Downscaling Sentinel-3 Data

Edoardo Albergo, University of Padua

31/07/2024

Abstract

The study focuses on downscaling Sentinel-3 Land Surface Temperature (LST) data using Sentinel-2 Normalized Difference Vegetation Index (NDVI) data through Area-to-Area CoKriging. The project involves acquiring and processing raster and vector data for an area above Venice, Italy, with data clipped to the specific study area using a defined polygon. The methodology addresses common geostatistical issues by employing area-to-area kriging, which respects the spatial nature of data collected over areas rather than points. This approach integrates a correlated variable (NDVI) to enhance the downscaling resolution. The study includes extensive data preprocessing, variogram analysis, and the application of the ataCoKriging function, highlighting techniques to mitigate anomalies in the predictions. The results indicate a significant improvement in spatial resolution while maintaining the original data's statistical properties, with validation tests confirming the method's effectiveness.

1 Introduction

Remote sensing technology provides critical data for environmental monitoring and research, with satellite missions like Sentinel-3 offering valuable insights into land surface temperatures (LST). However, the coarse spatial resolution of Sentinel-3 LST data often limits its applicability for detailed local studies. To address this, downscaling techniques can be employed, improving spatial resolution by integrating

higher-resolution ancillary data. This study focuses on downscaling Sentinel-3 LST data using Sentinel-2 NDVI data through Area-to-Area CoKriging, a geostatistical method that leverages spatial correlations between variables. The area above Venice in Italy serves as the study area due to its diverse land cover and the availability of both LST and NDVI data. The chosen date for data acquisition, January 8, 2022, ensures minimal cloud cover, facilitating accurate remote sensing measurements.

2 Methods and Data

2.1 Dataset

This project required the use of raster data and vectorial data. The main source of data were two raster files containing information about the Land Surface Temperature (LST) and the Normalized Index Vegetation Index (NDVI) covering the area of the Veneto region in Italy. One vector layer containing the polygon of the selected area of study has then been produced and used to clip the rasters. The date chosen for the data acquisition is the 08 January 2022 due to the optimal cloud cover of the images.

A summary of the raster tiles' specifications is available in table 1.

Tile name	**Source**	**Resolution**
S2A_MSIL2A_20220108T101411_N0301_R022_T32TQR_20220108T115544.SAFE	Copernicus Browser	20x20m
S3A_SL_2_LST____20220108T093548_20220108T093848_20220109T183249_0179_080_307_2160_LN2_O_NT_004.SEN3	Copernicus Browser	1000x1000m

Table 1: Tile information with source and resolution.

After the download, Sentinel-3 tile has been reprojected to EPSG:32632 using the software SNAP. Both tiles have been clipped using a polygon named "Marina" that has been defined by the author, which represents an area above Venice very close to the sea and has a rich presence of temperature monitoring station that can be used in the future to further validate the results. The clipped rasters are shown in Fig. 1 where the negative values for the NDVI have been removed to improve the visualization

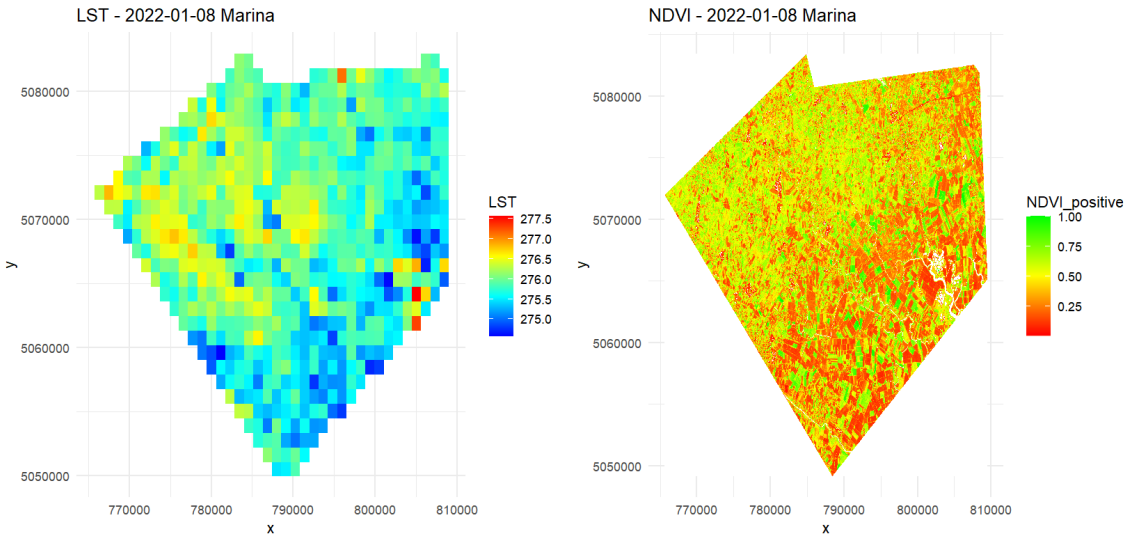

Fig. 1: Clipped rasters, LST and NDVI

2.2 Methods: Area-to-Area CoKriging

2.2.1 The support problem

When interpolating areal data, a common approach in geostatistics is to reduce the area to its centroid and apply point-scale methods[6][11][4][5][3][8]. Although practical, this approach presents several problems. The first problem is physical; the values collected on a specific support, such as a pixel, are not point values but a synthesis (typically a mean) of the values of each point within the support. Ignoring the multipoint nature of each area can lead to underestimating the spatial relationships within a phenomenon and is a rigid approach. Additionally, if the areal units vary in shape and size, calculating distances between centroids may not accurately represent the actual distances between points within the areas[9].

Furthermore, when treating count data as single points, such as sanitary data related to the population in a territory, the small number problem may occur. For example, if a rare event happens in a sparsely populated area, the percentage incidence of that event may be exaggerated, leading to high variability in estimation. Aggregating data on an areal level can reduce estimation variance and stabilize results.

2.2.2 Area-to-area kriging

To address these issues, area-to-area kriging (ATA kriging) can be adopted. This method operates on an area-scale support and considers areal values as a synthesis of the points composing them. This section provides a high-level overview of the area-to-area kriging and CoKriging algorithms. For further details, refer to Goovaert (2008) and Pardo-Igúzquiza et al. (2005).

The ordinary kriging estimation for the value $\hat{z}$ of the areal unit ν_β can be expressed as

$$\hat{z}(\nu_\beta) = \sum_{i=1}^{K} \lambda_i(\nu_\beta) z(\nu_i)$$

Where $\lambda_i(\nu_\beta)$ are the kriging weights related to the area ν_β for the area ν_i.

The difference between a point-scale kriging and an area-to-area kriging lies in the covariance structure. The areal variogram of a stationary process is related to the point-scale variogram by the following equation:

$$\gamma_\nu(h) = \bar{\gamma}(\nu, \nu_h) - \gamma(\nu, \nu)$$

where $\bar{\gamma}(\nu, \nu_h)$ is the average variogram between an arbitrary point in the support ν and another in the translated support ν_h and $\gamma(\nu, \nu)$ is the variance within the points of the ν area.

Goovaert's approach, which is the one used in this paper, leverages this relation to deconvolute the area-scale variogram into a point-scale variogram. The first step is to apply a discretizing grid to the entire area. The discretizing points can have different weights depending on their distance from the centroid. An initial value is imputed to $\bar{\gamma}(\nu, \nu_h)$ (usually the area-scale empirical variogram $\hat{\gamma}_\nu(h)$) and $\gamma(\nu, \nu)$ is estimated by applying the imputed variogram to the discretizing points. Then, $\gamma_\nu(h)$ is calculated following the formula above and the difference index $d_t = |\gamma_\nu(h) - \hat{\gamma}_\nu(h)|$ is computed.

An iterative approach then adjusts the parameters of $\bar{\gamma}(\nu, \nu_h)$ to minimize d. The algorithm stops when the maximum number of iterations has been reached, when d meets a predefined threshold or when d does not decrease for a certain number of

iterations.

Calculating the point-scale variogram enables kriging at the level of discretizing points, modeling the relationships between individual points within the areas and allowing for fine-grained interpolation.

Area-to-area CoKriging integrates this method by exploiting the correlation between the target variable and a correlated variable. In Pardo-Iguizquiza et al. (2006), it is shown that when the target variable needs to be downscaled to a higher resolution, incorporating an ancillary variable at the target resolution improves interpolation results.

The CoKriging model with one covariate can be expressed as

$$\hat{z}(\nu_\beta) = \sum_{i=1}^{K} \lambda_i(\nu_\beta) z(\nu_i) + \sum_{j=1}^{J} \alpha_j(\nu_\beta) y(\nu_j)$$

where $\alpha_j(\nu_\beta)$ are the CoKriging weights for the covariate variable $y(\nu_j)$ relative to the β area for the j^{th} observation.

Introducing the additional variable in the kriging model changes the shape of the covariance matrix C and the covariance vector D. In the ATA kriging method, C contains nxn points representing the covariance between each of the discretizing points, while in the ata CoKriging method C contains (n x $nVars$) + $nVars$ x (n x $nVars$) + $nVars$ points, where $nVars$ is the number of covariates. The additional $nVars$ term at the end of each dimension represents the Lagrange multipliers. The D vector, representing the covariance between the unknown point subject to the estimation and each of the known points, follows a similar logic and is of size (n x $nVars$) + $nVars$.

Area-to-area CoKriging techniques are used in this paper to downscale Sentinel-3 data, using Sentinel-2 aggregated data as an ancillary variable.

3 Implementation

3.1 Descriptive Statistics

As shown in Fig. 1, the values for the LST ranged from a minimum of 274.57 K to a maximum of 277.56, while the NDVI ranged from -1 to +1, with negative values located almost exclusively in the middle-right area, corresponding to the river *Canale Nicesolo* and the above *Laguna di Caorle*. The percentile distribution of the two measures can be observed in Table 2, while the histograms of the values are shown in Fig. 2.

	5%	**25%**	**50%**	**75%**	**95%**	**Mean**	**SD**
LST	275.13	275.56	275.85	276.15	276.49	275.85	0.43
NDVI	0.13	0.24	0.44	0.60	0.77	0.42	0.24

Table 2: Percentile distribution, mean and standard deviation for LST and NDVI

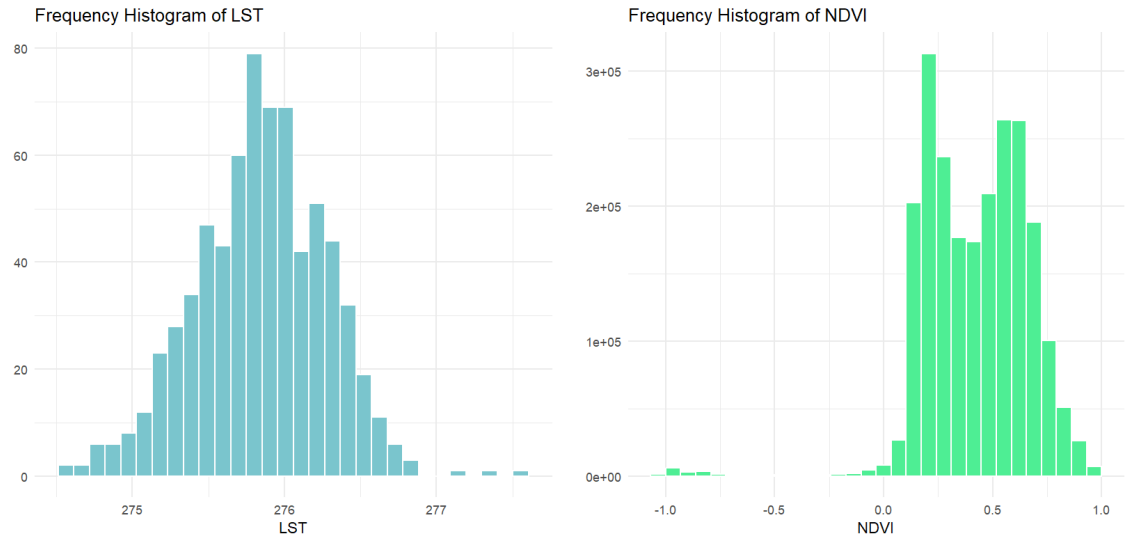

Fig. 2: Frequency histograms for LST and NDVI

The frequency distributions suggest a Gaussian behavior for the LST, while the NDVI seems to be distributed according to a Bimodal Gaussian distribution with

few anomalies on the left tail. Looking at the percentiles of the NDVI it's clear how the area has predominantly positive values.

3.2 Preprocessing

3.2.1 NDVI aggregation

Due to the scarcity of computational resources, the initial intent of downscaling the LST image to a 20m resolution has been set aside. For such reason, the NDVI tiles have been aggregated to obtain a 100m resolution image, which is also the final goal of the downscaling.

To obtain the coarser NDVI image, the *aggregate* function from the *raster* package has been used, inputting a factor of 5 and indicating a mean function for the aggregation.

From now on, any mention of the NDVI will refer to the raster with a 100m resolution.

3.2.2 Variogram investigation

In order to test the spatial process underlying the data, a variogram analysis of the two rasters has been performed. The data have been tested for anisotropy and for stationarity, both resulting in an isotropic process with a clear trend effect on both the x and the y axis.

The resulting variogram models are shown in Fig.3

The best fitting variogram model for LST was found to be a nested variogram combining two Spherical models with sill = 0.080 and range = 3159, and sill = 0.027 and range = 17559. The fitted nugget effect is 0.042.

For the NDVI, a Spherical model with sill = 0.006, range = 3653, and nugget effect = 0.030 has been fitted.

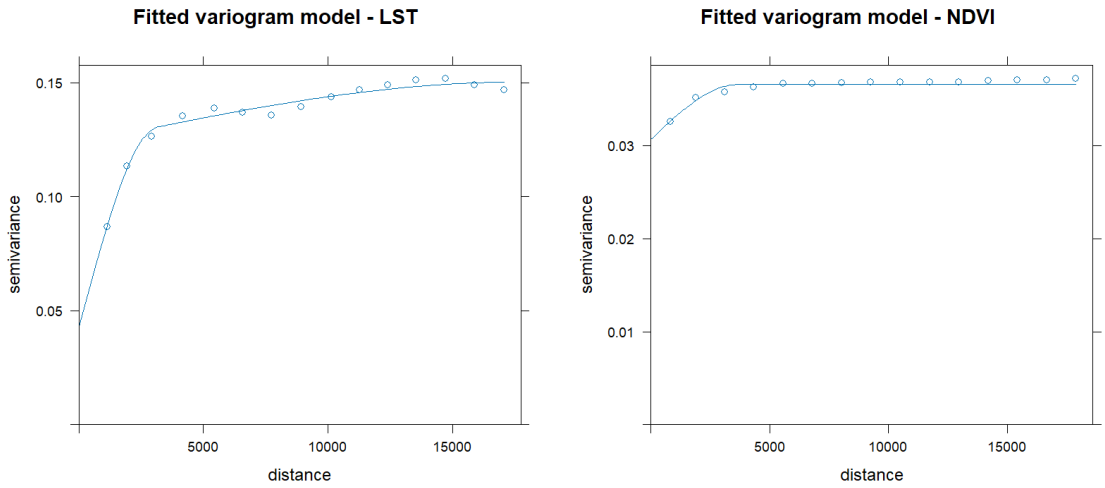

Fig 3: Fitted variogram models for LST and NDVI

3.2.3 Detrending

Since the stationarity of the processes couldn't be assessed in the variogram analysis and given that one of the assumptions for the application of the CoKriging model is that the input data are stationary, a trend removal procedure has been applied. Both the LST and the NDVI have been modeled using a linear regression [1] of the form:

$value \sim \beta_0 + \beta_1 x + \beta_2 y + \epsilon$ where x and y represent the x and y cohordinates. The idea behind this type of detrending technique is to remove any variation in the mean of the process that can be expressed as a function of its position. The residual part will then be modeled as the stationary spatial process underlying the data.

The histograms of the residuals can be seen in Fig 4.

Both distributions exhibit a Gaussian-like shape and are symmetric around the zero value, meeting the requirements for the application of the CoKriging model.

3.3 Modelling

The *atakrig* package [10] is an R package designed by Maogui Hu for multivariate area-to-area and area-to-point kriging predictions, offering robust geostatistical

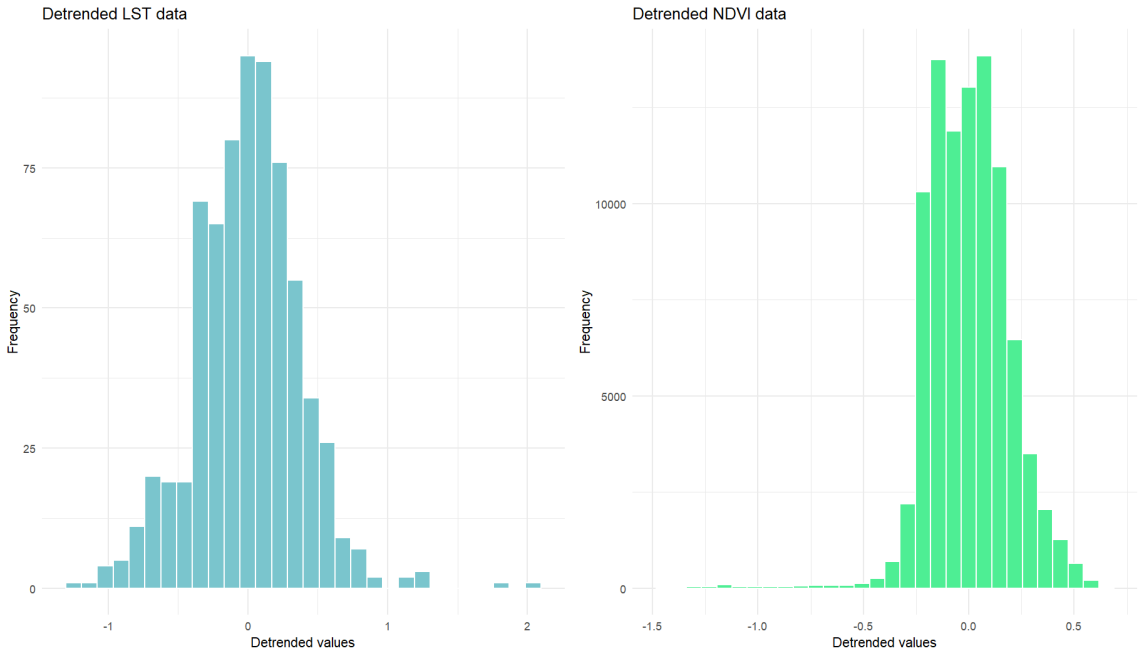

Fig 4: Histograms of the detrended data

methods for spatial data interpolation. It supports a wide range of applications, including environmental science, ecology, and hydrology, by providing tools for spatial data conversion between different scales.

The workflow that leads to the downscaled results consists of three steps: raster discretization, deconvolution of the variograms, and area-to-area CoKriging.

3.3.1 Rasters Discretization

The *discretizeRaster* function in *atakrig* is used to convert raster data into fine resolution points, a process known as discretization. This function takes a raster object and divides it into smaller grid cells of a specified size, such as 100 meters in this case. Each cell's center coordinates and values are extracted, and a point spread function (PSF) is applied to assign weights to the discretized points. The Gaussian PSF, specified by the *psf* parameter, ensures that the weights follow a Gaussian distribution, providing a smoother representation of the raster data. Using Gaussian weights instead of equal weights is advantageous in this case because it emphasizes

the central values of the raster cells, reducing the influence of outlier points and capturing spatial patterns more accurately. This results in a more realistic interpolation of spatial variables, particularly when dealing with Sentinel data[7].

While *atakrig* requires both the LST and NDVI data to be discretized, the Gaussian weights apply only to the LST raster. This is because the NDVI raster is already at a 100m resolution, so only one discretizing point falls within each pixel.

3.3.2 Variograms deconvolution

After the LST and the NDVI data have undergone the discretization process, the deconvolution of the variogram, handled by the *deconvPointVgmForCoKriging* function, is applied. This function deconvolutes point-scale variograms and cross-variograms from spatial areal samples, allowing for the accurate modeling of spatial variability at finer scales. In practice the function iteratively adjusts the parameters of the point-scale variogram to match the experimental variogram derived from the areal data. Initially, an area-scale semivariogram is fitted to the areal data, and an initial guess for the point-scale semivariogram is made. Following Goovaerts's approach[9], the areal-scale semivariogram is then regularized using this initial guess, and the difference between the modeled and experimental semivariograms is calculated. The parameters of the point-scale semivariogram are iteratively adjusted in order to minimize the difference between their areal reaggregation and the theoretic areal variogram. Such difference is calculated using the mean absolute error for each of the variograms' bin. This iterative process ensures that the deconvoluted variogram aims to represent the underlying spatial structure of the data, estimating the spatial correlations on a finer scale.

The deconvolution process, whose output can be seen in Fig. 5, produces the deconvoluted point variogram for LST and NDVI, as well as their deconvoluted cross-variogram.

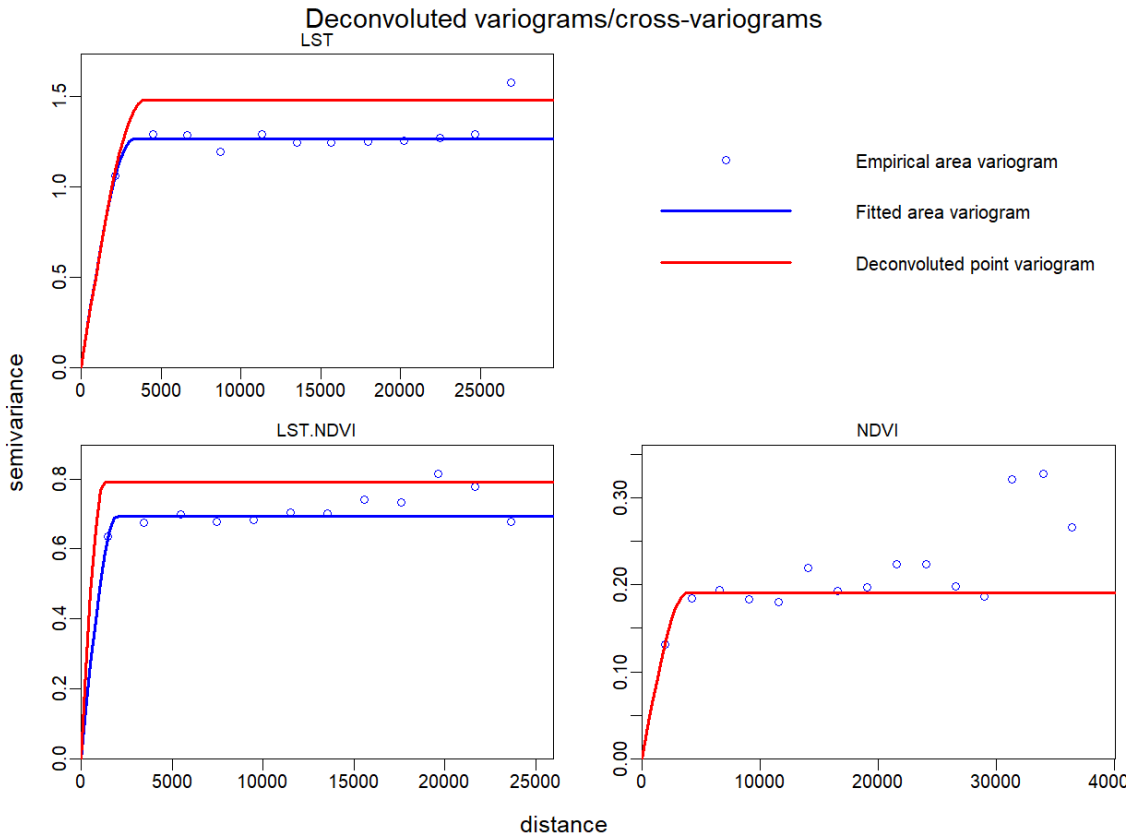

Fig 5. Deconvoluted variograms and cross-variogram

3.3.3 Area-to-Area CoKriging

Once the rasters have been discretized and the deconvoluted variograms computed, the *AtaCoKriging* function can be applied. The function is called with the *nmax* parameter set to 100, indicating that a maximum of 100 neighboring areas are used for the interpolation. The CoKriging method leverages the spatial correlations between the known data points (LST and NDVI) and the unknown points to predict the areal values of an empty raster, generated using the frame of the NDVI raster. After calculating the matrix of distances, the covariance matrix between the known areas (C) and the covariance vector between the unknown area and the known areas (D) are determined using the deconvoluted variograms. The function then operates

recursively on each area, considering only the number of neighboring area indicated by *nmax*. The CoKriging weights are ultimately calculated by solving the following system of linear equations: $W_{CoKriging}C = D$. Solving this equation requires the inversion of the C matrix, which is a complex operation due to the matrix's characteristics. This topic will be further discussed in the next paragraph regarding anomalies produced by the model.

Once the CoKriging weights vector has been computed, the resulting values are used to obtain the LST (detrended) estimation for the single 100m area considered. A recursive application on each of the unknown areas returns the predicted Raster, which is then retransformed into LST in Kelvin degrees using the parameters derived from the linear model used for detrending.

3.4 Anomalies Reduction Techniques

The *atakrig*'s *ataCoKriging* function produced, on average, 2% anomalous predictions, ranging between -1600 and +7800, which are completely out of scale compared to the expected values between 274 and 280 Kelvin degrees. Therefore, some mitigation techniques were adopted. The first approach was to test a range of *nmax* values from 5 to 100. This operation did not affect the number of anomalies or the magnitude of the anomalous values.

The second approach considered the quality of the sample values, so different preprocessing techniques were adopted: rescaling the values to a [0,100] interval, rounding to the third decimal for both LST and NDVI, and applying standard Kriging to exclude a problem in the interaction between NDVI and LST. Rounding to the third decimal proved to be the most effective technique, reducing the magnitude of the anomalies to the interval [-700, +1200]. Still, 2% of the values were completely unacceptable.

Comparing the results produced by standard Kriging performed by *atakrig* and *gstat*, it became evident that the main issue resided in how *atakrig*'s function produced the estimates. For this reason, a copy of the package's code was made to enable modifications to the functions. By introducing some monitoring messages inside the

function, it was possible to assess that the C matrix (covariance between the known areas) was ill-conditioned, with a conditioning number $k \sim 88$. To address this, a small normalization constant was added to the determinant of C. Additionally, the *Matrix*[2] package was used to resolve the system of linear equations to determine the CoKriging weights. This package adopts specific methods for sparse matrices, avoiding the explosion of small values during the C matrix inversion.

These amendments helped reduce the percentage of anomalous predictions to 0.6%, with a range of predicted values between 240 K and 318 K. While some of these values are still unrealistic, the benefit of adopting matrix normalization and an algorithm specific to sparse matrices is evident.

No further actions to mitigate anomalies were taken, and the anomalous values were filtered out from the dataset. Fig. 6 shows the location and magnitude of the anomalies. It is easy to notice that the vast majority of the points are located at the border of the area, with central points being much smaller than those located on the perimeter.

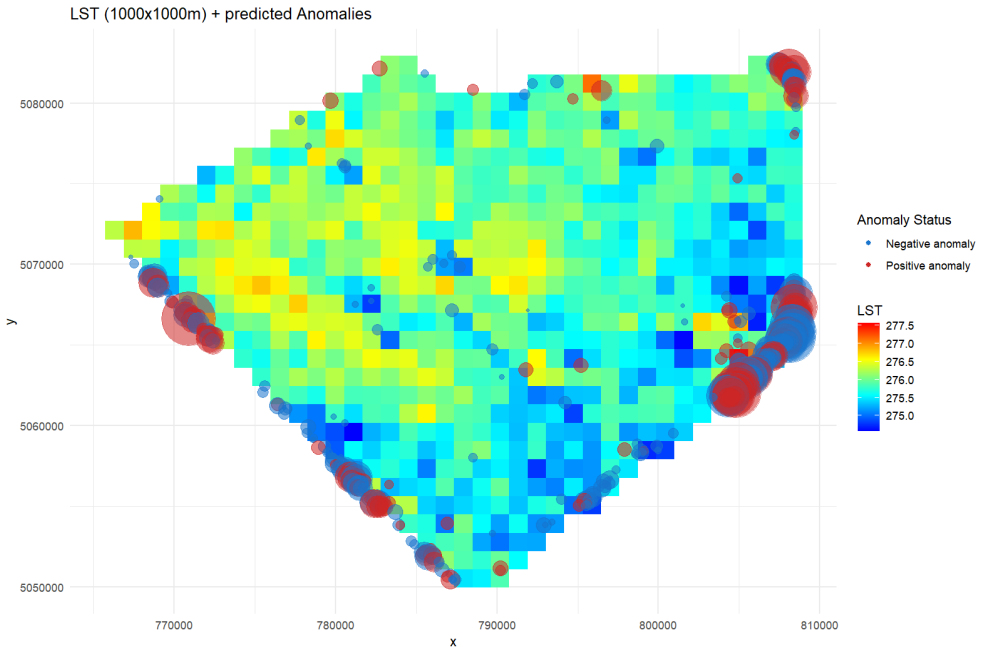

Fig. 6: LST raster vs anomalous predictions. The size of the bubble is proportional to the magnitude of the anomaly

4 Results

4.1 Outcome of the model

Fig. 7 shows the comparison between the original and the downscaled raster. It can be noticed that the downscaled model reproduces the two warmer areas at the top-left of the raster accurately. However, the area at the center-right, corresponding to the *Canale Nicesolo* and the *Laguna di Caorle*, is where the model has more issues producing coherent predictions, as confirmed by the high number of anomalies observed in Fig. 6.

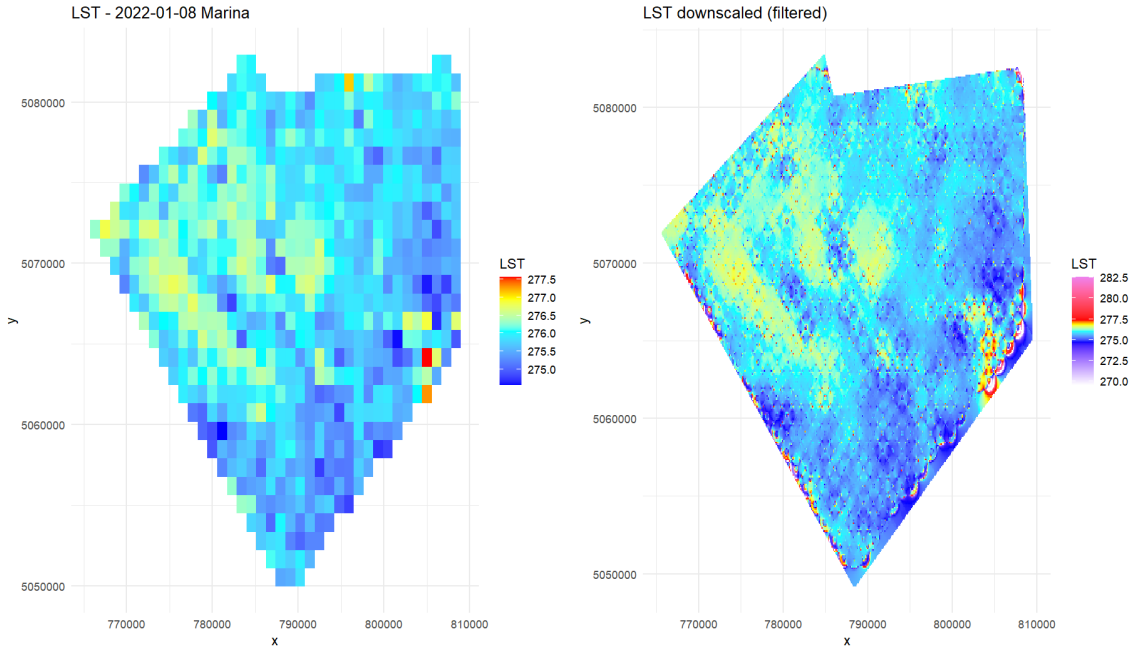

Fig. 7: Original vs downscaled LST rasters. Homologated color scales

4.2 Validation

While it is very difficult to validate results of a downscaling process, given that no benchmarks are available, it is still possible to conduct a visual comparison (such as the one proposed in Fig. 7) to check that the downscaling respects the original

macro-pattern. Figure 8 presents a comparison of the histograms of the original raster and the downscaled raster values. It is noticeable that, while the downscaled raster exhibits a much smoother Gaussian shape, the two distributions are quite similar. The average values for the original and downscaled data are 275.8436 and 275.8408, respectively, and a t-test could not reject the null hypothesis of different means for the two samples (t = -0.1743, df = 713.62, p-value = 0.8617). A similar result is produced by the Mann-Whitney test, which rejects the null hypothesis that the two samples come from different distributions (W = 31938445, p-value = 0.9604).

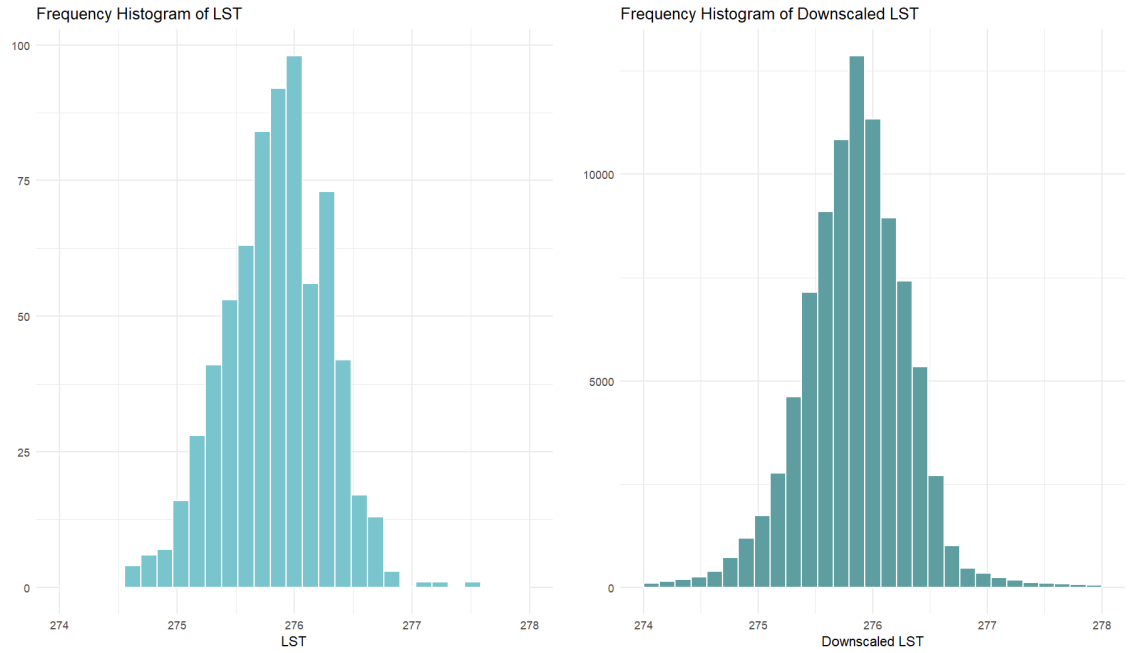

Fig. 8: Original vs downscaled LST histograms.

5 Conclusions and further improvements

The study successfully demonstrates the potential of Area-to-Area CoKriging for downscaling Sentinel-3 LST data using Sentinel-2 NDVI data. The methodology effectively enhances the spatial resolution of LST data while preserving the statistical integrity of the original dataset. Despite some anomalies in the predictions, the

integration of matrix normalization and specific algorithms for sparse matrices significantly reduces their occurrence. Future research should focus on refining anomaly mitigation techniques and exploring additional ancillary variables to enhance the downscaling process. This may also require modifying the variogram fitting procedure, which currently does not support nested variogram structures. Additionally, implementing area-to-area CoKriging algorithms for non-stationary data would be highly beneficial. The findings highlight the value of advanced geostatistical methods in remote sensing applications, offering a framework for improving the spatial resolution of satellite-derived environmental data.

References

[1] ArcGIS Documentation. *Understanding how to remove trends from the data*. Visited on 15/06/2024. 2024. URL: https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/understanding-how-to-remove-trends-from-the-data.htm.

[2] Douglas Bates, Martin Maechler, and Mahendra Jagan. *Package 'Matrix'*. 2024.

[3] Olaf Berke. "Exploratory disease mapping: kriging the spatial risk function from regional count data". In: *International Journal of Health Geographics* 3.1 (2004), p. 18. DOI: 10.1186/1476-072X-3-18.

[4] George Christakos and Jian Lai. "A study of the breast cancer dynamics in North Carolina". In: *Social Science & Medicine* 45.10 (1997), pp. 1503–1517.

[5] George Christakos and Marc L Serre. "Spatiotemporal analysis of environmental exposure-health effect associations". In: *Journal of Exposure Analysis & Environmental Epidemiology* 10.2 (2000), pp. 168–187.

[6] Noel Cressie. *Statistics for spatial data*. Wiley, New York, 1993, p. 900.

[7] ESA. *ESA Sentinel Wiki*. Visited on 17/06/2024. 2024. URL: https://sentiwiki.copernicus.eu/web/s2-mission.

[8] Pierre Goovaerts. "Geostatistical analysis of disease data: estimation of cancer mortality risk from empirical frequencies using Poisson kriging". In: *International Journal of Health Geographics* 4.1 (2005), p. 31. DOI: 10.1186/1476-072X-4-31.

[9] Pierre Goovaerts. "Kriging and Semivariogram Deconvolution in the Presence of Irregular Geographical Units". In: (2008). Unpublished.

[10] Maogui Hu. *Package 'atakrig'*. 2023. URL: https://cran.r-project.org/web/packages/atakrig/atakrig.pdf.

[11] Michael A Oliver et al. "Binomial cokriging for estimating and mapping the risk of childhood cancer". In: *IMA Journal of Mathematics Applied in Medicine and Biology* 15.3 (1998), pp. 279–297.