from satellite observations, these approaches rely on cloud-free acquisitions. This represents a significant obstacle in regions which are prone to cloud cover.

In this paper, a graph-based propagation method, referred to as GraphProp, is introduced. This method can accurately obtain LST values which would otherwise have been missing due to cloud cover. To validate this approach, a series of experiments are presented using synthetically-obscured Landsat acquisitions. The validation takes place over scenarios ranging from between 10% and 90% cloud cover across three urban locations. In presented experiments, GraphProp recovers missing LST values with a mean absolute error of less than 1.1°C, 1.0°C and 1.8°C in 90% cloud cover scenarios across the studied locations respectively.

1 Introduction

Land Surface Temperature (LST) has been identified as an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS) (Zemp et al., 2022). As an ECV it is relevant to the study of a number of phenomena that characterize Earth's climate including urban heat island effects (Nazarian et al., 2022; Zhou et al., 2018; Morabito et al., 2016), water exchanges (Knipper et al., 2019; Anderson et al., 2016) and vegetation health (Bento et al., 2018; Masitoh & Rusydi, 2019). In regions of persistent cloud cover, however, LST rasters often contain missing data. LST retrieval algorithms rely on the ability to measure the thermal infrared (TIR) energy emitted from the land surface (Wan & Dozier, 1996), meaning cloud-occluded TIR observations cannot be used to measure LST. Given the importance of LST as an ECV, it is important to overcome such barriers to measurement in order to have access to regularly-sensed values so as to allow for the subsequent study of the processes to which it pertains.

Existing methods for gap-filling LST data can be categorised into one of two groupings: model-based methods or statistical methods (Mo et al., 2021). Within the former, temperature cycle models including Quan et al. (2016); Sobrino and Julien (2013); Fu and Weng (2015); Zhan et al. (2014) have been proposed which construct physical models of the temperature fluctuations and fit parameters to available observations. Although these models have the strength that they can be used to estimate continuous LST time series, they struggle to capture the spatial variability and higher frequency dynamics of the data. In Zou et al. (2018), although the authors propose a model-based approach which aims to better capture short-term LST fluctuations, they acknowledge that the approach struggles in built-up regions.

In the statistical category of methods, there are a number of approaches which could be applied to the problem of LST gap-filling (Mo et al., 2021). These range in their sophistication from simple imputation methods such as mean filling or linear interpolation to more rigorous methods such as tensor completion methods. Although not directly studied for the problem of tackling LST gaps caused by cloud-obfuscated observations, tensor completion methods including Ng et al. (2017); Srindhuna and Baburaj (2020); He et al. (2019); Chen et al. (2019) have been proposed to address cloud-covered acquisitions. These methods operate under the assumption that the data lies within a low-rank subspace and exploits the observations to complete the missing regions so as to satisfy this assumption. In studies which consider the limits of recoverability for these methods, they generally assume that observations are randomly distributed (Ashraphijuo et al., 2017). This does not hold in the case of cloud-obfuscated data, where missing regions generally form contiguous regions, which these methods struggle to recover (Rolland et al., 2023).

In this manuscript, we propose a graph-based propagation approach which completes missing values in LST rasters caused by cloud-obfuscated Landsat data more accurately than methods in existing literature. The proposed approach avoids LST gaps by completing the missing information in the inputs used to compute LST, specifically, the cloud-obfuscated Landsat data. The major advantage of tackling the problem upstream in this way is that we integrate fully with the downstream LST calculations. The physics which are embedded within the downstream LST equations are utilised fully and therefore any results remain physically consistent with observations. Methods which tackle the gaps at the output stage do not have this same guarantee.

The proposed graph completion approach, referred to as GraphProp, constructs a graph-based representation of the region, where graph nodes represent pixels and graph edges connect pixels that exhibited similar spectral signatures in an earlier cloud-free acquisition of the same region. The graph is then used to complete the partially-observed acquisition by allowing propagation between the observed and missing regions of the image to occur. By removing clouds from the acquisition, the algorithm is able to provide a full input to the downstream LST calculations and thus provide a complete LST raster.

The results presented in this work validate the proposed approach by performing experiments using synthetically-introduced gaps within a Landsat dataset. This allows for the quality of the LST outputs to be assessed against the LST values computed using the original data. The GraphProp approach is shown to more accurately reconstruct the missing LST information than benchmark completion methods and is shown to do so in even extreme cloud cover scenarios where 90% of the image is obscured.

2 Materials and Methods

2.1 GraphProp completion

The proposed approach tackles the problem of filling the gaps in the LST raster by tackling the gaps in the Landsat inputs. If the gaps in the Landsat inputs can be filled, then the Statistical Mono-Window (SMW) algorithm can be applied to the filled inputs to obtain LST values for the entire region. To do so, the proposed graph-based propagation method, GraphProp, constructs a graph-based representation of the area of interest. A graph, $\mathcal{G}$, consists of a set of nodes, $\mathcal{V}$, and a set of edges, $\mathcal{E}$, where each edge connects two nodes. In the context of this study, edges are used to connect pairs of pixels which have been observed to exhibit similarity.

In order to construct the graph-based representation of the region, the GraphProp method makes use of an earlier cloud-free acquisition of the region. Using $\mathcal{H}^0$ and $\mathcal{H}^1$ to denote the reference and partially-observed rasters with three dimensions respectively (having shape $H \times W \times C$ where H, W and C are the height, width and number of image channels respectively), the graph is constructed as follows. By indexing the first two dimensions, i.e. both spatial dimensions, of the reference acquisition, $\mathcal{H}^0[i, j, :]$, we reference a specific pixel and obtain a vector of length C which represents the spectral signature of the location captured by the Landsat 8 platform. Given that the reference acquisition is cloud-free, we can do so for all i and j so as to obtain a set of HW vectors in C-dimensional space. The graph-based representation of the region is constructed using a k-nearest neighbors graph, such that each pixel is connected to its k-nearest neighbors. By introducing an undirected and unweighted edge between a node and its k-nearest neighbors, a graph structure is obtained. This graph provides the structure upon which the observations from the partially-observed acquisition are propagated.

The values which are propagated are the spectral signatures of the pixels in the partially-observed acquisition rather than the reference acquisition and there is no requirement for the spectral signature to remain close across the two acquisitions. This allows temporal changes to take place between the acquisitions and therefore the dynamic nature of the measured spectral signals to be incorporated.

The assumption made by adopting this approach is that pixels which were observed to exhibit spectral similarity in the reference image are likely to also exhibit spectral similarity in the partially-observed acquisition. The reference image will therefore ideally have been captured on a date close to the partially-observed acquisition to ensure the assumption holds. The presented results suggest that this assumption is also reasonable even over longer time frames, provided the region of interest has not underwent significant changes in land cover between acquisitions.

To mathematically describe this propagation approach it is necessary to introduce some notation. First, a function FlattenSpatialDimensions$(\cdot)$ is defined which takes a raster with three dimensions and returns a matrix with two dimensions by flattening the two spatial dimensions, giving $\boldsymbol{F}^0 = \text{FlattenSpatialDimensions}(\mathcal{H}^0) \in \mathbb{R}^{HW\times C}$ and $\boldsymbol{F}^1 = \text{FlattenSpatialDimensions}(\mathcal{H}^1) \in \mathbb{R}^{HW\times C}$. The inverse operation is also defined, UnflattenSpatialDimensions$(\cdot)$, such that $\mathcal{H}^1 = \text{UnflattenSpatialDimensions}(\boldsymbol{F}^1)$. A mask, Ω, is defined which is a set used to index the pixels that were observed in the partially-observed acquisition, such that $\boldsymbol{F}^1_\Omega$ gives the matrix when only the rows of $\boldsymbol{F}^1$ relating to the observed pixels in the partially-observed acquisition are indexed. The complement set, Ω_c, is used to define the missing pixels in the partially-observed acquisition, such that $\boldsymbol{F}^1_{\Omega_c}$ contains the missing entries which are to be recovered.

In this notation, the finite difference approximation to heat diffusion on a graph, as described by Kondor, Risi and Lafferty, John (2002), can be written using the graph's Laplacian matrix, $\mathbf{L}$, as

$$\frac{\partial \boldsymbol{F}^1}{\partial t} \propto -\mathbf{L}\boldsymbol{F}^1. \tag{1}$$

The diffusion equation is modified to hold observed entries fixed, which is achieved by considering the temporal derivative as zero for these rows in $\mathbf{F}^1$. The Laplacian is indexed by its rows and columns such that $\mathbf{L}_{\Omega\Omega_c}$ denotes the submatrix consisting of the rows corresponding to observed pixels and the columns corresponding to the missing pixels.

This allows us to represent the propagation for the unobserved rows in $\mathbf{F}^1$ as

$$\frac{\partial \boldsymbol{F}^1_{\Omega_c}}{\partial t} \propto -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_{\Omega} - \mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c}. \tag{2}$$

The steady state can be found either by iteratively applying steps proportional to the derivative in (2) or by setting the derivative to zero and obtaining $\boldsymbol{F}^1_{\Omega_c}$ as the solution to

$$\mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c} = -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_{\Omega}. \tag{3}$$

The implementation steps of GraphProp are summarised in Algorithm 1.

2.2 LST Calculations

The SMW algorithm, developed by Climate Monitoring Satellite Application Facility (CM-SAF), allows LST values to be calculated from a satellite's TIR band. The SMW models use coefficients obtained by fitting linear regression models that relate measured 11 μm radiance values and the total column water vapor (TCWV) to LST. Once these coefficients are obtained for a given satellite platform, they can then be used to

Algorithm 1 GraphProp algorithm

Input: $\Omega, \mathcal{H}^0, \mathcal{H}^1_\Omega$ ▷ Observation mask, reference input, partially-observed input

1: $\boldsymbol{F}^0 = \text{FlattenSpatialDimensions}(\mathcal{H}^0) \in \mathbb{R}^{HW\times C}$ ▷ Flatten reference input
2: $\boldsymbol{F}^1_\Omega = \text{FlattenSpatialDimensions}(\mathcal{H}^1_\Omega) \in \mathbb{R}^{HW\times C}$ ▷ Flatten partially-observed input
3: $\mathcal{E} \leftarrow \text{kNN}\left(\boldsymbol{F}^0\right)$ ▷ k-nearest neighbors graph (using reference input)
4: $\mathbf{L} = \text{Laplacian}(\mathcal{E})$ ▷ Laplacian matrix of graph
5: $\boldsymbol{F}^1_{\Omega_c} \leftarrow \text{Solve}\left(\mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c} = -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_\Omega\right)$ ▷ Solve diffusion for missing entries
6: $\mathcal{H}^1 = \text{UnflattenSpatialDimensions}(\text{Merge}(\boldsymbol{F}^1_\Omega, \boldsymbol{F}^1_{\Omega_c}))$

Output: $\mathcal{H}^1$

map the satellite-derived inputs to a value representing the LST which would be measured at that location. This approach is adopted by Ermida et al. (2020), where they integrate the process into the Google Earth Engine (GEE) platform to provide a tool for obtaining LST from Landsat observations. The algorithm provided in GEE by Ermida et al. (2020) is depicted schematically in Figure 1.

In addition to the TIR measurements captured by Landsat 8's Band 10, a number of other inputs are used. Firstly, a dynamic estimation of the ground's emissivity is obtained, where emissivity is defined as the ratio of energy emitted by a body to the amount of energy which a black body would emit in equivalent conditions. To estimate this quantity dynamically, the Landsat 8 acquisition is used to obtain an instantaneous measure of fractional vegetation cover (FVC), a quantity describing the fraction of total area covered by vegetation. This value is used to update the static measure of FVC, obtained using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset. Additionally, NCEP/NCAR Reanalysis Data is used to quantify the precipitable water in the atmosphere. Through quantification of TCWV and subsequently the total precipitable water (TPW), the effect that the atmosphere has on the measured brightness temperatures is accounted for (Ermida et al., 2020).

The primary cause of missing LST values is the presence of clouds in the Landsat acquisition. Since the SMW algorithm depicted in Figure 1 is applied in a pixel-wise fashion, in scenes that are only partially obscured by clouds, the gaps in the LST output match the cloud mask pattern which is provided by the quality assessment band.

2.3 Experimental Design

In order to assess the accuracy of the LST values recovered, a set of experiments involving synthetically-obscured data have been performed. This involved the creation of a cloud-free dataset of observations from which the SMW algorithm can compute values against which recovered LST values can be compared. As the graph-based propagation approach involves exploiting an earlier acquisition of the same region, the dataset therefore includes a cloud-free pair of acquisitions for each of the studied locations.

In this study, three urban locations were selected: Jakarta, Indonesia; London, United Kingdom; and Paris, France. Urban locations were selected as the primary focus of the LST accuracy study as they exhibit smaller scale land cover heterogeneities that result in localized land surface temperature variations (Xiao et al., 2007) and therefore represent a challenging gap-filling task. Jakarta falls within a tropical rainforest region while London and Paris sit within an oceanic climate according to Köppen climate definitions (Beck et al., 2018). With Jakarta and London representing cities near a coastline and Paris falling further inland, the three cities therefore might be expected to exhibit a range of differing temperate dynamics and were therefore chosen such that analysis covers wide ranging temperature dynamic characteristics.

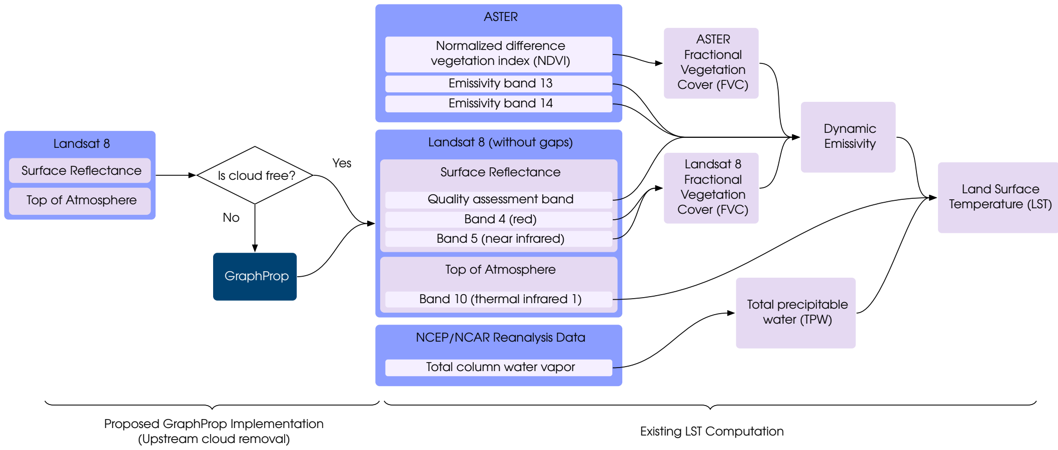


Figure 1: Schematic showing how the proposed approach integrates with existing algorithm for computing LST (Ermida et al., 2020). The advantage of adopting an upstream completion approach is that it does not alter the physics that relate measured thermal infrared energy to LST.

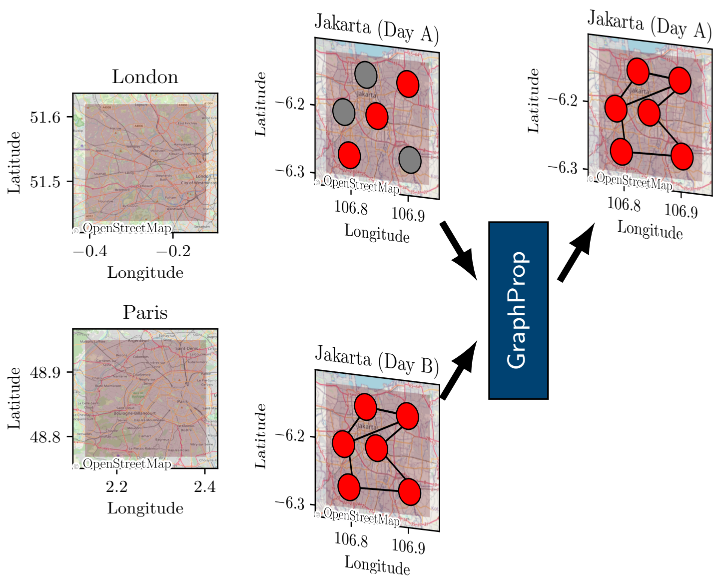


Figure 2: Schematic representing the GraphProp method. The reference acquisition captured on a different date (Day B) is used to construct a graph-based representation of the region to complete the partially-observed acquisition captured on Day A. The graph structure is used to propagate observed values (denoted by red nodes) to complete missing values (denoted by gray nodes). The 20 km $\times$ 20 km region of interest for the respective study regions is depicted by the shaded square.

For each of the three cities, a pair of cloud-free Landsat 8 acquisitions were obtained, as outlined in Table 1. In each location, the region of interest covers a square with side length 20 km and acquisitions were acquired at a resolution of 30 m per pixel (giving rasters made up of 670$\times$670 pixels). In the studied application of the proposed approach the

GraphProp method makes use of an earlier cloud-free reference acquisition to assist the completion task. In order to obtain a cloud-free acquisition of the region it may be necessary to look back further in time, resulting in temporal separations such as seen in Table 1, with the time between acquisitions varying between the three locations, ranging from 9 to 32 weeks. These temporal gaps are representative of the temporal separation which might be expected in a real-world application of the proposed approach.

In order to realistically synthetically obscure the latter acquisitions, genuine cloud masks from other cloud-obscured Landsat 8 acquisitions were collected. These masks were collected such that they represent cloud cover scenarios which range from 10%-90% cloud cover. At each of these 10% intervals, 10 different cloud masks were collected, giving a total of 90 cloud masks. A figure depicting the cloud masks is included within supporting information provided.

2.4 Evaluation

To quantify the accuracy of each method's recovered LST values, the root mean square error (RMSE) and mean absolute error (MAE) were computed over the missing regions of each scene. Using $f(\cdot)$ to summarize the function for deriving the LST from the satellite data, the RMSE and MAE are defined as follows. The algorithm-recovered LST values are contained within the matrix $\overline{\mathrm{LST}}$, where

$$\overline{\mathrm{LST}} = f(\mathcal{H}^1) \in \mathbb{R}^{H \times W}.$$

By flattening the spatial dimensions of the LST matrix, a vector of LST values can be obtained, $\overline{\mathrm{LST}}_{\mathrm{flat}}$, where

$$\overline{\mathrm{LST}}_{\mathrm{flat}} = \mathrm{FlattenSpatialDimensions}(\mathrm{LST}) \in \mathbb{R}^{HW},$$

and then subsequently indexed using Ω_c to consider only the missing regions of the LST matrix:

$$\left(\overline{\mathrm{LST}}_{\mathrm{flat}}\right)_{\Omega_c} \in \mathbb{R}^{|\Omega_c|}.$$

By then comparing against the ground truth, in the matrix LST, the RMSE and MAE were computed as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\left|\left|\left(\overline{\mathrm{LST}}_{\mathrm{flat}}\right)_{\Omega_c} - (\mathrm{LST}_{\mathrm{flat}})_{\Omega_c}\right|\right|_2^2}{|\Omega_c|}} \quad \mathrm{MAE} = \frac{\left|\left|\left(\overline{\mathrm{LST}}_{\mathrm{flat}}\right)_{\Omega_c} - (\mathrm{LST}_{\mathrm{flat}})_{\Omega_c}\right|\right|_1}{|\Omega_c|},$$

where $||\cdot||_2$ and $||\cdot||_1$ are the l_2 and l_1 norms respectively.

2.5 Benchmarked methods

2.5.1 Mean-filled LST

In the absence of more sophisticated methods, the simplest imputation method is to take the partially-computed LST raster and to fill gaps with the mean of the observed values. This provides a baseline against which to compare more sophisticated approaches.

2.5.2 Band-wise mean-filled Landsat 8 inputs

Rather than imputation using the average LST value, a second baseline approach is provided by mean-filling the inputs. To do so, the mean value of the observed portions of each Landsat 8 band is computed and used to fill corresponding missing entries in the input. From the mean-filled inputs, a value of LST can be found for each pixel thus completing the LST raster.

Table 1: Cloud-free Landsat 8 image acquisition pairs

Location	First Acquisition Date	First Acquisition Ground Truth LST Mean (°C)	Second Acquisition Date	Second Acquisition Ground Truth LST Mean (°C)	Days Between Acquisitions
Jakarta	2019-09-11	44.7	2020-04-22	40.5	224
London	2020-06-25	39.1	2020-09-13	27.1	80
Paris	2022-03-06	11.2	2022-05-09	29.6	64

2.5.3 Low-rank tensor completion

A third baseline approach takes a more rigorous approach to the problem of missing data. The low-rank tensor completion approach to data imputation is based on the assumption that the data can be represented by a low-rank tensor.

There have been many studies which present tensor completion algorithms which make this assumption (Cai et al., 2010; Liu et al., 2013; He et al., 2019; Yuan et al., 2019) to recover missing entries. While the low-rank assumption is a powerful tool for tensor completion, it is not always appropriate. For example, it assumes that entries are missing at random which is not the case for cloud-obfuscated satellite imagery.

In this study, the high accuracy low rank tensor completion algorithm (HaLRTC) (Liu et al., 2013) provides a benchmark from this family of methods. It was applied by stacking the partially observed top of atmosphere B10 band with the two surface reflectance bands used to compute LST (SR_B4 and SR_B5). The result is a tensor of size $H \times W \times 3$ which was provided as an input to HaLRTC alongside the observation mask, Ω.

3 Results

Applying a cloud mask to synthetically obscure the latter image of each city's image pair, each method is applied to provide LST values in the missing regions. Doing so for 10 random cloud masks between 10%-90% cloud cover at 10% intervals, the MAE and RMSE was computed for each method. The results, shown in Figure 3, show the mean of each metric when averaged over the 10 random masks at a given cloud cover fraction. The errors are computed over the missing pixels in each experiment only and not over the whole scene, since the errors are zero or negligible for the observed pixels.

Across the three cities studied, the low-rank tensor completion method (HaLRTC) provides improved accuracy versus the naive mean-filling approaches when cloud cover is minimal, i.e. in the region of $10\%-30\%$. In cases which are more severely cloud-obscured, the mean-filling provides equally or more accurate results. This could perhaps be explained by the fact that as cloud cover increases, the large contiguous missing regions deviate further from the founding assumption of the low-rank completion approaches that consider entries to be missing uniform at random.

The proposed method, however, improves the LST accuracy across all cloud cover fractions. GraphProp provides missing LST values with a MAE of less than 1.1°C, 1.0°C and 1.8°C across all tested scenarios for the Jakarta, London and Paris studies respectively. Results do not show a significant deterioration in accuracy as cloud cover increases, unlike in results applying HaLRTC. It has been estimated (Santamouris et al., 2015) that each 1°C increase in temperature within urban heat islands can increase energy demands by between 0.5-5%. The accuracy of LST estimates, therefore, will have direct implications on energy resource planning and management.

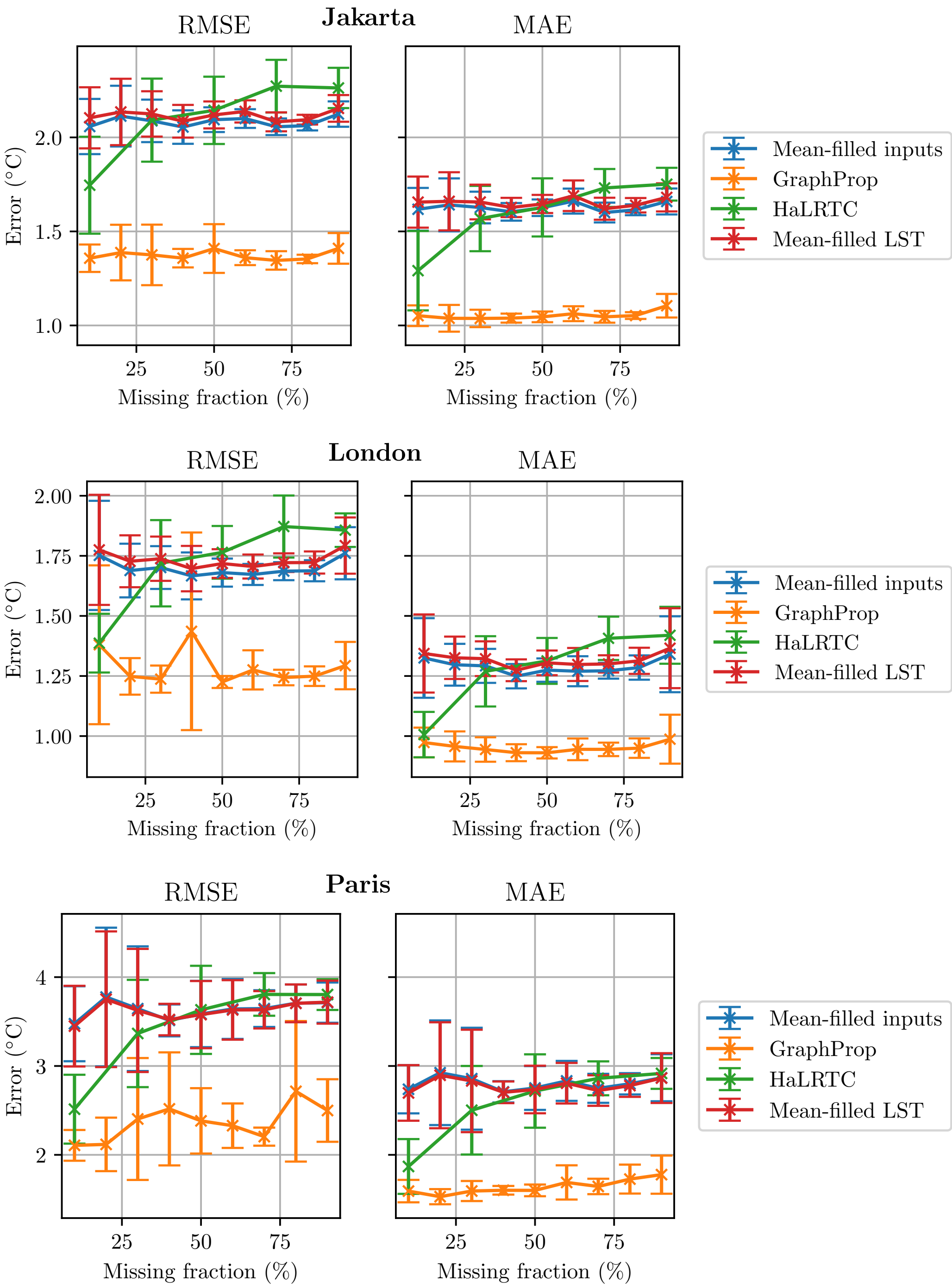
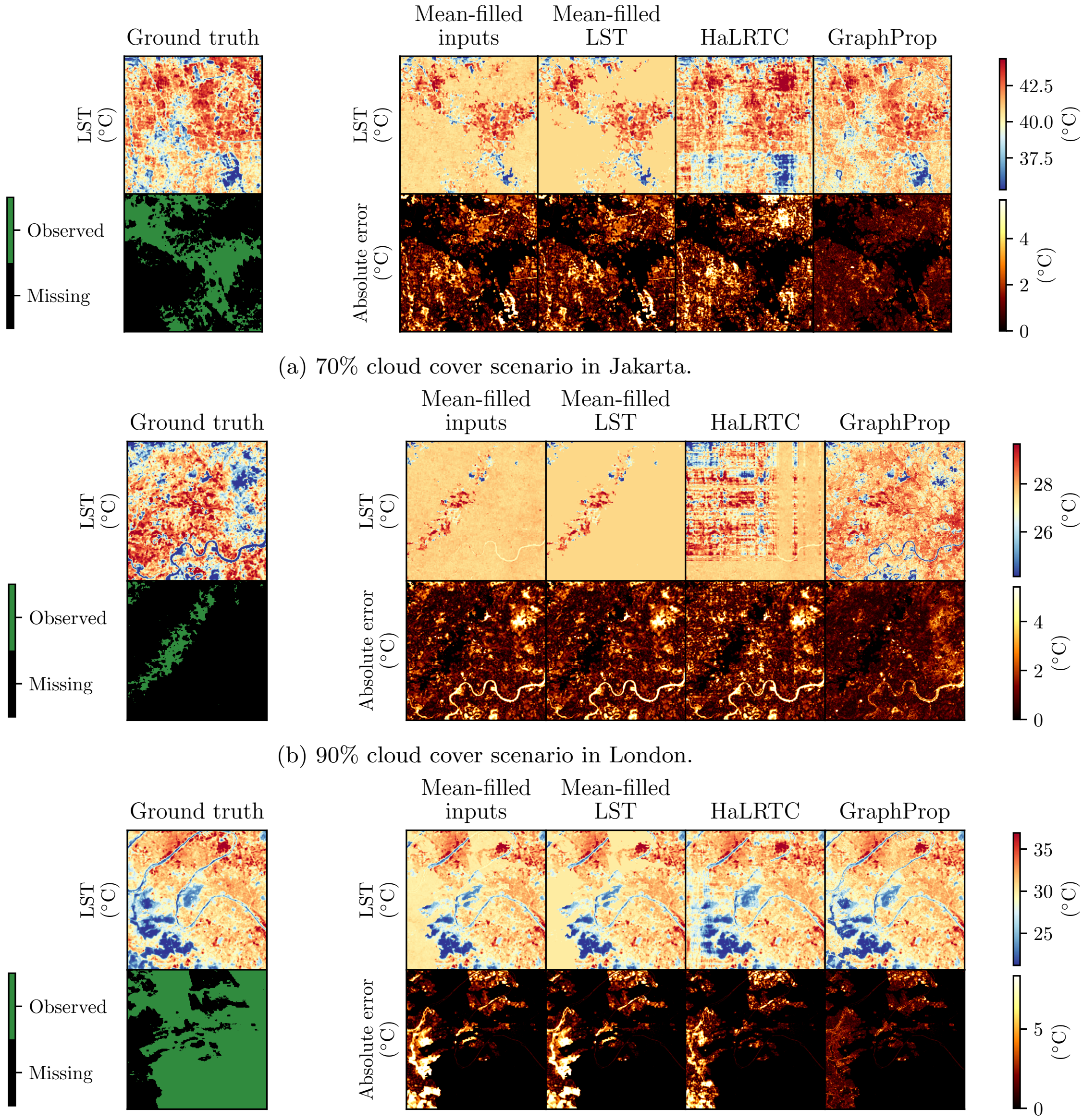


Figure 3: Accuracy of completed LST values as a function of the percentage of the acquisition removed. Each plotted value represents the mean value computed across the 10 random obfuscation masks at the given missing fraction percentage and the error bars represent one standard deviation.

A selection of results are displayed in Figure 4 to illustrate the spatial accuracy characteristics of each method. These absolute error distributions demonstrate that Graph-



(a) 70% cloud cover scenario in Jakarta.

(b) 90% cloud cover scenario in London.

(c) 30% cloud cover scenario in Paris.

Figure 4: Qualitative comparison of the LST imputation methods for at varying amounts of cloud cover.

Prop can recover fine-scale LST variations more accurately than other methods and therefore gives results with fewer neighborhoods of high error magnitude.

Across all the tested methods the errors are larger in the Paris study than for the other two locations. One theory for this behaviour is that the Paris study represents the largest temperature difference between the respective acquisitions (29.6°C versus 11.2°C) and therefore the graph-based representation derived from the colder reference acquisition which was used to complete the partially-observed acquisition may less perfectly characterize the region on the later date. This hypothesis would require further study to prove or disprove conclusively. Nonetheless, the GraphProp results remain the most accurate for the Paris study despite the larger errors observed across methods.

These errors are computed with reference to the values obtained using the SMW algorithm when taking the synthetically-obscured pixel values as inputs. The contribution of Ermida et al. (2020) provides an analysis of the accuracy of this approach with reference to ground truth from in-situ measurements. In their study, performed across 12 locations, they found the RMSE error of the SMW algorithm to be 1.9°C when using Landsat 8 inputs.

4 Discussion and Conclusions

The results presented in this study show that the proposed method, GraphProp, is able to provide accurate LST values in the presence of cloud cover. The method is able to provide more accurate results than the low-rank tensor completion method (HaLRTC) and other more naive mean-filling approaches. The presented results also show this approach to be very robust against the extent of cloud cover present in the scene. The contribution, therefore, represents a useful and practical tool for the analysis of LST in the presence of cloud cover. This will assist applications involving the analysis of LST dynamics, by reducing gaps in time series caused by cloud cover and therefore provide a more complete picture of temporal trends that might otherwise have been difficult to observe.

This study has focused on the analysis of LST in urban areas, however the authors expect similar results to be achievable if applied to other land cover types. The method can also be straightforwardly extended to time series involving more than two acquisitions, for example by combining graphs obtained from each. This would allow for the case where it is not required that any one of the reference acquisitions is fully cloud-free, provided that each location is observed without cloud in at least one of the acquisitions to identify its spectral similarities allowing it to be incorporated into the graph-based representation of the region.

There is scope to extend this research to analyse how the accuracy of the proposed approach varies as other variables such as the temporal or seasonal distance between the respective acquisitions is changed.

Open Research Section

The data used in this study were obtained using the Google Earth Engine platform (https://earthengine.google.com/). The code used to generate the results presented in this paper is available at https://github.com/IMPACTSquad/LST-Gaps.

Acknowledgments
This work was funded through the following sources: the UK Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/T517847/1]; the Centre for Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA) and the Research Council of Norway [RCN grant number 237906]; the IMPETUS project funded by the European Union Horizon 2020 research and innovation program under grant agreement nr. 101037084; the Isaac Newton Trust; and Newnham College, Cambridge, United Kingdom.

References

Anderson, M. C., Zolin, C. A., Sentelhas, P. C., Hain, C. R., Semmens, K., Yilmaz, M. T., … Tetrault, R. (2016). The evaporative stress index as an indicator of agricultural drought in brazil: An assessment based on crop yield impacts. *Remote Sensing of Environment*, *174*, 82–99.

Ashraphijuo, M., Aggarwal, V., & Wang, X. (2017). On deterministic sampling patterns for robust low-rank matrix completion. *IEEE Signal Processing Letters*, *25*(3), 343–347.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific data*, *5*(1), 1–12.

Bento, V. A., Gouveia, C. M., DaCamara, C. C., & Trigo, I. F. (2018). A climatological assessment of drought impact on vegetation health index. *Agricultural and forest meteorology*, *259*, 286–295.

Cai, J.-F., Candès, E. J., & Shen, Z. (2010). A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, *20*(4), 1956-1982. doi: 10.1137/080738970

Chen, Y., He, W., Yokoya, N., & Huang, T.-Z. (2019). Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition. *ISPRS Journal of Photogrammetry and Remote Sensing*, *157*, 93–107.

Ermida, S. L., Soares, P., Mantas, V., Göttsche, F.-M., & Trigo, I. F. (2020). Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series. *Remote Sensing*, *12*(9), 1471.

Fu, P., & Weng, Q. (2015). Temporal dynamics of land surface temperature from landsat tir time series images. *IEEE Geoscience and Remote Sensing Letters*, *12*(10), 2175–2179.

He, W., Yokoya, N., Yuan, L., & Zhao, Q. (2019). Remote sensing image reconstruction using tensor ring completion and total variation. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(11), 8998-9009. doi: 10.1109/TGRS.2019.2924017

Knipper, K. R., Kustas, W. P., Anderson, M. C., Alfieri, J. G., Prueger, J. H., Hain, C. R., . . . others (2019). Evapotranspiration estimates derived using thermal-based satellite remote sensing and data fusion for irrigation management in california vineyards. *Irrigation Science*, *37*, 431–449.

Kondor, Risi and Lafferty, John. (2002, 05). Diffusion kernels on graphs and other discrete input spaces. *ICML, Vol. 2*.

Liu, J., Musialski, P., Wonka, P., & Ye, J. (2013). Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 208-220. doi: 10.1109/TPAMI.2012.39

Masitoh, F., & Rusydi, A. (2019). Vegetation health index (vhi) analysis during drought season in brantas watershed. In *Iop conference series: Earth and environmental science* (Vol. 389, p. 012033).

Mo, Y., Xu, Y., Chen, H., & Zhu, S. (2021). A review of reconstructing remotely sensed land surface temperature under cloudy conditions. *Remote Sensing*, *13*(14). Retrieved from `https://www.mdpi.com/2072-4292/13/14/2838` doi: 10.3390/rs13142838

Morabito, M., Crisci, A., Messeri, A., Orlandini, S., Raschi, A., Maracchi, G., & Munafò, M. (2016). The impact of built-up surfaces on land surface temperatures in italian urban areas. *Science of the Total Environment*, *551*, 317–326.

Nazarian, N., Krayenhoff, E. S., Bechtel, B., Hondula, D. M., Paolini, R., Vanos, J., . . . Santamouris, M. (2022). Integrated assessment of urban overheating impacts on human life. *Earth's Future*, *10*(8), e2022EF002682. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022EF002682` (e2022EF002682 2022EF002682) doi: https://doi.org/10.1029/2022EF002682

Ng, M. K.-P., Yuan, Q., Yan, L., & Sun, J. (2017). An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(6), 3367-3381. doi: 10.1109/TGRS.2017.2670021

Quan, J., Zhan, W., Chen, Y., Wang, M., & Wang, J. (2016). Time series decomposition of remotely sensed land surface temperature and investigation of trends and seasonal variations in surface urban heat islands. *Journal of Geophysical Research: Atmospheres*, *121*(6), 2638–2657.

Rolland, I., Selvakumaran, S., & Marinoni, A. (2023). Remote sensing image completion using a diffusion-based propagation algorithm. In *Image and signal processing for remote sensing xxix* (Vol. 12733, pp. 57–66).

Santamouris, M., Cartalis, C., Synnefa, A., & Kolokotsa, D. (2015). On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—a review. *Energy and buildings*, *98*, 119–124.

Sobrino, J. A., & Julien, Y. (2013). Time series corrections and analyses in thermal remote sensing. In *Thermal infrared remote sensing: Sensors, methods, applications* (pp. 267–285). Springer.

Srindhuna, M., & Baburaj, M. (2020). Estimation of missing data in remote sensing images using t-svd based tensor completion. In *2020 international conference for emerging technology (incet)* (pp. 1–5).

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on geoscience and remote sensing*, *34*(4), 892–905.

Xiao, R., Ouyang, Z., Zheng, H., Li, W., Schienke, E. W., & Wang, X. (2007). Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences*, *19*(2), 250-256. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1001074207600412` doi: https://doi.org/10.1016/S1001-0742(07)60041-2

Yuan, L., Li, C., Mandic, D., Cao, J., & Zhao, Q. (2019). Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 33, pp. 9151–9158).

Zemp, M., Chao, Q., Han Dolman, A. J., Herold, M., Krug, T., Speich, S., … Yu, W. (2022). Gcos 2022 implementation plan. *Global Climate Observing System GCOS*(244), 85.

Zhan, W., Zhou, J., Ju, W., Li, M., Sandholt, I., Voogt, J., & Yu, C. (2014). Remotely sensed soil temperatures beneath snow-free skin-surface using thermal observations from tandem polar-orbiting satellites: An analytical three-time-scale model. *Remote Sensing of Environment*, *143*, 1–14.

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., … Sobrino, J. A. (2018). Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sensing*, *11*(1), 48.

Zou, Z., Zhan, W., Liu, Z., Bechtel, B., Gao, L., Hong, F., … Lai, J. (2018). Enhanced modeling of annual temperature cycles with temporally discrete remotely sensed thermal observations. *Remote Sensing*, *10*(4), 650.

Improving Land Surface Temperature Estimation in Cloud Cover Scenarios using Graph-Based Propagation

Iain Rolland[1], Sivasakthy Selvakumaran[1], Shaikh Fairul Edros Ahmad Shaikh[2], Perrine Hamel[2], and Andrea Marinoni[1,3]

[1]University of Cambridge, Department of Engineering, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom
[2]Nanyang Technological University, Asian School of the Environment, 50 Nanyang Avenue, Singapore, 639798
[3]UiT the Arctic University of Norway, Department of Physics and Technology, P.O. box 6050 Langnes, NO-9037, Tromsø, Norway

Key Points:

- Gaps in satellite-derived land surface temperature (LST) measurements caused due to clouds can be tackled using graph-based propagation
- The proposed approach, GraphProp, recovers missing LST values more accurately than existing tensor completion methods from literature
- The presented results show the approach to be robust even in highly challenging settings including up to 90% cloud cover

Corresponding author: Iain Rolland, `imr27@cam.ac.uk`

Abstract

Land surface temperature (LST) serves as an important climate variable which is relevant to a number of studies related to energy and water exchanges, vegetation growth and urban heat island effects. Although LST can be derived from satellite observations, these approaches rely on cloud-free acquisitions. This represents a significant obstacle in regions which are prone to cloud cover.

In this paper, a graph-based propagation method, referred to as GraphProp, is introduced. This method can accurately obtain LST values which would otherwise have been missing due to cloud cover. To validate this approach, a series of experiments are presented using synthetically-obscured Landsat acquisitions. The validation takes place over scenarios ranging from between 10% and 90% cloud cover across three urban locations. In presented experiments, GraphProp recovers missing LST values with a mean absolute error of less than 1.1°C, 1.0°C and 1.8°C in 90% cloud cover scenarios across the studied locations respectively.

1 Introduction

Land Surface Temperature (LST) has been identified as an Essential Climate Variable (ECV) by the Global Climate Observing System (GCOS) (Zemp et al., 2022). As an ECV it is relevant to the study of a number of phenomena that characterize Earth's climate including urban heat island effects (Nazarian et al., 2022; Zhou et al., 2018; Morabito et al., 2016), water exchanges (Knipper et al., 2019; Anderson et al., 2016) and vegetation health (Bento et al., 2018; Masitoh & Rusydi, 2019). In regions of persistent cloud cover, however, LST rasters often contain missing data. LST retrieval algorithms rely on the ability to measure the thermal infrared (TIR) energy emitted from the land surface (Wan & Dozier, 1996), meaning cloud-occluded TIR observations cannot be used to measure LST. Given the importance of LST as an ECV, it is important to overcome such barriers to measurement in order to have access to regularly-sensed values so as to allow for the subsequent study of the processes to which it pertains.

Existing methods for gap-filling LST data can be categorised into one of two groupings: model-based methods or statistical methods (Mo et al., 2021). Within the former, temperature cycle models including Quan et al. (2016); Sobrino and Julien (2013); Fu and Weng (2015); Zhan et al. (2014) have been proposed which construct physical models of the temperature fluctuations and fit parameters to available observations. Although these models have the strength that they can be used to estimate continuous LST time series, they struggle to capture the spatial variability and higher frequency dynamics of the data. In Zou et al. (2018), although the authors propose a model-based approach which aims to better capture short-term LST fluctuations, they acknowledge that the approach struggles in built-up regions.

In the statistical category of methods, there are a number of approaches which could be applied to the problem of LST gap-filling (Mo et al., 2021). These range in their sophistication from simple imputation methods such as mean filling or linear interpolation to more rigorous methods such as tensor completion methods. Although not directly studied for the problem of tackling LST gaps caused by cloud-obfuscated observations, tensor completion methods including Ng et al. (2017); Srindhuna and Baburaj (2020); He et al. (2019); Chen et al. (2019) have been proposed to address cloud-covered acquisitions. These methods operate under the assumption that the data lies within a low-rank subspace and exploits the observations to complete the missing regions so as to satisfy this assumption. In studies which consider the limits of recoverability for these methods, they generally assume that observations are randomly distributed (Ashraphijuo et al., 2017). This does not hold in the case of cloud-obfuscated data, where missing regions generally form contiguous regions, which these methods struggle to recover (Rolland et al., 2023).

In this manuscript, we propose a graph-based propagation approach which completes missing values in LST rasters caused by cloud-obfuscated Landsat data more accurately than methods in existing literature. The proposed approach avoids LST gaps by completing the missing information in the inputs used to compute LST, specifically, the cloud-obfuscated Landsat data. The major advantage of tackling the problem upstream in this way is that we integrate fully with the downstream LST calculations. The physics which are embedded within the downstream LST equations are utilised fully and therefore any results remain physically consistent with observations. Methods which tackle the gaps at the output stage do not have this same guarantee.

The proposed graph completion approach, referred to as GraphProp, constructs a graph-based representation of the region, where graph nodes represent pixels and graph edges connect pixels that exhibited similar spectral signatures in an earlier cloud-free acquisition of the same region. The graph is then used to complete the partially-observed acquisition by allowing propagation between the observed and missing regions of the image to occur. By removing clouds from the acquisition, the algorithm is able to provide a full input to the downstream LST calculations and thus provide a complete LST raster.

The results presented in this work validate the proposed approach by performing experiments using synthetically-introduced gaps within a Landsat dataset. This allows for the quality of the LST outputs to be assessed against the LST values computed using the original data. The GraphProp approach is shown to more accurately reconstruct the missing LST information than benchmark completion methods and is shown to do so in even extreme cloud cover scenarios where 90% of the image is obscured.

2 Materials and Methods

2.1 GraphProp completion

The proposed approach tackles the problem of filling the gaps in the LST raster by tackling the gaps in the Landsat inputs. If the gaps in the Landsat inputs can be filled, then the Statistical Mono-Window (SMW) algorithm can be applied to the filled inputs to obtain LST values for the entire region. To do so, the proposed graph-based propagation method, GraphProp, constructs a graph-based representation of the area of interest. A graph, $\mathcal{G}$, consists of a set of nodes, $\mathcal{V}$, and a set of edges, $\mathcal{E}$, where each edge connects two nodes. In the context of this study, edges are used to connect pairs of pixels which have been observed to exhibit similarity.

In order to construct the graph-based representation of the region, the GraphProp method makes use of an earlier cloud-free acquisition of the region. Using $\mathcal{H}^0$ and $\mathcal{H}^1$ to denote the reference and partially-observed rasters with three dimensions respectively (having shape $H \times W \times C$ where H, W and C are the height, width and number of image channels respectively), the graph is constructed as follows. By indexing the first two dimensions, i.e. both spatial dimensions, of the reference acquisition, $\mathcal{H}^0[i, j, :]$, we reference a specific pixel and obtain a vector of length C which represents the spectral signature of the location captured by the Landsat 8 platform. Given that the reference acquisition is cloud-free, we can do so for all i and j so as to obtain a set of HW vectors in C-dimensional space. The graph-based representation of the region is constructed using a k-nearest neighbors graph, such that each pixel is connected to its k-nearest neighbors. By introducing an undirected and unweighted edge between a node and its k-nearest neighbors, a graph structure is obtained. This graph provides the structure upon which the observations from the partially-observed acquisition are propagated.

The values which are propagated are the spectral signatures of the pixels in the partially-observed acquisition rather than the reference acquisition and there is no requirement for the spectral signature to remain close across the two acquisitions. This allows temporal changes to take place between the acquisitions and therefore the dynamic nature of the measured spectral signals to be incorporated.

The assumption made by adopting this approach is that pixels which were observed to exhibit spectral similarity in the reference image are likely to also exhibit spectral similarity in the partially-observed acquisition. The reference image will therefore ideally have been captured on a date close to the partially-observed acquisition to ensure the assumption holds. The presented results suggest that this assumption is also reasonable even over longer time frames, provided the region of interest has not underwent significant changes in land cover between acquisitions.

To mathematically describe this propagation approach it is necessary to introduce some notation. First, a function FlattenSpatialDimensions$(\cdot)$ is defined which takes a raster with three dimensions and returns a matrix with two dimensions by flattening the two spatial dimensions, giving $\boldsymbol{F}^0 = \text{FlattenSpatialDimensions}(\mathcal{H}^0) \in \mathbb{R}^{HW\times C}$ and $\boldsymbol{F}^1 = \text{FlattenSpatialDimensions}(\mathcal{H}^1) \in \mathbb{R}^{HW\times C}$. The inverse operation is also defined, UnflattenSpatialDimensions$(\cdot)$, such that $\mathcal{H}^1 = \text{UnflattenSpatialDimensions}(\boldsymbol{F}^1)$. A mask, Ω, is defined which is a set used to index the pixels that were observed in the partially-observed acquisition, such that $\boldsymbol{F}^1_\Omega$ gives the matrix when only the rows of $\boldsymbol{F}^1$ relating to the observed pixels in the partially-observed acquisition are indexed. The complement set, Ω_c, is used to define the missing pixels in the partially-observed acquisition, such that $\boldsymbol{F}^1_{\Omega_c}$ contains the missing entries which are to be recovered.

In this notation, the finite difference approximation to heat diffusion on a graph, as described by Kondor, Risi and Lafferty, John (2002), can be written using the graph's Laplacian matrix, $\mathbf{L}$, as

$$\frac{\partial \boldsymbol{F}^1}{\partial t} \propto -\mathbf{L}\boldsymbol{F}^1. \tag{1}$$

The diffusion equation is modified to hold observed entries fixed, which is achieved by considering the temporal derivative as zero for these rows in $\mathbf{F}^1$. The Laplacian is indexed by its rows and columns such that $\mathbf{L}_{\Omega\Omega_c}$ denotes the submatrix consisting of the rows corresponding to observed pixels and the columns corresponding to the missing pixels.

This allows us to represent the propagation for the unobserved rows in $\mathbf{F}^1$ as

$$\frac{\partial \boldsymbol{F}^1_{\Omega_c}}{\partial t} \propto -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_{\Omega} - \mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c}. \tag{2}$$

The steady state can be found either by iteratively applying steps proportional to the derivative in (2) or by setting the derivative to zero and obtaining $\boldsymbol{F}^1_{\Omega_c}$ as the solution to

$$\mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c} = -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_{\Omega}. \tag{3}$$

The implementation steps of GraphProp are summarised in Algorithm 1.

2.2 LST Calculations

The SMW algorithm, developed by Climate Monitoring Satellite Application Facility (CM-SAF), allows LST values to be calculated from a satellite's TIR band. The SMW models use coefficients obtained by fitting linear regression models that relate measured 11 μm radiance values and the total column water vapor (TCWV) to LST. Once these coefficients are obtained for a given satellite platform, they can then be used to

Algorithm 1 GraphProp algorithm

Input: $\Omega, \mathcal{H}^0, \mathcal{H}^1_\Omega$ ▷ Observation mask, reference input, partially-observed input
1: $\boldsymbol{F}^0 = \text{FlattenSpatialDimensions}(\mathcal{H}^0) \in \mathbb{R}^{HW \times C}$ ▷ Flatten reference input
2: $\boldsymbol{F}^1_\Omega = \text{FlattenSpatialDimensions}(\mathcal{H}^1_\Omega) \in \mathbb{R}^{HW \times C}$ ▷ Flatten partially-observed input
3: $\mathcal{E} \leftarrow \text{kNN}\left(\boldsymbol{F}^0\right)$ ▷ k-nearest neighbors graph (using reference input)
4: $\mathbf{L} = \text{Laplacian}(\mathcal{E})$ ▷ Laplacian matrix of graph
5: $\boldsymbol{F}^1_{\Omega_c} \leftarrow \text{Solve}\left(\mathbf{L}_{\Omega_c\Omega_c}\boldsymbol{F}^1_{\Omega_c} = -\mathbf{L}_{\Omega_c\Omega}\boldsymbol{F}^1_\Omega\right)$ ▷ Solve diffusion for missing entries
6: $\mathcal{H}^1 = \text{UnflattenSpatialDimensions}(\text{Merge}(\boldsymbol{F}^1_\Omega, \boldsymbol{F}^1_{\Omega_c}))$
Output: $\mathcal{H}^1$

map the satellite-derived inputs to a value representing the LST which would be measured at that location. This approach is adopted by Ermida et al. (2020), where they integrate the process into the Google Earth Engine (GEE) platform to provide a tool for obtaining LST from Landsat observations. The algorithm provided in GEE by Ermida et al. (2020) is depicted schematically in Figure 1.

In addition to the TIR measurements captured by Landsat 8's Band 10, a number of other inputs are used. Firstly, a dynamic estimation of the ground's emissivity is obtained, where emissivity is defined as the ratio of energy emitted by a body to the amount of energy which a black body would emit in equivalent conditions. To estimate this quantity dynamically, the Landsat 8 acquisition is used to obtain an instantaneous measure of fractional vegetation cover (FVC), a quantity describing the fraction of total area covered by vegetation. This value is used to update the static measure of FVC, obtained using the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Emissivity Dataset. Additionally, NCEP/NCAR Reanalysis Data is used to quantify the precipitable water in the atmosphere. Through quantification of TCWV and subsequently the total precipitable water (TPW), the effect that the atmosphere has on the measured brightness temperatures is accounted for (Ermida et al., 2020).

The primary cause of missing LST values is the presence of clouds in the Landsat acquisition. Since the SMW algorithm depicted in Figure 1 is applied in a pixel-wise fashion, in scenes that are only partially obscured by clouds, the gaps in the LST output match the cloud mask pattern which is provided by the quality assessment band.

2.3 Experimental Design

In order to assess the accuracy of the LST values recovered, a set of experiments involving synthetically-obscured data have been performed. This involved the creation of a cloud-free dataset of observations from which the SMW algorithm can compute values against which recovered LST values can be compared. As the graph-based propagation approach involves exploiting an earlier acquisition of the same region, the dataset therefore includes a cloud-free pair of acquisitions for each of the studied locations.

In this study, three urban locations were selected: Jakarta, Indonesia; London, United Kingdom; and Paris, France. Urban locations were selected as the primary focus of the LST accuracy study as they exhibit smaller scale land cover heterogeneities that result in localized land surface temperature variations (Xiao et al., 2007) and therefore represent a challenging gap-filling task. Jakarta falls within a tropical rainforest region while London and Paris sit within an oceanic climate according to Köppen climate definitions (Beck et al., 2018). With Jakarta and London representing cities near a coastline and Paris falling further inland, the three cities therefore might be expected to exhibit a range of differing temperate dynamics and were therefore chosen such that analysis covers wide ranging temperature dynamic characteristics.

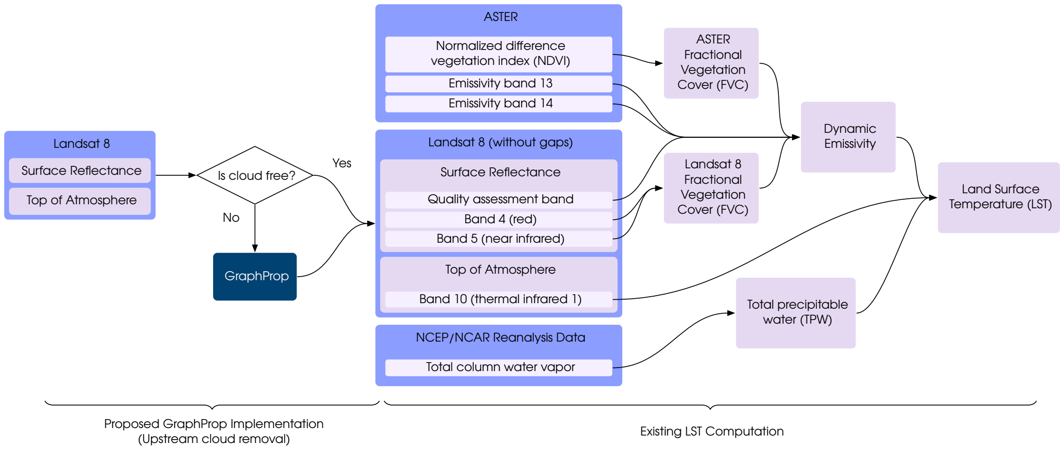


Figure 1: Schematic showing how the proposed approach integrates with existing algorithm for computing LST (Ermida et al., 2020). The advantage of adopting an upstream completion approach is that it does not alter the physics that relate measured thermal infrared energy to LST.

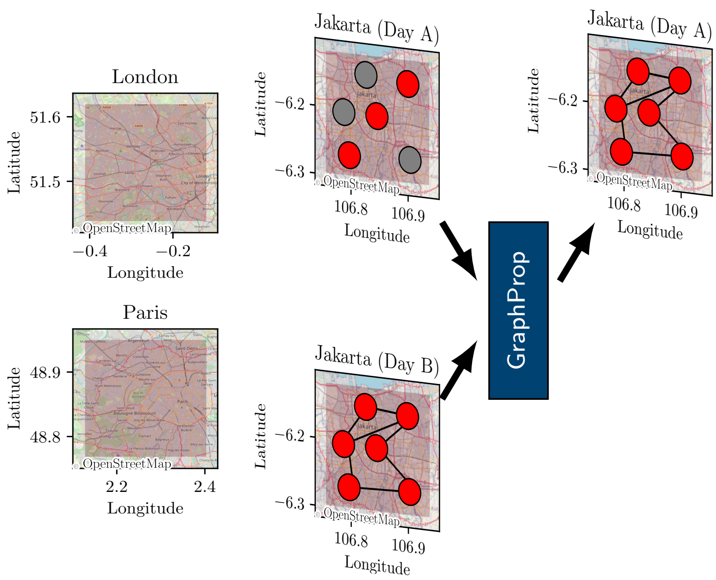


Figure 2: Schematic representing the GraphProp method. The reference acquisition captured on a different date (Day B) is used to construct a graph-based representation of the region to complete the partially-observed acquisition captured on Day A. The graph structure is used to propagate observed values (denoted by red nodes) to complete missing values (denoted by gray nodes). The 20 km × 20 km region of interest for the respective study regions is depicted by the shaded square.

For each of the three cities, a pair of cloud-free Landsat 8 acquisitions were obtained, as outlined in Table 1. In each location, the region of interest covers a square with side length 20 km and acquisitions were acquired at a resolution of 30 m per pixel (giving rasters made up of 670×670 pixels). In the studied application of the proposed approach the

GraphProp method makes use of an earlier cloud-free reference acquisition to assist the completion task. In order to obtain a cloud-free acquisition of the region it may be necessary to look back further in time, resulting in temporal separations such as seen in Table 1, with the time between acquisitions varying between the three locations, ranging from 9 to 32 weeks. These temporal gaps are representative of the temporal separation which might be expected in a real-world application of the proposed approach.

In order to realistically synthetically obscure the latter acquisitions, genuine cloud masks from other cloud-obscured Landsat 8 acquisitions were collected. These masks were collected such that they represent cloud cover scenarios which range from 10%-90% cloud cover. At each of these 10% intervals, 10 different cloud masks were collected, giving a total of 90 cloud masks. A figure depicting the cloud masks is included within supporting information provided.

2.4 Evaluation

To quantify the accuracy of each method's recovered LST values, the root mean square error (RMSE) and mean absolute error (MAE) were computed over the missing regions of each scene. Using $f(\cdot)$ to summarize the function for deriving the LST from the satellite data, the RMSE and MAE are defined as follows. The algorithm-recovered LST values are contained within the matrix $\overline{\text{LST}}$, where

$$\overline{\text{LST}} = f(\mathcal{H}^1) \in \mathbb{R}^{H \times W}.$$

By flattening the spatial dimensions of the LST matrix, a vector of LST values can be obtained, $\overline{\text{LST}}_{\text{flat}}$, where

$$\overline{\text{LST}}_{\text{flat}} = \text{FlattenSpatialDimensions}(\text{LST}) \in \mathbb{R}^{HW},$$

and then subsequently indexed using Ω_c to consider only the missing regions of the LST matrix:

$$\left(\overline{\text{LST}}_{\text{flat}}\right)_{\Omega_c} \in \mathbb{R}^{|\Omega_c|}.$$

By then comparing against the ground truth, in the matrix LST, the RMSE and MAE were computed as follows:

$$\text{RMSE} = \sqrt{\frac{\left|\left|\left(\overline{\text{LST}}_{\text{flat}}\right)_{\Omega_c} - (\text{LST}_{\text{flat}})_{\Omega_c}\right|\right|_2^2}{|\Omega_c|}} \quad \text{MAE} = \frac{\left|\left|\left(\overline{\text{LST}}_{\text{flat}}\right)_{\Omega_c} - (\text{LST}_{\text{flat}})_{\Omega_c}\right|\right|_1}{|\Omega_c|},$$

where $||\cdot||_2$ and $||\cdot||_1$ are the l_2 and l_1 norms respectively.

2.5 Benchmarked methods

2.5.1 Mean-filled LST

In the absence of more sophisticated methods, the simplest imputation method is to take the partially-computed LST raster and to fill gaps with the mean of the observed values. This provides a baseline against which to compare more sophisticated approaches.

2.5.2 Band-wise mean-filled Landsat 8 inputs

Rather than imputation using the average LST value, a second baseline approach is provided by mean-filling the inputs. To do so, the mean value of the observed portions of each Landsat 8 band is computed and used to fill corresponding missing entries in the input. From the mean-filled inputs, a value of LST can be found for each pixel thus completing the LST raster.

Table 1: Cloud-free Landsat 8 image acquisition pairs

Location	First Acquisition Date	First Acquisition Ground Truth LST Mean (°C)	Second Acquisition Date	Second Acquisition Ground Truth LST Mean (°C)	Days Between Acquisitions
Jakarta	2019-09-11	44.7	2020-04-22	40.5	224
London	2020-06-25	39.1	2020-09-13	27.1	80
Paris	2022-03-06	11.2	2022-05-09	29.6	64

2.5.3 Low-rank tensor completion

A third baseline approach takes a more rigorous approach to the problem of missing data. The low-rank tensor completion approach to data imputation is based on the assumption that the data can be represented by a low-rank tensor.

There have been many studies which present tensor completion algorithms which make this assumption (Cai et al., 2010; Liu et al., 2013; He et al., 2019; Yuan et al., 2019) to recover missing entries. While the low-rank assumption is a powerful tool for tensor completion, it is not always appropriate. For example, it assumes that entries are missing at random which is not the case for cloud-obfuscated satellite imagery.

In this study, the high accuracy low rank tensor completion algorithm (HaLRTC) (Liu et al., 2013) provides a benchmark from this family of methods. It was applied by stacking the partially observed top of atmosphere B10 band with the two surface reflectance bands used to compute LST (SR_B4 and SR_B5). The result is a tensor of size $H \times W \times 3$ which was provided as an input to HaLRTC alongside the observation mask, Ω.

3 Results

Applying a cloud mask to synthetically obscure the latter image of each city's image pair, each method is applied to provide LST values in the missing regions. Doing so for 10 random cloud masks between 10%-90% cloud cover at 10% intervals, the MAE and RMSE was computed for each method. The results, shown in Figure 3, show the mean of each metric when averaged over the 10 random masks at a given cloud cover fraction. The errors are computed over the missing pixels in each experiment only and not over the whole scene, since the errors are zero or negligible for the observed pixels.

Across the three cities studied, the low-rank tensor completion method (HaLRTC) provides improved accuracy versus the naive mean-filling approaches when cloud cover is minimal, i.e. in the region of $10\%-30\%$. In cases which are more severely cloud-obscured, the mean-filling provides equally or more accurate results. This could perhaps be explained by the fact that as cloud cover increases, the large contiguous missing regions deviate further from the founding assumption of the low-rank completion approaches that consider entries to be missing uniform at random.

The proposed method, however, improves the LST accuracy across all cloud cover fractions. GraphProp provides missing LST values with a MAE of less than 1.1°C, 1.0°C and 1.8°C across all tested scenarios for the Jakarta, London and Paris studies respectively. Results do not show a significant deterioration in accuracy as cloud cover increases, unlike in results applying HaLRTC. It has been estimated (Santamouris et al., 2015) that each 1°C increase in temperature within urban heat islands can increase energy demands by between 0.5-5%. The accuracy of LST estimates, therefore, will have direct implications on energy resource planning and management.

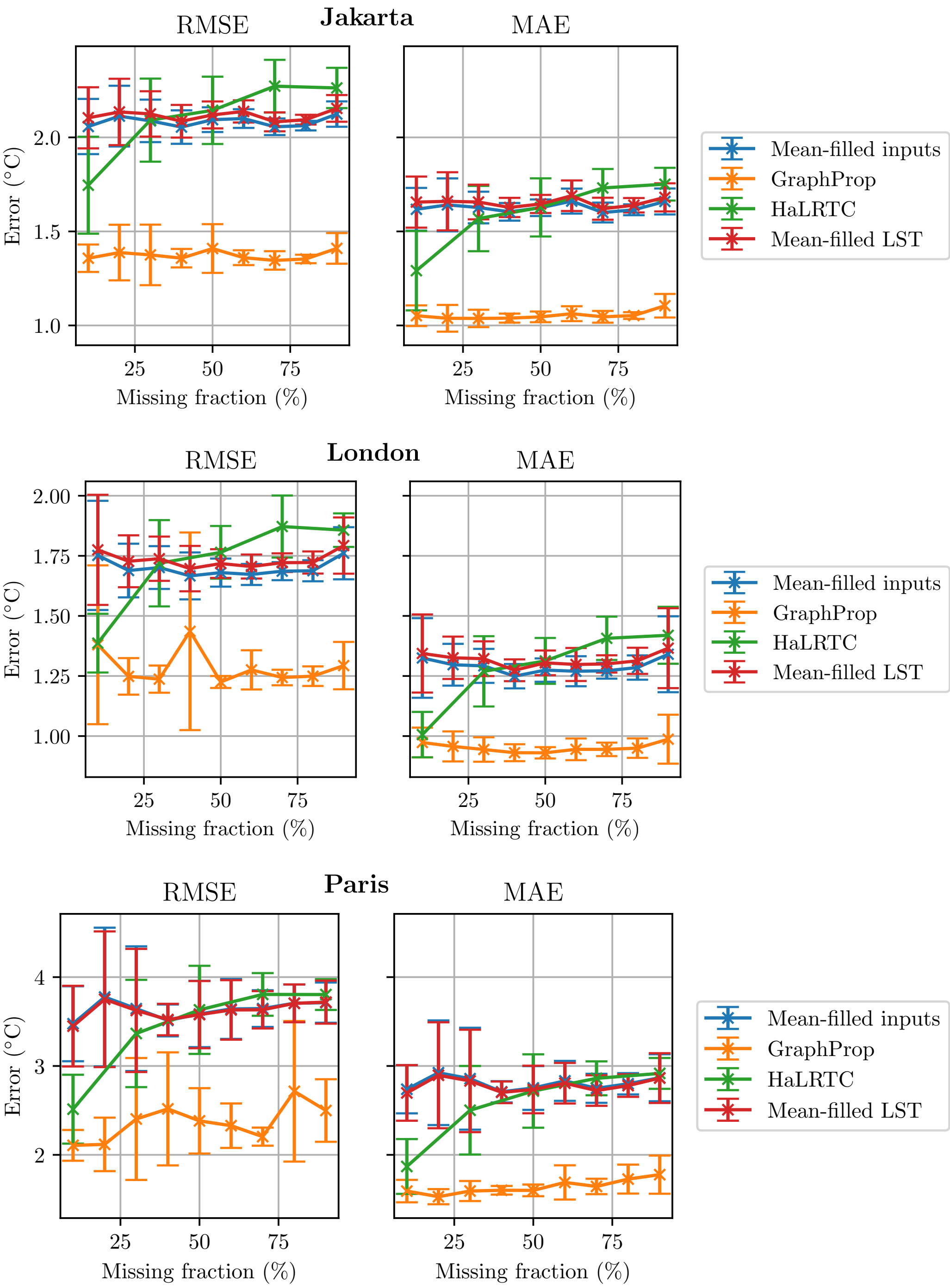
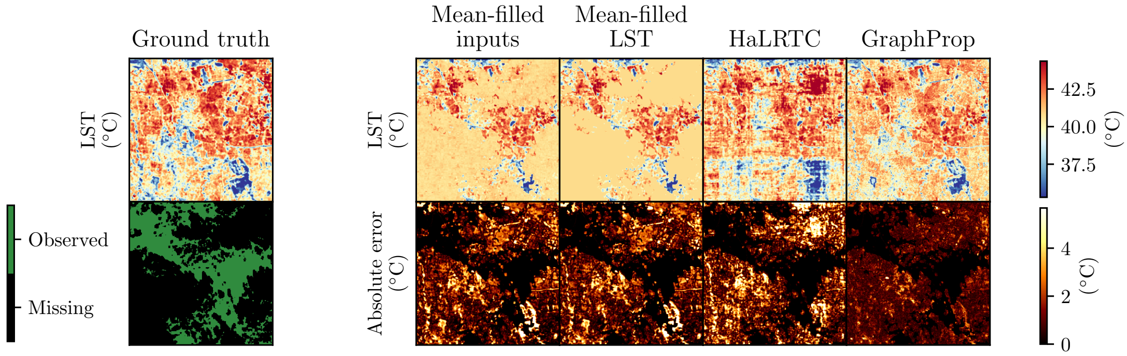
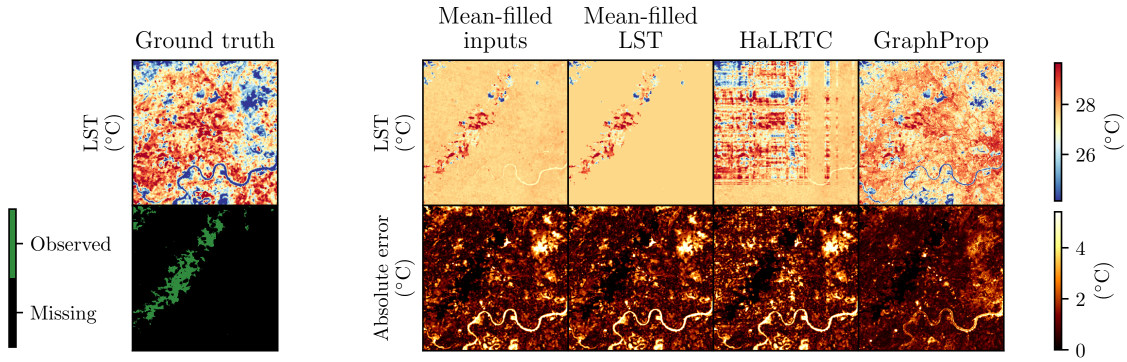


Figure 3: Accuracy of completed LST values as a function of the percentage of the acquisition removed. Each plotted value represents the mean value computed across the 10 random obfuscation masks at the given missing fraction percentage and the error bars represent one standard deviation.

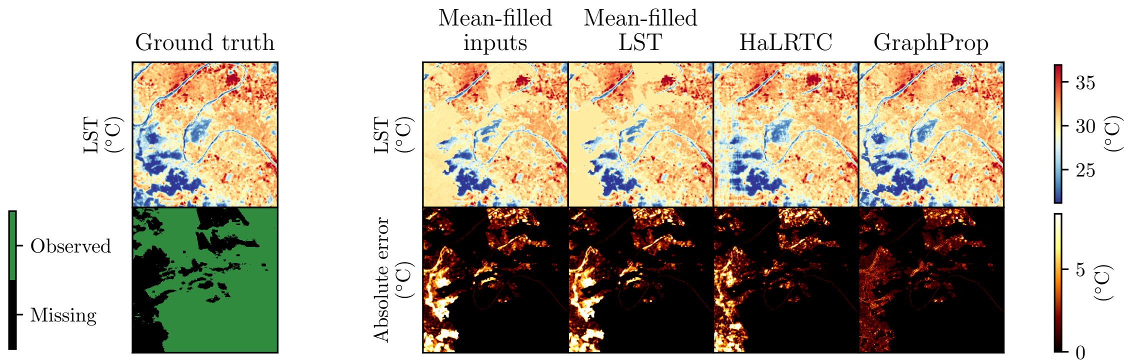
A selection of results are displayed in Figure 4 to illustrate the spatial accuracy characteristics of each method. These absolute error distributions demonstrate that Graph-



(a) 70% cloud cover scenario in Jakarta.



(b) 90% cloud cover scenario in London.



(c) 30% cloud cover scenario in Paris.

Figure 4: Qualitative comparison of the LST imputation methods for at varying amounts of cloud cover.

Prop can recover fine-scale LST variations more accurately than other methods and therefore gives results with fewer neighborhoods of high error magnitude.

Across all the tested methods the errors are larger in the Paris study than for the other two locations. One theory for this behaviour is that the Paris study represents the largest temperature difference between the respective acquisitions (29.6°C versus 11.2°C) and therefore the graph-based representation derived from the colder reference acquisition which was used to complete the partially-observed acquisition may less perfectly characterize the region on the later date. This hypothesis would require further study to prove or disprove conclusively. Nonetheless, the GraphProp results remain the most accurate for the Paris study despite the larger errors observed across methods.

These errors are computed with reference to the values obtained using the SMW algorithm when taking the synthetically-obscured pixel values as inputs. The contribution of Ermida et al. (2020) provides an analysis of the accuracy of this approach with reference to ground truth from in-situ measurements. In their study, performed across 12 locations, they found the RMSE error of the SMW algorithm to be 1.9°C when using Landsat 8 inputs.

4 Discussion and Conclusions

The results presented in this study show that the proposed method, GraphProp, is able to provide accurate LST values in the presence of cloud cover. The method is able to provide more accurate results than the low-rank tensor completion method (HaLRTC) and other more naive mean-filling approaches. The presented results also show this approach to be very robust against the extent of cloud cover present in the scene. The contribution, therefore, represents a useful and practical tool for the analysis of LST in the presence of cloud cover. This will assist applications involving the analysis of LST dynamics, by reducing gaps in time series caused by cloud cover and therefore provide a more complete picture of temporal trends that might otherwise have been difficult to observe.

This study has focused on the analysis of LST in urban areas, however the authors expect similar results to be achievable if applied to other land cover types. The method can also be straightforwardly extended to time series involving more than two acquisitions, for example by combining graphs obtained from each. This would allow for the case where it is not required that any one of the reference acquisitions is fully cloud-free, provided that each location is observed without cloud in at least one of the acquisitions to identify its spectral similarities allowing it to be incorporated into the graph-based representation of the region.

There is scope to extend this research to analyse how the accuracy of the proposed approach varies as other variables such as the temporal or seasonal distance between the respective acquisitions is changed.

Open Research Section

The data used in this study were obtained using the Google Earth Engine platform (https://earthengine.google.com/). The code used to generate the results presented in this paper is available at https://github.com/IMPACTSquad/LST-Gaps.

Acknowledgments
This work was funded through the following sources: the UK Engineering and Physical Sciences Research Council (EPSRC) [grant number EP/T517847/1]; the Centre for Integrated Remote Sensing and Forecasting for Arctic Operations (CIRFA) and the Research Council of Norway [RCN grant number 237906]; the IMPETUS project funded by the European Union Horizon 2020 research and innovation program under grant agreement nr. 101037084; the Isaac Newton Trust; and Newnham College, Cambridge, United Kingdom.

References

Anderson, M. C., Zolin, C. A., Sentelhas, P. C., Hain, C. R., Semmens, K., Yilmaz, M. T., ... Tetrault, R. (2016). The evaporative stress index as an indicator of agricultural drought in brazil: An assessment based on crop yield impacts. *Remote Sensing of Environment*, *174*, 82–99.

Ashraphijuo, M., Aggarwal, V., & Wang, X. (2017). On deterministic sampling patterns for robust low-rank matrix completion. *IEEE Signal Processing Letters*, *25*(3), 343–347.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific data*, *5*(1), 1–12.

Bento, V. A., Gouveia, C. M., DaCamara, C. C., & Trigo, I. F. (2018). A climatological assessment of drought impact on vegetation health index. *Agricultural and forest meteorology*, *259*, 286–295.

Cai, J.-F., Candès, E. J., & Shen, Z. (2010). A Singular Value Thresholding Algorithm for Matrix Completion. *SIAM Journal on Optimization*, *20*(4), 1956-1982. doi: 10.1137/080738970

Chen, Y., He, W., Yokoya, N., & Huang, T.-Z. (2019). Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition. *ISPRS Journal of Photogrammetry and Remote Sensing*, *157*, 93–107.

Ermida, S. L., Soares, P., Mantas, V., Göttsche, F.-M., & Trigo, I. F. (2020). Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series. *Remote Sensing*, *12*(9), 1471.

Fu, P., & Weng, Q. (2015). Temporal dynamics of land surface temperature from landsat tir time series images. *IEEE Geoscience and Remote Sensing Letters*, *12*(10), 2175–2179.

He, W., Yokoya, N., Yuan, L., & Zhao, Q. (2019). Remote sensing image reconstruction using tensor ring completion and total variation. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(11), 8998-9009. doi: 10.1109/TGRS.2019.2924017

Knipper, K. R., Kustas, W. P., Anderson, M. C., Alfieri, J. G., Prueger, J. H., Hain, C. R., . . . others (2019). Evapotranspiration estimates derived using thermal-based satellite remote sensing and data fusion for irrigation management in california vineyards. *Irrigation Science*, *37*, 431–449.

Kondor, Risi and Lafferty, John. (2002, 05). Diffusion kernels on graphs and other discrete input spaces. *ICML, Vol. 2*.

Liu, J., Musialski, P., Wonka, P., & Ye, J. (2013). Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 208-220. doi: 10.1109/TPAMI.2012.39

Masitoh, F., & Rusydi, A. (2019). Vegetation health index (vhi) analysis during drought season in brantas watershed. In *Iop conference series: Earth and environmental science* (Vol. 389, p. 012033).

Mo, Y., Xu, Y., Chen, H., & Zhu, S. (2021). A review of reconstructing remotely sensed land surface temperature under cloudy conditions. *Remote Sensing*, *13*(14). Retrieved from `https://www.mdpi.com/2072-4292/13/14/2838` doi: 10.3390/rs13142838

Morabito, M., Crisci, A., Messeri, A., Orlandini, S., Raschi, A., Maracchi, G., & Munafò, M. (2016). The impact of built-up surfaces on land surface temperatures in italian urban areas. *Science of the Total Environment*, *551*, 317–326.

Nazarian, N., Krayenhoff, E. S., Bechtel, B., Hondula, D. M., Paolini, R., Vanos, J., . . . Santamouris, M. (2022). Integrated assessment of urban overheating impacts on human life. *Earth's Future*, *10*(8), e2022EF002682. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022EF002682` (e2022EF002682 2022EF002682) doi: https://doi.org/10.1029/2022EF002682

Ng, M. K.-P., Yuan, Q., Yan, L., & Sun, J. (2017). An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(6), 3367-3381. doi: 10.1109/TGRS.2017.2670021

Quan, J., Zhan, W., Chen, Y., Wang, M., & Wang, J. (2016). Time series decomposition of remotely sensed land surface temperature and investigation of trends and seasonal variations in surface urban heat islands. *Journal of Geophysical Research: Atmospheres*, *121*(6), 2638–2657.

Rolland, I., Selvakumaran, S., & Marinoni, A. (2023). Remote sensing image completion using a diffusion-based propagation algorithm. In *Image and signal processing for remote sensing xxix* (Vol. 12733, pp. 57–66).

Santamouris, M., Cartalis, C., Synnefa, A., & Kolokotsa, D. (2015). On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings—a review. *Energy and buildings*, *98*, 119–124.

Sobrino, J. A., & Julien, Y. (2013). Time series corrections and analyses in thermal remote sensing. In *Thermal infrared remote sensing: Sensors, methods, applications* (pp. 267–285). Springer.

Srindhuna, M., & Baburaj, M. (2020). Estimation of missing data in remote sensing images using t-svd based tensor completion. In *2020 international conference for emerging technology (incet)* (pp. 1–5).

Wan, Z., & Dozier, J. (1996). A generalized split-window algorithm for retrieving land-surface temperature from space. *IEEE Transactions on geoscience and remote sensing*, *34*(4), 892–905.

Xiao, R., Ouyang, Z., Zheng, H., Li, W., Schienke, E. W., & Wang, X. (2007). Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences*, *19*(2), 250-256. Retrieved from `https://www.sciencedirect.com/science/article/pii/S1001074207600412` doi: https://doi.org/10.1016/S1001-0742(07)60041-2

Yuan, L., Li, C., Mandic, D., Cao, J., & Zhao, Q. (2019). Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 33, pp. 9151–9158).

Zemp, M., Chao, Q., Han Dolman, A. J., Herold, M., Krug, T., Speich, S., … Yu, W. (2022). Gcos 2022 implementation plan. *Global Climate Observing System GCOS*(244), 85.

Zhan, W., Zhou, J., Ju, W., Li, M., Sandholt, I., Voogt, J., & Yu, C. (2014). Remotely sensed soil temperatures beneath snow-free skin-surface using thermal observations from tandem polar-orbiting satellites: An analytical three-time-scale model. *Remote Sensing of Environment*, *143*, 1–14.

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., … Sobrino, J. A. (2018). Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sensing*, *11*(1), 48.

Zou, Z., Zhan, W., Liu, Z., Bechtel, B., Gao, L., Hong, F., … Lai, J. (2018). Enhanced modeling of annual temperature cycles with temporally discrete remotely sensed thermal observations. *Remote Sensing*, *10*(4), 650.