Comparison of direct and inverse methods of satellite observations downscaling for the coastal zone area

Andrzej CHYBICKI, Zbigniew ŁUBNIEWSKI

Gdansk University of Technology
Faculty of Electronics, Telecommunications and Informatics
Narutowicza 11/12, 80-233 Gdansk, Poland
Andrzej.Chybicki@eti.pg.gda.pl

The Earth observation satellite imaging systems have known limitations, especially regarding their spatial and temporal resolution. Therefore, approaches which aim to combine data retrieved from sensors of higher temporal and lower spatial resolution with the data characterized by lower temporal but higher spatial resolution are of high interest. This allows for joint utilization of the advantages of both these types of sensors. As there are several ways to achieve this goal, in this paper two approaches, direct and inverse, of downscaling the land surface temperature (LST) derived from low resolution imagery acquired by the Advanced Very High Resolution Radiometer (AVHRR) were evaluated. The applied downscaling methods utilize biophysical properties of the surface sensed using short wave infrared and thermal band. The presented algorithm evaluation was performed on the basis of a specific test case: the coastal zone area of the Gulf of Gdańsk, Poland. In this context, the objective presented in the study was to compare two methods of downscaling for a specific test case in order to evaluate how the proposed approaches cope with the specific conditions of the coastal zone area.

Keywords: Earth observation, land surface temperature, sea surface temperature, PBIM downscaling method, surface effective emissivity

1. Introduction

Remote satellite Earth Observation (EO) systems are capable of providing valuable information related to environmental monitoring. In particular, observations acquired from visible, Shortwave Infrared (SWIR) and Thermal Infrared (TIR) electromagnetic band sensors are of high interest as they allow one to monitor and explore several phenomena taking place in the atmosphere or on the Earth's surface.

Most important limitations regarding EO observations are related to spatial and temporal resolution of data. The spatial resolution generally refers to the way the observation data is stored in a grid. In the digital form of observation, each pixel is considered to represent observation of a particular area. The size of a single pixel is dependent on the EO system design, and determines the resolution of the image. The second term, temporal resolution, specifies the visiting frequency of a satellite sensor over a specific location providing information on the time period between two successive acquisitions from the same area.

In general, higher spatial resolution of the sensor causes swath observation to be narrower, and thus exceeds the revisit time, and vice versa. Therefore, none of the available space-borne observation systems, allow frequent imaging at high spatial resolutions. Since the spatial and temporal resolution of imaging systems, e.g. in the thermal infrared band, are negatively correlated, research on combining several observations in order to virtually improve both spatial and temporal imaging resolution, is of high interest.

Generally, the process of combining data and observations characterized by different spatial and temporal resolution is known as downscaling or sub-pixel mapping [1][2]. The goal of these methodologies is to retrieve datasets that are a combination of high spatial resolution imagery, with datasets characterized by lower spatial resolution, but being more frequently registered. The appropriate combination of both sources enables one to generate a new product, because the imagery is more frequently delivered, and has higher spatial resolution. In consequence, this enables the potential use of downscaled products in microclimate studies, human activities monitoring, urban heat island effects observations [3], urban and peri-urban areas canopy characterization or urban climate energy budget estimation [4].

In the paper, the evaluation of two approaches: direct and inverse, to the downscaling procedure was presented. In order to present the evaluation results, the authors of the paper prepared a specific test case for the investigated downscaling methods. Direct and inverse methods of downscaling were tested for satellite acquisitions of the coastal zone area for the 13th of September, 2013. Since the method of downscaling refers to the biophysical properties of the imaged surface, the purpose of the research was to evaluate the performance of the presented approaches, and compare them in terms of Root Mean Squared Error (RMSE), correlation coefficient (R) and visual comparison to time co-incident high resolution Landsat 8 TIRS LST, which was treated as a validation dataset.

2. Theory

Downscaling is the process of increasing the spatial resolution of a set of spatial information (e.g., 2- or 3-dimensional). In the case of satellite acquisition, downscaling relies on the observational data spatial resolution increase obtained by intelligent merging of at least two datasets [5][6]. This methodology can be described as follows: let us assume two input data sources - such that both sources represent acquisition over the same area, and that the contribution of each surface cover component of low-resolution thermal pixels can be

estimated as the weighted sum for all subareas corresponding to pixels in finer resolution contained by the coarse resolution pixel area. It can be expressed as:

$$LST_{1000} = \sum_{n=1}^{N} LST_{100}^{(n)} \frac{s_{100}^{(n)}}{S} \tag{1}$$

where

$$S = \sum_{n=1}^{N} s_{100}^{(n)} \tag{2}$$

where: LST_{1000} - is the land surface temperature of the pixel in coarse ($\sim 1000m\ x\ 1000m$) resolution, $LST_{100}^{(n)}$ - is the LST value of the *n*-th pixel in a high resolution ($\sim 100m\ x\ 100m$) image belonging to the mentioned high pixel resolution area, $s_{100}^{(n)}$- is the area of the given pixel in high resolution, *S* - is the area of a single pixel in coarse resolution, *N* - is the number of pixels in high resolution imagery that are contained in the corresponding pixel of a low resolution image.

In general, N depends on the ratio of data sources spatial resolution taken under consideration, and may also depend on the number of high resolution pixels being excluded from the analysis for a given low resolution pixel due to cloud presence and other limitations. The graphical overview of this approach is presented in Fig. 1.

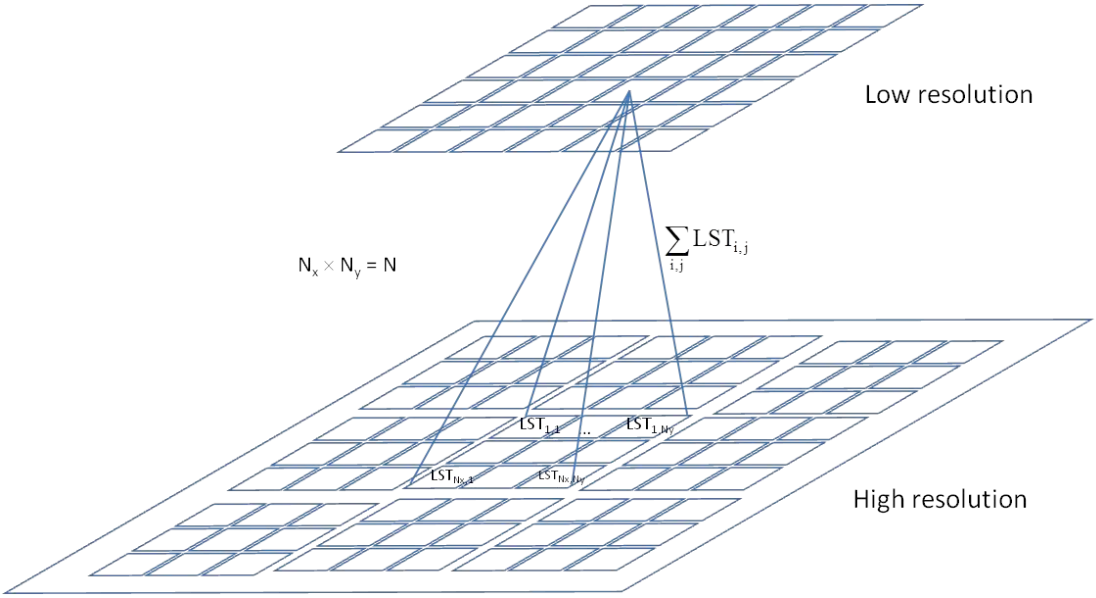

Fig. 1. Diagram representing the relation between spatial distribution of a measured quantity in high resolution and low resolution.

The dependency mentioned above describes the relation between high and low scale spatial datasets. However, according to (1), the transformation of pixel values from high to low scale results in the loss of information, so the reverse transformation is obviously ambiguous. Therefore; additional, high scaled information about the characteristics of a specific area, that influences the pixel values LST_{100} and LST_{1000}, i.e. some independently

measured quasi-static quantity, usually derived from satellite imagery or an auxiliary database, must be provided as well. This can be achieved by assigning a representative spectral emissivity value to each surface cover type, or by using dedicated models for estimating the relation between spectral response function and LST. In that case, the formula (3) below, proposed in [7], is used: which assumes a dependency between effective emissivity ε and land surface temperature.

At this point, there are two ways to proceed with the problem of finding the best high resolution estimate for a given low-resolution input dataset:

1. Direct method (DM), that a priori assumes the dependence between effective emissivity and land surface temperature. It takes into account the quasi-linear dependence between surface effective emissivity ε and LST expressed as [7]:

$$LST_{100} = \frac{LST_{1000} \cdot \varepsilon_{100}}{\varepsilon_{100 \to 1000}}, \tag{3}$$

where ε_{100} is high resolution effective emissivity and $\varepsilon_{100 \to 1000}$ is average effective emissivity for low resolution pixels.

Effective emissivity computation was based on a high resolution Normalized Difference Vegetation Index (NDVI) over the respective region. The maximum NDVI ($NDVI_{max}$) and the minimum NDVI ($NDVI_{min}$) values for the study area were determined, which were further used for computing the fractional vegetation cover [8][9]:

$$\mathrm{FVC} = \left(\frac{\mathrm{NDVI} - \mathrm{NDVI_{min}}}{\mathrm{NDVI_{max}} - \mathrm{NDVI_{min}}}\right)^2 \tag{4}$$

and the effective emissivity ε was calculated as:

$$\varepsilon = 0.98(1 - \mathrm{FVC}) + 0.93\ \mathrm{FVC}. \tag{5}$$

2. Inverse method (IM) that aims to find the min-error solution for the aforementioned dependence (1) between high resolution and low resolution pixel values for the whole imagery.

In the inverse approach, it is assumed that there exists some dependence between effective emissivity and LST expressed as:

$$\mathrm{LST}_{100} = \mathrm{f}(\varepsilon_{100}), \tag{6}$$

where f(·) need not be linear (as it is assumed in the direct method). Combining (1) and (3) yields:

$$\mathrm{LST}_{100} = \sum_{n=1}^{N} \mathrm{f}\left(\varepsilon_{100}\right)\frac{\mathrm{S_n}}{\mathrm{S}} \tag{7}$$

for a given low scale pixel. The form of f(·) is generally unknown for a given terrain, but assuming that it is stable for the entire scene, and applying (7) to each low scale pixel, we

may express the problem as a system of linear equations. For this purpose, let us calculate the histogram $[h_1, h_2, ..., h_K]$ of $\varepsilon_{\text{eff high}}$ (where h_k, k = 1, ..., K is the percentage of high scale pixels whose ε value belongs to a given range) for each low scale pixel, using always the same ε minimum/maximum values and the same number of bars K. Then, if we express f(·) as a vector of values (weights) $[w_1, w_2, ..., w_K]$ corresponding to the central values of $\varepsilon_{\text{eff high}}$ ranges used in calculating the histogram, and assuming that the single pixel area s_n is constant, we may write (7) as:

$$\mathrm{LST}_{1000}^{(m)} = \sum_{k=1}^{K} h_k^{(m)} \cdot w_k \,, \tag{8}$$

where *m* is the number of a "coarse" pixel in a scene *(m = 1, ..., M)*. When iterating through all M pixels of low resolution imagery, we obtain the system of M equations (6) presented in matrix form:

$$\mathbf{LST_{1000}} = \mathbf{H} \cdot \mathbf{w}, \tag{9}$$

where $\mathbf{LST_{1000}}$, is the column vector, **H** is a matrix and *w* is the column vector to be found. The solution to this problem is the following estimate vector $\widehat{\boldsymbol{w}}$:

$$\widehat{\boldsymbol{w}} = \boldsymbol{x_0} + (\boldsymbol{H}^T\boldsymbol{H} + \boldsymbol{\Gamma})^{-1}(\boldsymbol{H}^T(\boldsymbol{LST_{1000}} - \boldsymbol{H}\boldsymbol{x_0})) \tag{10}$$

where $\boldsymbol{x_0}$ is the expected solution, calculated from the direct method (3). We used the simple form of Tikhonov regularization $\boldsymbol{\Gamma} = \lambda\mathbf{I}$, where **I** is the identity matrix and the λ value (regularization parameter) was found using the near-optimal parameter calculation method proposed by [10].

3. Experiment setup

In order to verify and evaluate the proposed algorithms of downscaling, we have prepared the following experiment. We have created the composite 100m (high resolution) dataset Landsat 8 OLI 2 scenes registered on 2013-08-05 at 09:45:37 UTC (scene center time) and on 2014-07-23 at 09:43:29 UTC (scene center time) and we have combined it with Corine Land Cover database (100m resolution) in order to retrieve the effective emissivity (ε) composite product. In the experiment, this quasi-static high resolution dataset was used as input data together with temporal low resolution ($\sim 1000\ m$) MetopB/AVHRR/3 LST, which can be generally registered several times per day.

In order to evaluate results obtained using the investigated approaches, we used Landsat 8 TIRS LST product registered on 2013-08-05 at 09:45:37 UTC (scene center time). Then we compared the output of the downscaling algorithms, i.e. the downscaled LST, with Landsat 8 TIRS LST using visual comparison, cross plots, root mean squared error, and Pearson correlation coefficient of datasets. An overview diagram showing the processing and validation strategy of the proposed approach is presented in Fig. 2.

Data acquisition in all presented cases was made under clear sky conditions. For calculating LST we applied a split window technique using two thermal bands according to the following formula:

$$\mathrm{LST} = T_i + c_1 \cdot (T_i - T_j) + c_2 \cdot (T_i - T_j)^2 + c_0 + (c_3 + c_4 W)(1 - \bar{\varepsilon}) + (c_5 + c_6 W)\Delta\varepsilon$$

where W is the water vapor content retrieved from an external data source, T_i and T_j are thermal bands for TIRS and AVHRR/3 respectively, $\bar{\varepsilon}$ is surface average emissivity and $\Delta\varepsilon$ is surface emissivity difference for channels i and j, and $c_0 - c_6$ values are the split window coefficients determined from algorithm calibration [11], as presented in Table. 1.

The proposed method was tested for the case study of the coastal zone of the Gulf of Gdańsk, Poland, Southern Baltic area. The study area was extracted as a part of the land situated in the coastal zone, i.e. no further than 5 km from the coastline, which is presented in Fig. 2.

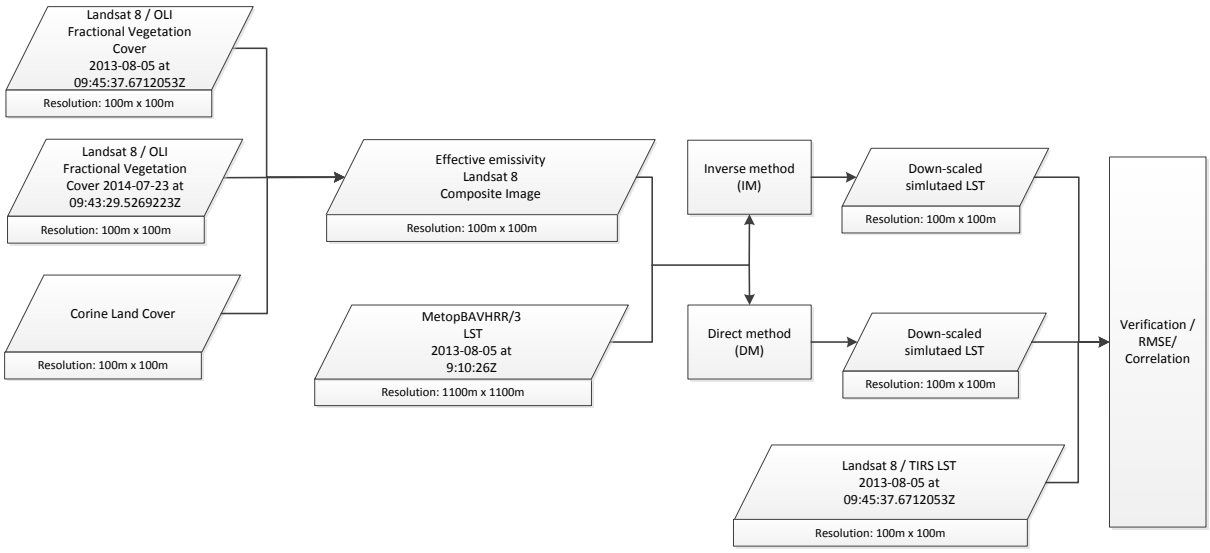

Fig. 2. Overall diagram of data processing flow and verification.

Tab. 1. LST estimation algorithm parameters for MetopB/AVHRR/3 and Landsat 8 TIRS.

Platform/ sensor	λ_i-λ_j	c_0 [K]	c_1[-]	c_2 [K^{-1}]	c_3 [K]	c_4 [$Kcm^2\ g^{-1}$]	c_5 [K]	c_6 [$Kcm^2\ g^{-1}$]
METOP B/ AVHRR3	10.82-11.97	-0.045	1.733	0.307	44.3	0.61	-150	18.7
Landsat 8/ TIRS	10.8-12	-0.268	1.378	0.183	54.3	-2.238	-129.2	16.4

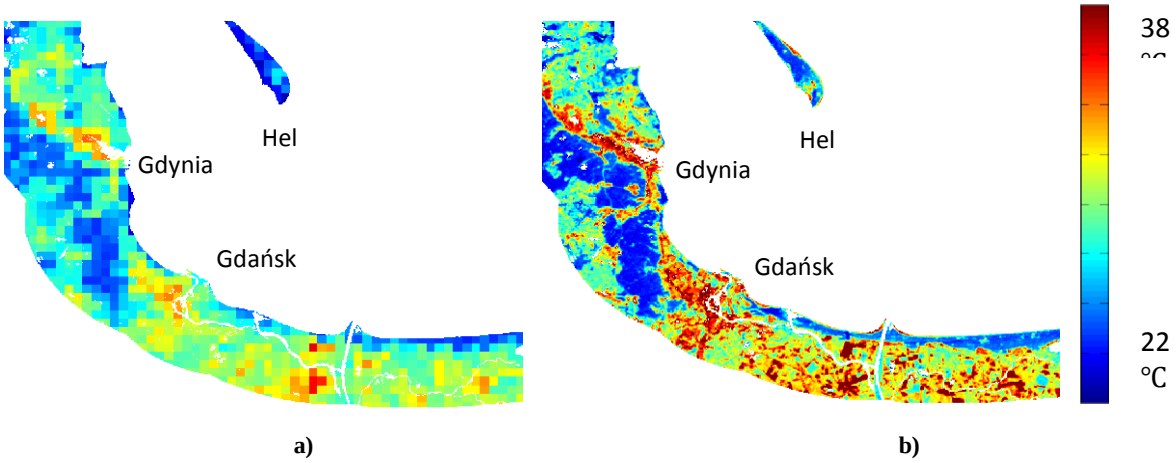

Fig. 2. Analyzed area - 5km coastal zone of the Gulf of Gdańsk. Left picture (a) shows low resolution ($\sim 1000m$) input dataset to be downscaled - MetopB/AVHRR/3 LST. Right picture (b) shows proxy validation dataset - ($\sim 100m$) Landsat 8 OLI LST.

4. Results

In this section we present the results obtained by two methods downscaling the analyzed area, namely DM and IM, using MetOpB/AVHRR/3 1000 m LST as an input dataset. The output of the algorithms is the simulated downscaled 100m LST product.

Fig. 3 presents the map visualizations of the obtained DM and IM LST downscaling results (Fig. 3d and f) along with the scatter plots of the downscaled LST values Vs reference values from 100m Landsat 8 OLI LST (Fig. 3c and e). For comparison, the map of the original low scale MetOpB/AVHRR/3 1000m LST (Fig. 3a) as well as the map of the reference high scale Landsat 8 OLI 100m LST (Fig. 3b) are also presented.

As it can be observed in scatter plots (presented on left side of Fig. 3), the correlation between the high resolution proxy validation dataset and the downscaled product, is visible. However, it can be seen that since DM is not optimized (it a priori assumes the linear character of the dependence between ε and LST), it underestimates (particularly for higher temperature pixels) the downscaled LST value. Note that this is not the case when using the inverse method (IM).

The deeper analysis of spatial distribution for the obtained LST in the investigated area shows that the results obtained by the IM algorithm better represent thermal characteristics for the area, and are visually closer to the original Landsat 8 LST than the product obtained using DM. Specifically, it can be observed that the influence of local surface high resolution characteristics is reflected in IM rather than in local observations in LST retrieved from AVHRR. What is more, as in DM the LST_{1000} term is directly used in the formula and strongly influences the obtained LST_{100} value, it results in the appearance of coarse resolution artifacts in the downscaled product, which effect does not occur in the IM result case.

Table. 2 presents analytical results obtained by the evaluated approaches (IM and DM rows in a table), and original AVHRR/3 data resampled using a nearest neighbor technique (presented in the third row). Second column of the table represents Pearson correlation coefficient (R), the third represents RMSE with respect to reference LST Landsat 8/OLI LST dataset. The last column shows average BIAS between reference dataset and the result of downscaling products, in order to verify if the methods retain radiometric properties of the results. Results presented below, show that the IM method yields results closest to Landsat 8/OLI LST validation dataset, having overall lowest RMSE and the highest correlation

coefficient among the presented approaches. It is generally consistent with conclusions which may be made from visual analysis of the results presented in Fig. 3. For the considered test case, IM yields 23% RMSE reduction in comparison to DM, and correlation increases from 0.7438 to 0.8486. When comparing IM to nearest neighbor resampling, the inverse technique approach yields 27% RMSE reduction and an 18% correlation increase (from 0.6987 to 0.8486).

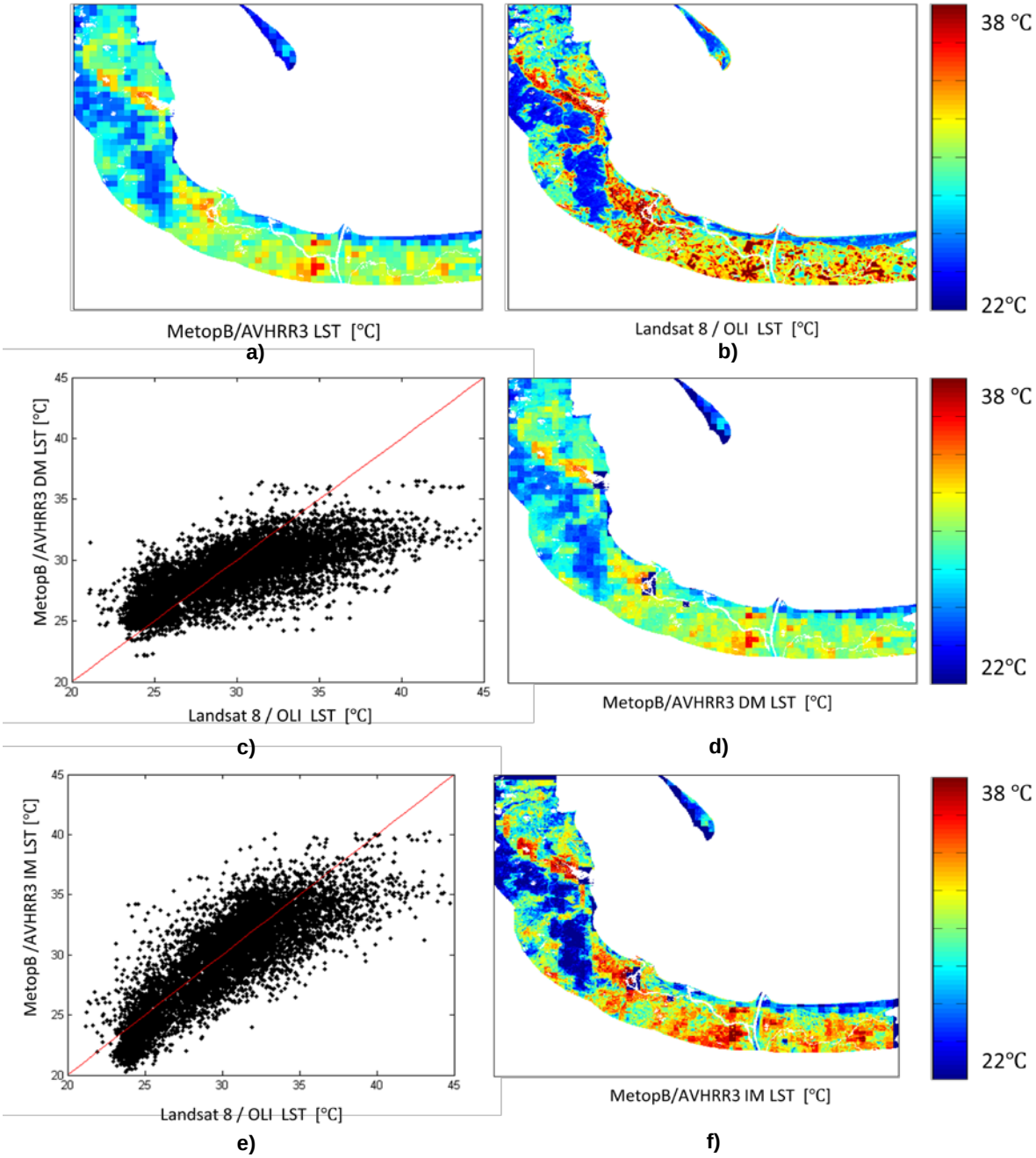

Fig. 3. Scatter plots and map visualizations of the results obtained by DM and IM LST downscaling for the investigated area: a) low scale MetOpB/AVHRR/3 1000m LST, b) high scale Landsat 8 OLI 100m LST used as reference, c) scatterplot of Landsat 8 OLI LST vs. DM LST downscaling results, d) map visualization of DM LST downscaling results, e) scatterplot of Landsat 8 OLI LST vs. IM LST downscaling results, f) map visualization of IM LST downscaling results. The red line in the scatter plots represents 1:1 ratio.

Tab. 2. Analytical results of evaluated IM and DM approaches.

Downsampling method	Correlation coefficient R with respect to Landsat 8 OLI LST	RMSE [°C] with respect to Landsat 8 OLI LST	BIAS[°C] with respect to Landsat 8 OLI LST
IM	0.8486	2.2970	0.5637
DM	0.7438	2.9934	1.0318
Nearest neighbor	0.6987	3.1361	1.0318

5. Conclusions

In the paper, the comparison of direct and inverse methods of downscaling AVHRR LST was presented. The study was conducted on the basis of a prepared experiment for the coastal zone area, where we used time co-incident 100m Landsat 8 TIRS LST as a validation dataset. Results of evaluated approaches show that the method based on inverse problem technique yields better results, in comparison to the direct approach. The reason for that, in the analyzed case, is that surface biophysical properties like soil and air humidity, heat capacity, surface attributes etc. are spatially diverse which causes non-linearity and deviation in the vegetation index – land surface temperature (VI-LST) dependence within the analyzed area. This is particularly observable in the coastal zone area, where proximity of the sea results in the arising of specific VI-LST conditions. In this context, the optimization is particularly effective as it finds minimum-error of a linear model that connects low resolution and high resolution datasets.

Although the inverse problem approach yields much better results in the presented case, it should be pointed out that this approach is much more sensitive to distortions resulting from presence of clouds and aerosols, atmospheric effects etc. This is due to the fact that IM aims to minimize the global error and estimate an optimal weights vector. In this context, the presence of LST values disturbed by inappropriate atmospheric correction can perturb the entire process of optimization very strongly. In this sense, the direct methods are more resistant to those types of effects.

References

[1] J. J. Seetle, N. A. Drake, Linear mixing and the estimation of ground cover proportions b a Department of Geography, International Journal of Remote Sensing, vol. 14:6, pp. 1559-1177, 1997.

[2] H. D. Williamson, Estimating sub-pixel components of a semi-arid woodland, International Journal of Remote Sensing, vol. 15, pp. 3303-3307, 1994.

[3] A. J. Arnfield, Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island, International Journal of Climatology, 2003.

[4] J. Voogt, T. R. Oke, Thermal remote sensing of urban climates. Remote Sens. Environ., vol. 86, pp. 370–384, 2003.

[5] S. Liang, Validation and spatial scaling, in: J. Kong (Ed.), Quantitative Remote Sensing of Land Surfaces, pp. 431-471, New Jersey: Wiley & Sons, 2004.

[6] C. K. Munechika, J. S. Warnick, C. Salvaggio, J. R. Schott, Resolution enhancement of multispectral image data to improve classification accuracy, Photogrammetric Engineering and Remote Sensing, vol. 59, 1993.
[7] M. Stathopoulou, C. Cartalis, Downscaling AVHRR land surface temperatures for improved surface urban heat island intensity estimation, Remote Sensing of Environment, vol. 113, pp. 2592-2605, 2009.
[8] B. J. Choudhury, N. U. Ahmed, S. B. Idso, R. J. Reginato, C. S. T. Daughtry, Relations between evaporation coefficients and vegetation indices studied by model simulations. Remote Sensing of Environment, vol. 50, pp. 1-17, 1994.
[9] R. R. Gillies, T. N. Carlson, Thermal remote sensing of surface soil water content with partial vegetation cover for incorporation into climate models, Journal of Applied Meteorology, vol. 34, pp. 745-756, 1995.
[10] D. O'Leary, Near optimal parameters for Tikhonov regularization and other regularization methods, SIAM Journal on Scientific Computing, vol. 23:4, pp. 1161-1171, Society of Industrial and Applied Mathematics, 2001.
[11] J. C. Jimenez-Munoz, J. A. Sobrino, Split-window Coefficients for Land Surface Temperature Retrieval from Low-Resolution Thermal Infrared Sensors, IEEE Geoscience and Remote Sensing Letters, vol. 5:4, pp. 806-809, 2008.