Eiffel
GEOSS APPLICATIONS FOR CLIMATE CHANGE

REVEALING THE ROLE OF GEOSS AS THE DEFAULT DIGITAL PORTAL FOR BUILDING CLIMATE CHANGE ADAPTATION & MITIGATION APPLICATIONS

D4.2 Analytical framework for improved spatial resolution of CC datasets

Version 1.0
Date 19-Nov-2022

Editor Maria Dekavalla (ICCS)
Authors Maria Dekavalla (ICCS), Dimitris Bliziotis (ICCS)
Reviewers Roxanne S. Lorilla (NOA), Thanassis Drivas (NOA)

Dissemination Level Public

Call H2020-LC-CLA-2020-2
Topic LC-CLA-19-2020
Type of Action Research and Innovation Action
Start Date 01 June 2021
Duration 36 months
Project Information https://cordis.europa.eu/project/id/101003518

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101003518

The Members of the EIFFEL Consortium:

ID	Organisation	Short Name	Country
1	INSTITUTE OF COMMUNICATION AND COMPUTER SYSTEMS	ICCS	GREECE
2	ETHNIKO ASTEROSKOPEIO ATHINON	NOA	GREECE
3	PRODEVELOP SL	PRO	SPAIN
4	UNIVERSITAT POLITECNICA DE VALENCIA	UPV	SPAIN
5	DRAXIS ENVIRONMENTAL SA	DRAXIS	GREECE
6	STICHTING IHE DELFT INSTITUTE FOR WATER EDUCATION	IHE	NETHERLANDS
7	OPEN UNIVERSITEIT NEDERLAND	OUNL	NETHERLANDS
8	NOORD-BRABANT PROVINCIE	NOORD-BRABANT	NETHERLANDS
9	LIBRA MLI LTD	LIBRA	UNITED KINGDOM
10	DIABALKANIKO KENTRO PERIBALLONTOS	IBEC	GREECE
11	AUTORIDAD PORTUARIA DE BALEARES	BPA	SPAIN
12	UNIVERSIDAD AUTONOMA DE BARCELONA	UAB	SPAIN
13	PERIFEREIA ATTIKIS	ATTICA	GREECE
14	NATIONAL PAYING AGENCY	NPA	LITHUANIA
15	SCHWEIZERISCHES FORSCHUNGSINSTITUT FUER HOCHGEBIRGSKLIMA UND MEDIZIN IN DAVOS	PMOD WRC	SWITZERLAND
16	EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS	ECMWF	UNITED KINGDOM
17	SUOMEN YMPARISTOKESKUS	SYKE	FINLAND
18	RISA SICHERHEITSANALYSEN GMBH	RISA	GERMANY
19	EDGE IN EARTH OBSERVATION SCIENCES MONOPROSOPI IKE	EDGE	GREECE

Disclaimer

The information in this document is subject to change without notice. No warranty of any kind is made with regard to this document, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose. The Members of the EIFFEL Consortium shall not be held liable for errors contained herein or direct, indirect, special, incidental or consequential damages in connection with the use or performance of this material. The content of this document reflects only the authors' view. The European Commission and the Research Executive Agency are not responsible for any use that may be made of the information it contains.

Funded by the European Union

Document History

Version	Date	Editor	Comments
0.1	24-Oct-2022	Maria Dekavalla	Table of contents
0.2	9-Nov-2022	Maria Dekavalla	First complete draft
0.3	16-Nov-2022	Roxanne S. Lorilla	Review
0.4	18-Nov-2022	Maria Dekavalla	Integrated NOA review comments
0.5	21-Nov-2022	Dimitris Bliziotis	New additions on conclusions
1.0	23-Nov-2022	Dimitris Bliziotis	First final version

Funded by the European Union

Executive Summary

This document describes the machine learning models and data fusion techniques included in EIFFEL tools for spatial augmentation of Copernicus datasets and developed during the first 18 months of the project. These tools will be used to provide spatially augmented datasets in the service of EIFFEL pilot CC adaptation and mitigation applications. Five individual spatial augmentation tools will be developed: (i) Sentinel 2 super resolution to increase the spatial resolution of the coarser bands of Sentinel 2 to the resolution of the finer bands. (ii) Thermal Sharpening of Sentinel 3 SLSTR using Sentinel 2 Data. (iii) Spatio-temporal fusion of Sentinel 2 and Sentinel 3 OLCI and SYNERGY data to create data with fine spatial resolution and high temporal resolution. (iv) Sentinel-3 OLCI and SYNERGY data super resolution and (v) fusion of Sentinel 5p data with Sentinel 2, proxy and in-situ data if available. During the first 18 months of the project, two of the five spatial augmentation tools have been developed: (i) the Sentinel 2 super resolution and (ii) the thermal sharpening of Sentinel 3 SLSTR with Sentinel 2 data tool. Section 2.1 describes the developed Sentinel 2 super resolution tool. A detailed review of existing super resolution methods is included, and an explanation of the technical choices made during the development phase. In Section 2.2, the implementation of the thermal sharpening tool for augmenting the spatial resolution of Sentinel 3 SLSTR data is explained. Additionally, samples of the augmented datasets provided to Pilot users are demonstrated and their evaluation results are also presented. The developed components are going to be further improved based on the feedback obtained during pilot activities.

Funded by the European Union

Table of Contents

Funded by the European Union

List of Figures

Funded by the European Union

List of Tables

Funded by the European Union

List of Acronyms and Abbreviations

Acronym	Meaning
CAMS	Copernicus Atmosphere Monitoring Service
CC	Climate Change
DL	Deep Learning
GEOSS	Global Earth Observation System of Systems
GKH	GEO Knowledge Hub
ML	Machine Learning
PA	Paris Agreement
OLCI	Ocean Land Color Instrument
SBA	Social Benefit Area
SLSTR	Sea Land Surface Temperature Radiometer
SR	Super Resolution

Funded by the European Union

1 Introduction

This document describes in detail the ML models and data fusion techniques included in EIFFEL tools for spatial augmentation of Copernicus datasets and developed during the first 18 months of the project. The goal is to explain each developed component of the spatial augmentation tools and to document the technical choices made during the development phase. These tools will be used to provide spatially augmented datasets in the service of EIFFEL pilot applications. Feedback will be obtained during pilot activities, upon which the components will be further improved.

1.1 Context

1.1.1 Objectives

The deliverable report 4.2 contributes to 2 of 5 EIFFEL project objectives. The input provided is based on the functional and non-functional requirements defined in the D2.2, which has been the detailed outline of specifications for the CC applications per Pilot study (Figure 1) and the GEOSS tools implementations.

Figure 1: EIFFEL PILOT studies in the European Union.

Funded by the European Union

The Grant Agreement[1] lists five project objectives (O1-O5, below), of which two are indirectly linked to the findings published in this report due to their relevance for T4.2:

1. (**O2**) EIFFEL will leverage techniques of Explainable AI to develop tangible indicators for CC impacts; it will also make use of super resolution, data fusion and stochastic modelling techniques to generate spatially and temporally explicit information from the untapped pool of GEOSS.
2. (**O5**) EIFFEL will develop, using co-creation (O4), a set of CC adaptation and mitigation applications in different and quite diverse GEO SBAs, to demonstrate the project innovations: PILOT1/P1-Water/Land Management, PILOT2/P2-Sustainable Agriculture, PILOT3/P3-Transport Infrastructure, PILOT4/P4-Sustainable Urban Development, PILOT5/P5-Disaster Resilience.

The goal of T4.2 is to establish a ML-based framework and algorithms to augment the spatial resolution of Copernicus data beyond the capabilities of the original imaging systems. The spatially augmented data will be used as input to EIFFEL pilot CC adaptation and mitigation applications.

1.1.2 Work plan

This report, Deliverable **4.2**, corresponds to **T4.2: Augmenting spatial resolution of CC datasets using super resolution and data fusion (M3-M30) (WP4)**. It is part of **WP4: Improving temporal, spatial resolution and data quality of CC-related datasets**.

The spatial augmentation tools presented in **D4.2** are based on **D2.2** Report on EIFFEL specifications, which defines the functional and non-functional system requirements and specifications for both the GEOSS EIFFEL tools and CC applications[2] and **D2.3** EIFFEL System Architecture Report which defines the spatial augmentation tools to be developed.

The results presented in **D4.2** play an important role in the WPs that are directly related to the use of the implemented tools. Results will be used in **WP5**: Development of the EIFFEL CC applications based on GEOSS as well as during **WP7**: EIFFEL pilot demonstrations and impact assessment[3].

1.1.3 Milestones

D4.2 is linked to **MS8**:

- MS8: Alpha versions of D4.1-D4.3 related components. *Means of verification: Alpha versions of the tools in project repository and ready for internal testing*

1.1.4 Deliverables

D4.2. is based on the input of T2.2 (Lead beneficiary: DRAXIS) and is strongly correlated to all the forthcoming deliverables of WP3, WP4, WP5[4], WP6 and WP7[5].

[1] Grant Agreement No 101003518, Part B, p. 5
[2] Grant Agreement No 101003518, ANNEX 1, p. 16
[3] Grant Agreement No 101003518, ANNEX 1, p. 33
[4] Grant Agreement No 101003518, ANNEX 1, pp. 26-29
[5] Grant Agreement No 101003518, ANNEX 1, pp. 33-36

Funded by the European Union

1.2 Intended Readership and Document Structure

The dissemination level of this report is **public**. It is specifically intended for all the partners and especially the WP leaders of **all WPs as it relates to both GEOSS EIFFEL Tools and EIFFEL CC Applications**.

The section following the introduction describes in detail the following tools and components developed during the first 18 months of the project:

1. Sentinel 2 Super Resolution (Section 2.1)
2. Thermal Sharpening of Sentinel 3 SLSTR using Sentinel 2 Data (Section 2.2)

The progress of the remaining tools to be developed and the future work to be done are presented in Section 3.

2 EIFFEL Toolbox for Augmenting Spatial Resolution of CC Datasets

As described in WP4[6], the goal of T4.2 is to establish a ML-based framework and algorithms, that aim at enhancing the spatial resolution of Copernicus data beyond the capabilities of the original imaging systems. The framework will ensure Copernicus cross-platform exploitation, leveraging the inherent characteristics of the Sentinel data and SR modelling, including Deep Neural Network based architectures and techniques, towards enhancing the spatial characteristics of the observed properties in the context of CC applications. In particular, the spatial resolution of the coarser bands of Sentinel 2 (used by all pilots, P1-P5) will be increased to the resolution of the finer bands (i.e. from 60m & 20m to 10m), whilst at the same time preserving the integrity of spectral resolution. SR techniques will also be applied to Sentinel 3 data (used by P1, P3, P5) aiming to sharpen low resolution observations (i.e. from Sea and Land Surface Temperature Radiometer - SLSTR), using high resolution observations from Sentinel 2 satellites, towards improved monitoring of crucial parameters focusing on the interaction between the ground and the atmosphere. Additionally, another facet of this task involves the application of data fusion techniques enabling the combination of data from different sources (e.g. Sentinel with in-situ datasets) in the same processing cycle. Multi-sensor fusion shall predominantly involve the utilisation of in-situ data, that usually are point-based but of high accuracy, aiming to improve spatial characteristics of the observed parameters whilst enabling large scale mapping. ML models will be employed to upscale CC-related properties and transform them into regional or national maps. The employed ML models will be formulated as multiple-layer models to learn a representation of input/output data with abstractions from lower to higher levels. The overall process will include the semantic combination and assignment of in-situ measurements to EO data and products (i.e. Sentinel 5p, CAMS) as well as the characterisation of areas, from which related in-situ information may not be available.

Based on the above-mentioned goals and objectives and the user requirements specified in D2.2, five different spatial augmentation tools have been defined in D2.3 to fulfil the user requirements specified in D2.2. The first component is concerned with a DL-based SR tool for

[6] Grant Agreement No 101003518, ANNEX 1, pp. 23-25

Funded by the European Union

enhancing the spatial resolution of the coarser bands of Sentinel 2 to the resolution of its finer bands. The second component deals with a DL-based SR tool for augmenting the spatial resolution of Sentinel 3 OLCI and SYNERGY data through training with Sentinel 2 data. The third component is concerned with the spatio-temporal fusion of Sentinel 2 and Sentinel 3 OLCI and SYNERGY data to create synthetic data that simultaneously have the finer spatial resolution of Sentinel 2 and the temporal resolution of Sentinel 3. The last two components are involved with enhancing the spatial resolution of Sentinel 3 SLSTR and Sentinel 5p data by fusing them with Sentinel 2 images, proxy and in-situ data, if available.

During the first 18 months of the project, two of the five spatial augmentation tools have been developed: (i) the Sentinel 2 SR and (ii) the thermal sharpening of Sentinel 3 SLSTR with Sentinel 2 data tool. For a better interpretation of the developed components, a detailed description of each component is provided below.

2.1 Sentinel 2 Super Resolution

Sentinel 2 satellites carry an optical payload with visible, near-infrared and shortwave infrared sensors comprising 13 spectral bands: 4 bands at 10 m, 6 bands at 20 m and 3 bands at 60 m spatial resolution. Reasons for recording at varying spatial resolution include storage and transmission bandwidth restrictions, improved signal-to-noise ratio in some bands through larger pixels and bands designed for specific purposes that do not require high spatial resolution (e.g. atmospheric correction). However, spatial data analysis requires all the available bands to have a common and highest available spatial resolution to extract more detailed and accurate information. Given the wide range of CC-related applications that Sentinel 2 data can support, the first tool provided by Task 4.2 is responsible for increasing the spatial resolution of the coarser bands (i.e. 20 m and 60 m) to the target resolution of the finer bands (i.e. 10 m). Next follows a review of state-of-the-art methods for improving the spatial resolution of Sentinel 2 coarser bands concluding with the approach that was developed and a detailed description of the methodology, the presentation and validation of results.

2.1.1 Introduction

Sentinel 2 bands can be downsampled to 10 m resolution with simple interpolation techniques, like bicubic and bilinear interpolation. However, these methods return blurry images with no additional high-resolution information [1]. More sophisticated methods try to augment the coarser bands by inserting as much as possible of the spatial detail retrieved from finer resolution bands. These methods can be classified into three types: (1) pansharpening, (2) model-based, and (3) ML/DL approaches.

In pansharpening methods, a multispectral image of high spectral resolution is fused with a panchromatic image of higher spatial resolution to generate a synthetic image which has the same spatial resolution as the panchromatic image and the same spectral resolution as the original multispectral image. These methods require the availability of a high spatial resolution panchromatic image that has significant spectral overlap with the spectral bands of the coarse multispectral image. However, there are no available panchromatic bands that cover most of the Sentinel 2 sensor's spectral range. A widely used assumption in pansharpening is that a linear combination of bands gives an approximation to the

Funded by the European Union

panchromatic image. Several works explored this intermediate solution to increase the spatial resolution of Sentinel 2's 20 m bands up to 10 m by using heuristics to select or synthesize a "synthetic panchromatic" image from the 10 m bands. For instance, Vaiopoulos and Karantzalos [2] evaluated different pansharpening algorithms to enhance the spatial resolution of the 20 m visible near-infrared (VNIR) and short-wave infrared (SWIR) bands of Sentinel 2 by using fine information provided by the 10 m bands or synthetic panchromatic images created from combinations of the 10 m bands. Similarly, Park et al. [3] optimised the synthesis of simulated panchromatic images from the 10 m bands of Sentinel 2 to improve pansharpening results. However, the performance of pansharpening methods in the case of Sentinel 2 can be challenging given the important requirement of real or simulated PAN image. It is difficult to be obtained because Sentinel 2 data include more than one bands at the highest resolution, which also do not spectrally overlap with the lower resolution ones [4].

Model-based methods address the spatial augmentation problem by inverting an explicit observation model, which describes the blurring, spatial resolution degradation and noise processes that generated the coarse resolution image. Considering that the inverse model is ill-posed, since the number of unknown variables is much greater than the number of observed variables, an explicit regulariser (i.e., image prior) is required. Indicatively, Lanaras et al. [1] introduced a model-based super resolution approach for Sentinel 2 termed as SupReME (SUPer-REsolution for multispectral Multiresolution Estimation). In this approach, an edge-reserving regulariser is included which extracts the discontinuities from the fine resolution bands and transfers them to other bands. The correlation of the spectral bands is also exploited to reduce the dimensionality of the problem. Paris et al. [5] extended the SupReME approach by employing a patch-based regularisation that promotes self-similarity of images. Ulfarsson et al. [6] also extended SupReME by performing dimensionality reduction during optimisation and by using cyclic decent on a manifold. Brodu et al. [7] developed a two-stage method that at the first stage extracts the band-dependent spectral information that is common to all bands termed as "geometry of scene elements". Then applies this model to the coarser resolution bands such that they are consistent with the scene elements while preserving their reflectance using spectral unmixing techniques.

The third type of ML/DL-based methods includes SR techniques which directly learn the relation between coarse and fine spatial resolution images by training from data, instead of defining explicitly blurring, spatial resolution degradation and noise processes through a regulariser. CNNs are one of the most important paradigms in SR due to their potential to extract high-level features from images. As a result, several SR methods based on CNNs have been presented and tested. In the reference approach of Dong et al. [8] (Super-Resolution Convolutional Neural Network - SRCNN), a three-layer CNN is utilised to learn the mapping between coarse images, which are initially downscaled to the target resolution by interpolation, and their ground-truth fine resolution counterparts. Kim et al. [9], [10] developed two deep convolutional networks for SR enhancement. The first [9] was the Deeply Recursive Convolutional Network for Image Super Resolution (DRSN), which uses recursive or shared weights to reduce model parameters in a deep 20-layer CNN. The second [10] (Very Deep Super Resolution = VDSR) is also a deep CNN which includes a final residual connection that adds the interpolated input image to the output, and therefore the network needs only to learn the fine details not included in the interpolated input image, instead of regenerating the whole image.

Funded by the European Union

The residual (or skip) connections in CNNs were introduced by He et al. [11] for image object recognition. These connections bypass portions of the network and are added again later, such that skipped layers only need to estimate the residual with respect to their input state. In this way, the average effective path length through the network is reduced, which alleviates the vanishing gradient problem and greatly accelerates the learning. Mao et al. [12] employed the ResNet architecture in SR without further modification. However, applying ResNet architecture directly to low level vision problems like SR can be suboptimal concerning the fact that the original ResNet was proposed to solve higher-level computer vision problems like image classification and detection. To address this problem, Lim et al. [13] developed the enhanced deep residual network (EDSR) by removing all BatchNorm layers while stabilising the training process using a residual scaling layer after the last convolution of each residual block. With these modifications, EDRS has been shown to provide some performance advantages with respect to SRRN when considering moderate scaling ratios [14].

All the above-mentioned works have in common that they need to predict images of fine spatial resolution only based on previously seen fine resolution images. However, the problem of super-resolving the coarser bands of Sentinel 2 to the resolution of the finer bands deviates from the classic SR problem. In this case, we have access to the fine resolution bands during training and prediction and thus, we can transfer only the fine resolution information to the coarse resolution bands instead of generating the whole fine resolution image. Several DL-based methods have been proposed for this particular Sentinel 2 SR problem.

Lanaras et al. [15] were inspired by EDRS network and proposed the first CNN model to effectively super resolve all Sentinel 2 bands (i.e. 20m and 60m) to 10 m spatial resolution. The network was trained for two different scale factors (i.e. 2x and 6x) on a random selection of 60 Sentinel 2 images. Gargiulo [16] also presented a CNN architecture specially designed for super resolving 20 m Sentinel 2 bands to 10 m by introducing additional convolutional layers and a revisited loss function. A Deep Residual Network that fuses both fine and coarse spatial resolution bands was proposed by Palsson et al. [17]. Connection convolutional neural network (S2 SSC) extends on this but uses an unsupervised single image learning (zero-shot) approach in which the training and testing are performed on the same image [18]. The Remote Sensing Very Deep Super-Resolution (RSVDSR) [19] network retrains previously developed deep convolutional networks to work with Sentinel 2 data with promising results. Sentinel 2 Parallel Residual Network (SPRNet) [20] trained two sets of parallel residual networks and obtained good spatial fidelity and spectral preservation. The approach of Zhang et al [21] was inspired by DSen2 and developed two large neural networks for deployment and training on high performance computing clusters, one for 20 m bands and one for 60 m bands.

The existing DL-based Sentinel 2 SR methods are supervised and the required training dataset must include both the coarse resolution input and the ground-truth fine resolution output. However, 10 m resolution reference images are not available for the 20 m and 60 m Sentinel 2 bands. The lack of fine resolution reference images during training is addressed by reducing the resolution of images before training, by the scaling ratio between the coarse and fine bands. This strategy is inspired by Wald's Protocol [22], which assumes that the spectral correlation of an image is self-similar over a limited range of scales. Therefore, downsampling from 20 m to 10 m by inserting high resolution details across spectral bands can be achieved from images resampled at 40 m and 20 m, and similarly for the 60 m to 10 m case by resampling at 360 m and 60 m. Under this assumption, an unlimited amount of training data

Funded by the European Union

can be created by synthetically resampling original Sentinel 2 images by the desired factor and using them as input to generate the original data as output.

Given the ability to create training datasets without the need for high-resolution reference images, two different training approaches have been established to train Sentinel 2 DL-based SR models. The first approach is to train the neural network with an extensive training dataset with global coverage such that it can generalize well across different areas and can super-resolve arbitrary Sentinel 2 images without the need for retraining. In the second approach. models are trained and tested on the same image and thus the training is quicker than the first approach. This strategy is used when a single specific image is needed to be super-resolved. However, the network must be retrained each time a different Sentinel 2 image must be super-resolved. Concerning that CC mitigation and adaptation applications will be applied to various areas with different climatic conditions and land cover types, the first training approach with the extensive training dataset is considered more efficient because it is trained once and can super-resolve arbitrary Sentinel 2 images without further retraining.

The performance of DL-based SR methods can be sensitive to important training hyperparameters. Palsson et al. [17] trained a deep residual network with a single image training dataset and found that the optimal number of training patches depends on the complexity of the single image used for training, while a large number of epochs does not improve the performance because the network converges quickly. The network also appears insensitive to the number of residual blocks and patch size. Unfortunately, a similar performance evaluation of SR networks trained with large training datasets is missing from the literature. To fill this gap, three characteristic deep residual network architectures were trained with a single large training dataset to find the most efficient. To achieve the optimum performance, we also focused on the effects of various training hyperparameters (number and size of patches, number of training epochs, batch size and number of residual blocks) on the quantitative and qualitative performance of the SR models. The most efficient and fine-tuned architecture was used to create the final Sentinel 2 SR tool.

Summarizing our contributions, we have developed a large training dataset with a global distribution of samples to compare the performance of several deep residual network architectures and their hyperparameters and to select the most efficient which has global applicability for Sentinel 2 Level 2A data without retraining. The developed Sentinel 2 SR tool will provide Sentinel 2 images where all the included bands are at the highest available spatial resolution of 10 m.

2.1.2 Methodology

2.1.2.1 Dataset Creation

A large training dataset of Sentinel 2 data was considered mandatory to train a network that can super-resolve random Sentinel 2 images without retraining. Atmospherically corrected Sentinel 2 Level 2A tiles from 45 randomly selected locations were downloaded from the Copernicus Services Data Hub, aiming for wide distribution on the globe and for covering a variety of climate zones and land-cover types. An additional dataset of 5 Sentinel 2 Level 2A tiles located in possible areas of interest for the Pilots was also collected and used for testing. To simplify training and testing, we chose only images with low cloud cover and

Funded by the European Union

with no undefined ("black background") pixels. The locations of the training and test images are illustrated in Figure 2. Out of these images, 80% are used for training and the remaining 10% serve as validation set. The filenames of the Sentinel 2 images included in the training, validation and test datasets are defined in Appendix 1.

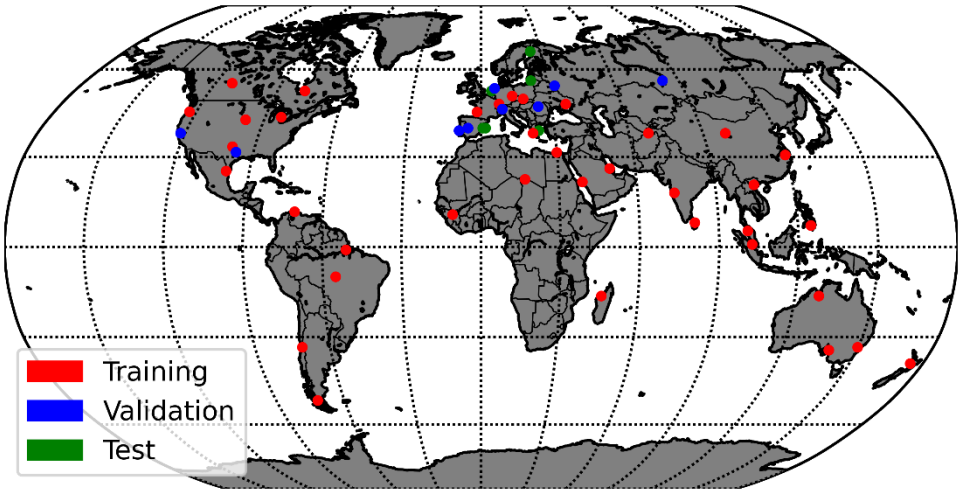

Figure 2 Locations of Sentinel 2 tiles selected for training and testing

CNNs are fully supervised and need a lot of training data including low resolution input images and their corresponding true high-resolution outputs. However, a central issue in our approach is how to construct the training, validation and testing datasets, given that ground truth images with 10m resolution are not available for the 20m and 60m bands. Since there is no high-resolution reference image available, the images need to be degraded in resolution by the desired scale factor to be used as input data and the original observed images are used as reference output. This approach is the most widely used method for dataset creation in DL-based Sentinel 2 SR approaches. And it is based on Wald's protocol [22] which assumes that the spectral correlation of an image is self-similar over a limited range of scales.

For better understanding of the dataset creation process, we use the following notation: the observed fine bands are denoted by $Y \in R^{d_1 \times d_2 \times L_1}$ and the observed coarse bands are denoted by $X \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2}$, where $d_1 \times d_2$ is the dimension of the fine resolution bands, L_1 is the number of fine resolution bands, L_2 is the number of coarse resolution bands and s is the scale ratio between the fine and coarse resolution bands. Downsampling by the factor s is denoted by D and upsampling by a factor s is denote by U. Finally, we denote the spatially degraded fine and coarse bands by DT and DX, respectively, where the operator D degrades their spatial dimensions by the factor s. Since DX is smaller than DY by a factor of s along each spatial dimension, it needs to be interpolated to the same size as DY to be added elementwise with the fine information extracted and compared with the DY during training. We denote the interpolated degraded coarse bands by $X^D \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2} = UDX$ and the degraded fine bands by $Y^D \in R^{\frac{d_1}{s} \times \frac{d_2}{s} \times L_2} = DY$. Now XD, YD and X have the same spatial dimensions.

In practice, to generate training data with the desired scale ratios of 2 and 6, we downsampled the original Sentinel 2 data by first blurring them with a Gaussian filter of

Funded by the European Union

standard deviation σ = 1/s pixels. Then, we downsampled by averaging over s×s windows, with s = 2 and s = 6, respectively. The process of generating the training data is illustrated in Figure 3. In this way, we obtained two datasets for training, validation and testing (Table 1). The first dataset consists of fine resolution images at 20m and coarse resolution images of 40 m, created by downsampling the original 10m and 20m bands by a factor of 2. It served to train a network for 2× super-resolution. The second one consists of images with 60m, 120m and 360m resolution, downsampled from the original 10m, 20m and 60m data. This dataset was used to learn a network for 6× super-resolution. We note that, due to unavailability of 10m ground truth, quantitative analysis of the results was also conducted at the reduced resolutions.

Figure 3. The process workflow applied for simulating the data for training and testing

Table 1. Spatial resolution of original data and training datasets according to the scale factor

Scale Factor	Original Data	Training Dataset
2	4 bands [B2,B3,B4,B8] (10 m), 6 bands [B5,B6,B7,B8A,B12] (20 m)	Input: 4 bands [B2,B3,B4,B8] (20 m), 6 bands [B5,B6,B7,B8A,B12] (40 m) Output: 4 bands [B2,B3,B4,B8] (20 m), 6 bands [B5,B6,B7,B8A,B12] (20 m)
6	4 bands [B2,B3,B4,B8] (10 m), 6 bands [B5,B6,B7,B8A,B12] (20 m), 2 bands [B1, B9] (60 m)	Input: 4 bands B2,B3,B4,B8] (60 m), 6 bands [B5,B6,B7,B8A,B12] (120 m), 2 bands [B1, B9] (360 m) Output: 4 bands B2,B3,B4,B8] (60 m), 6 bands [B5,B6,B7,B8A,B12] (60 m), 2 bands [B1, B9] (60 m)

To make the training of the network computationally feasible, each input image was divided into many small patches of suitable size. Thus, the input to the network is patches of the stacked degraded bands, i.e., patches of $[X^D, Y^D] \in R^{\frac{d_1}{r} \times \frac{d_2}{r} \times (L_1+L_2)}$, denoted by $[X_i^D, Y_i^D] \in R^{p \times p \times (L_1+L_2)}, i = 1, \dots, M$, where M is the number of patches and p is the patch-size. The target patches during training come from X and are denoted by $X_i \in R^{p \times p \times (L_1+L_2)}, i = 1, \dots, M$. As all the bands have the same size, the i^{th} patch covers the same part of the scene for all the images. In practice, we sampled 1000 random patches per training image. Three different training sets of 45000 patches each were created with size of 32x32, 64x64 and 128x128 pixels to define the optimum patch size during hyperparameter tuning.

The processing of the dataset was performed in Python. The Gaussian blurring was performed with the *gaussian filter* included in the multidimensional image processing package, *scipy.ndimage*. The downscaling by average was performed with the *skimage.measure.block_reduce* function. Finally, the upscaling of the coarse image patches to fit the dimensions of the fine image patches was implemented with the *skimage.transform.resize* function, which performs spline interpolation with default order of 1. Finally, the tiling into patches was performed with the *gdal.Translate* function.

2.1.2.2 Convolutional Neural Network Architectures

Funded by the European Union

Finding the right size and capacity for a CNN is largely an empirical choice. In the context of Sentinel 2 SR, three different deep residual network architectures have been proposed: (i) DSen2 architecture proposed by Lanaras et al. [15], (ii) ResNet proposed by Palsson et al. [17] and (iii) SPRNet proposed by Wu et al. [20]. The three neural network architectures are demonstrated in Figure 4.

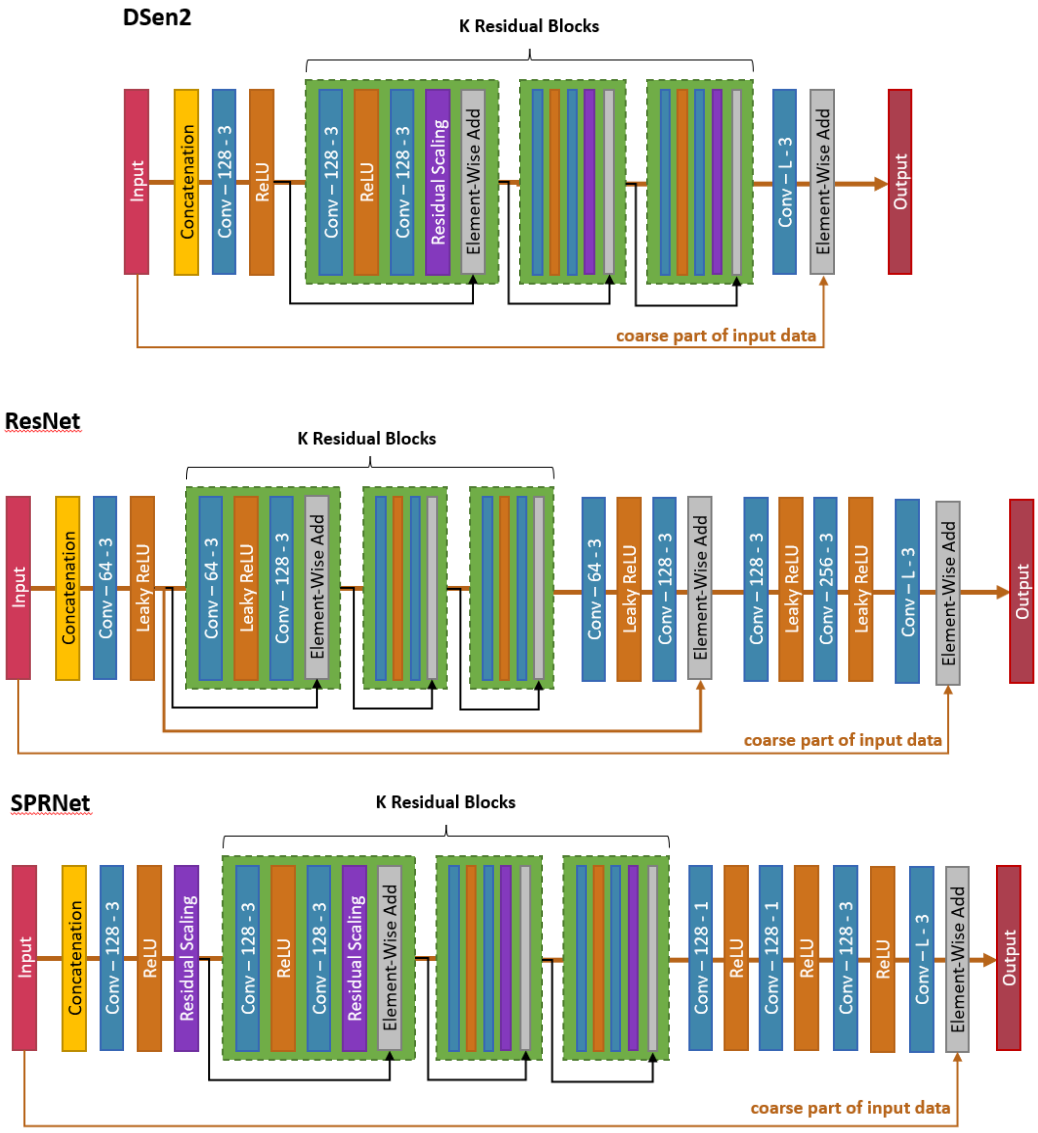

Figure 4 Sentinel 2 Super-Resolution Neural network architectures

All three networks were trained with the same training dataset of 45 images which were split into 45000 patches for 30 epochs with the same hyperparameter configuration to select the one that achieves the fastest convergence during training and the best evaluation results. Each network was implemented in the Keras framework with Tensorflow as the backend. The mini-batch size for SGD was set to 128 patches. The initial learning rate is 10^{-4} and it is reduced by a factor of 2 whenever the validation loss does not decrease for 5 consecutive epochs. For numerical stability, the raw reflectance values of each image patch were divided by 2000 before training. The network weights were initialized to small random

Funded by the European Union

values with the He uniform method [23] and optimised with the stochastic gradient descent. In specific, we used the Adam variant of SGD [24] with Nesterov momentum.

Various evaluation metrics were computed during validation at the end of each training epoch and during the evaluation with the test images: MAE, Peak Signal-Noise Ratio (PSNR), spectral angle mapper (SAM), ERGAS, universal image quality index (UIQ). According to the evaluation results (Table 2), the SPRNet architecture achieves the best results but within a long training time. The DSen2 architecture was selected since it achieves the second-best results and requires a shorter training time.

Table 2. Evaluation results for three different SR neural network architectures for super-resolving 20 m bands (ERGAS - relative dimensionless global error, MAE – mean absolute error, PSNR – peak signal to noise ratio, SAM – spectral angle mapper, UIQ – universal image quality)

	ERGAS	MAE	PSNR	SAM	SRE	UIQ	Time per Epoch	Epochs	Training Duration
SPRNet	**0.991924**	**0.013503**	**33.457920**	**0.937818**	**40.766830**	**0.999572**	1 h 49 min	30	54 h 30 min
ResNet	1.376152	0.018745	30.612630	1.301025	39.407833	0.999177	26 min	30	13 h
DSen2	1.130744	0.015817	32.318603	1.064090	39.190921	0.999445	45 min	30	**22 h 30 min**

To achieve the optimum performance, we also focused on the effects of various training hyperparameters (patch size, batch size, and loss function) on the quantitative and qualitative performance of the SR models. Figure 5 shows the spectral angle mapper values achieved with 4 different batch sizes while the other configuration settings were the same (Patch size = 32, Loss function = Mean Absolute Error (MAE), Residual blocks = 6, Optimizer = ADAM with Nesterov momentum, learning rate = 10^{-4}(reduced by a factor of 2 when stable for 5 epochs), epochs = 50). Additionally, Figures 6 and 7 show the spectral angle mapper values achieved when the network was trained with three training datasets of different patch sizes (i.e., 32, 64, 128 pixels) and loss functions (i.e., MAE, MSE, ERGAS), respectively.

The optimum configuration setting included a training dataset which was split into patches of 32 x 32 pixels size for 2x scaling ratio (i.e. SR of 20 m to 10 m) and 64 pixels size of 6x scaling ratio (i.e. SR of 60 m to 10 m). The network was trained on batches of 32 tiles with mean absolute error as the loss function and the ADAM with Nesterov momentum as an optimizer. The learning rate was 10^{-4}, which was reduced by a factor of 2 when stable for 5 epochs. The model was trained with the optimum configuration for 100 epochs. Figure 8 and Figure 9 depict the training curves while training with the optimum configurations for super resolving for scale factors of 2x and 6x. Table 3 includes the evaluation results of the fined tuned neural network architectures for super-resolving 20 m and 60 m bands to the 10 m target resolution.

Funded by the European Union

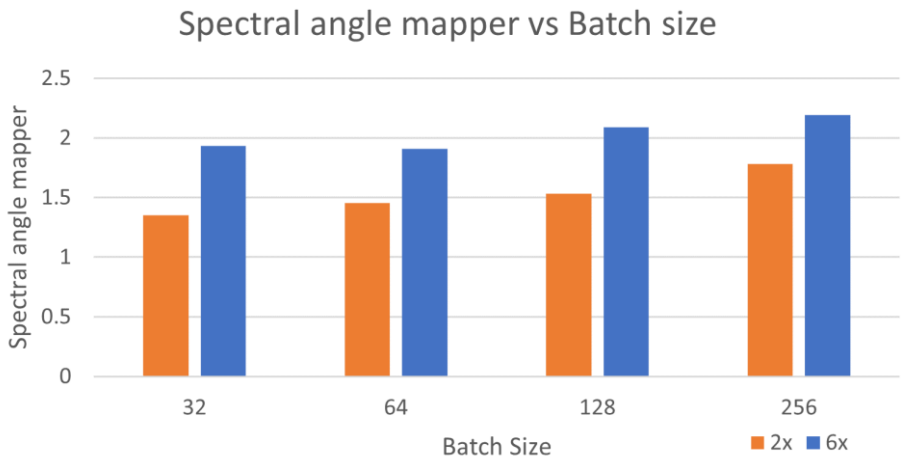

Figure 5. Spectral angle mapper values achieved with different batch sizes

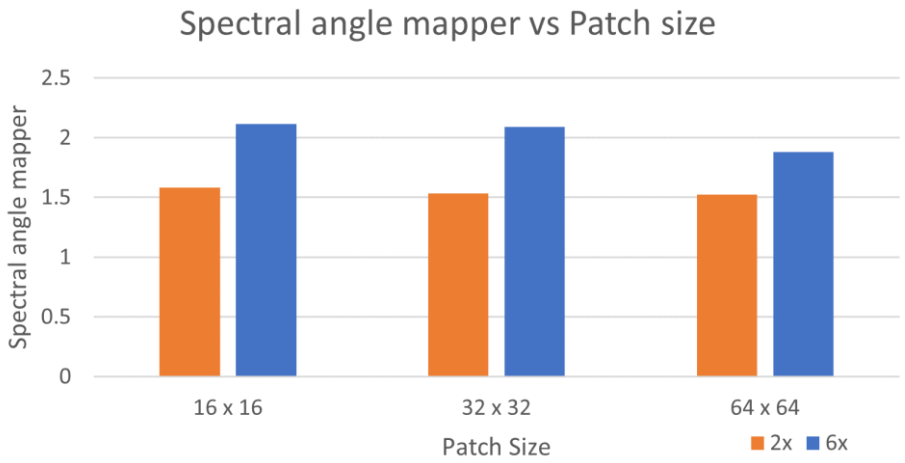

Figure 6. Spectral angle mapper values achieved with different patch sizes

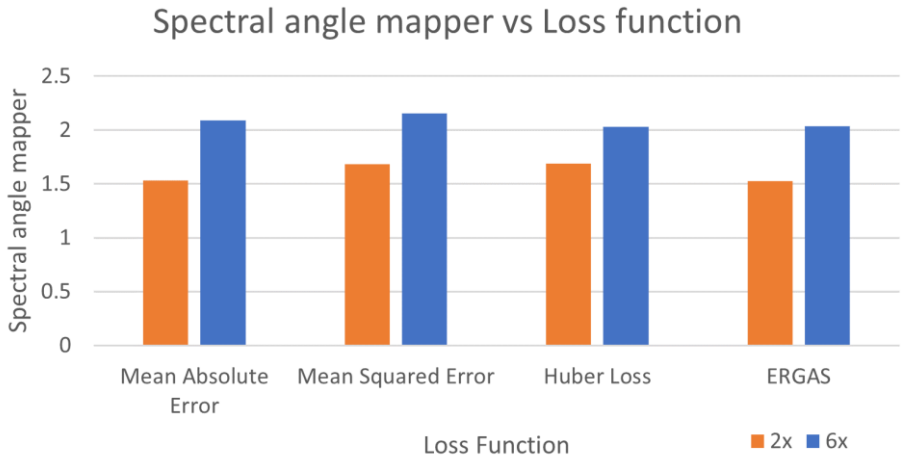

Figure 7. Spectral angle mapper values achieved with different loss functions

Funded by the European Union

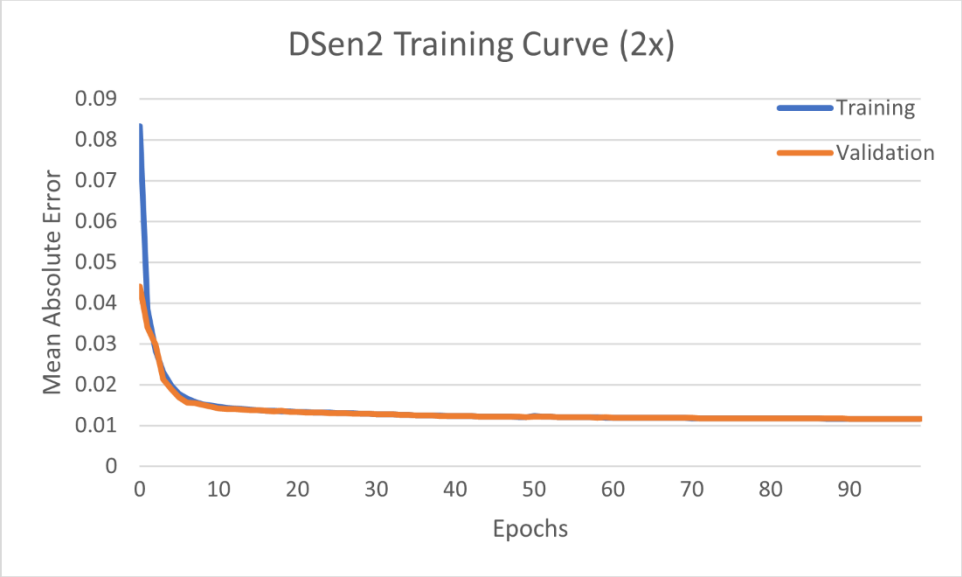

Figure 8. Training curve of DSen2 for scaling ratio 2 with the optimum configuration

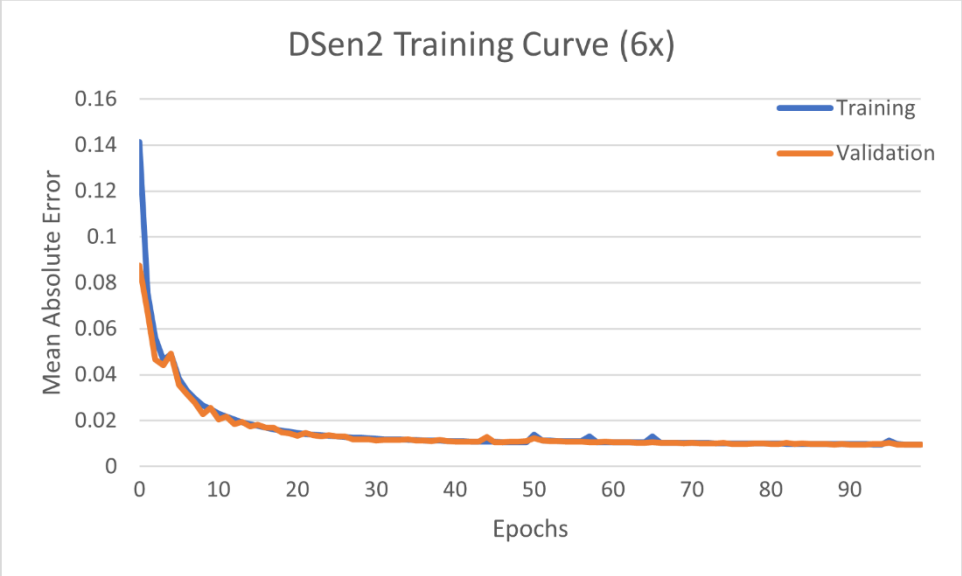

Figure 9. Training curve of DSen2 for scaling ratio 6 with the optimum configuration

Table 3. Evaluation results for the fine-tuned SR neural network architecture for super-resolving 20 m bands (2x) and 60 m bands (6x)

Scaling Ratio	MAE	PSNR	SAM	ERGAS	UIQ
2x	0.0121	34.7523	1.3308	1.4961	0.9991
6x	0.0159	32.2598	1.9371	1.4763	0.9986

2.1.3 Sentinel 2 SR Tool

The two trained neural networks for downscaling 20 m and 60 m bands to 10 m resolution are the basis of the Sentinel 2 SR tool, which is implemented with Python. The tool can super-resolve entire Sentinel 2 Level 2A images in SAFE format. The under-processing image is read partially with the GDAL library to reduce the memory requirements. Each accessed part of the

Funded by the European Union

Sentinel 2 image (both 20 m and 60 m bands) is super resolved to 10 m resolution and stored in the output file which is defined by the name of the original Sentinel 2 name and the initial 'SR_' at the front of the filename. The final super resolved image includes 12 bands with 10 m spatial resolution. Next, super resolved Sentinel 2 images acquired over the 5 different Pilot's AOI are presented as a proof of concept.

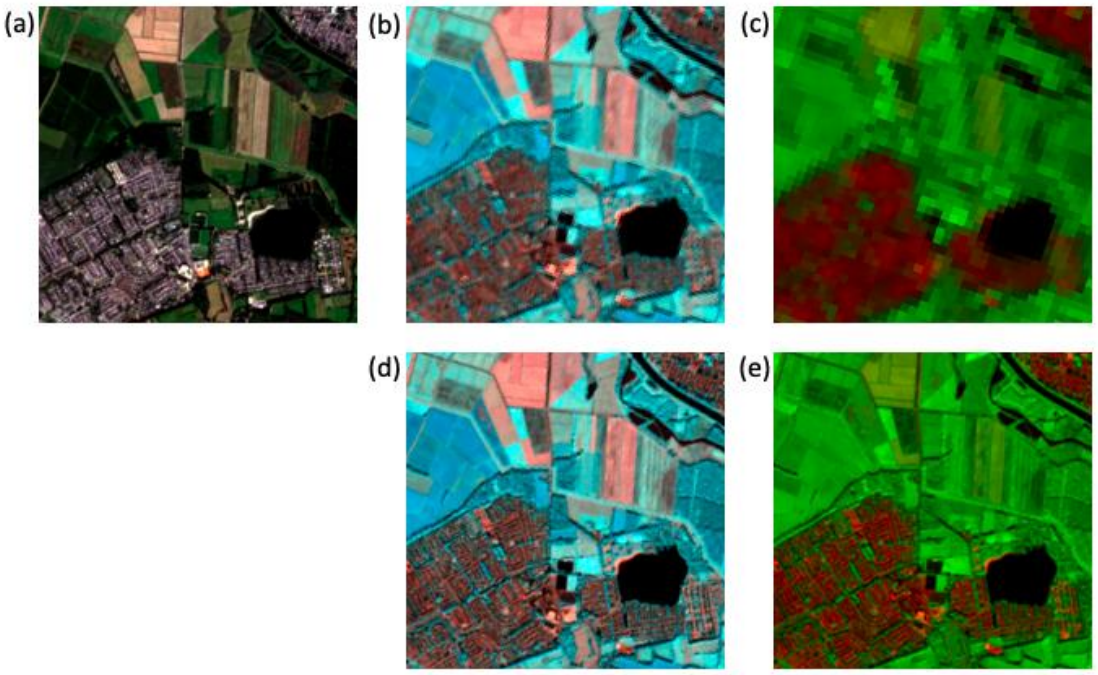

Figure 10. Top: (a) Input Sentinel 2 bands at 10 m (RGB Bands 4,3,2), (b) 20 m (RGB Bands 5,6,7) and (c) 60 m (RGB Bands 1,9) spatial resolution, Bottom: Super-resolved bands at 10 m spatial resolution (d) (RGB Bands 5,6,7) and (e) (RGB Bands 1,9). (Location: Area around Aa of Weerijs, Netherlands - Pilot 1)

Funded by the European Union

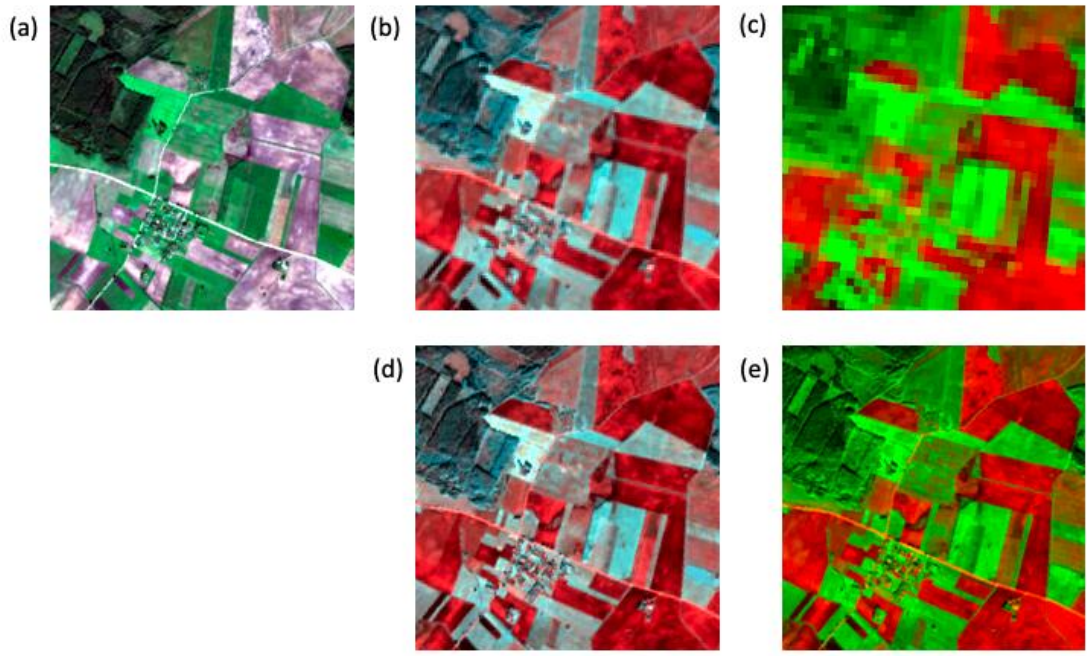

Figure 11. Top: (a) Input Sentinel 2 bands at 10 m (RGB Bands 4,3,2), (b) 20 m (RGB Bands 5,6,7) and (c) 60 m (RGB Bands 1,9) spatial resolution, Bottom: Super-resolved bands at 10 m spatial resolution (d) (RGB Bands 5,6,7) and (e) (RGB Bands 1,9). (Location: Lithuania - Pilot 2)

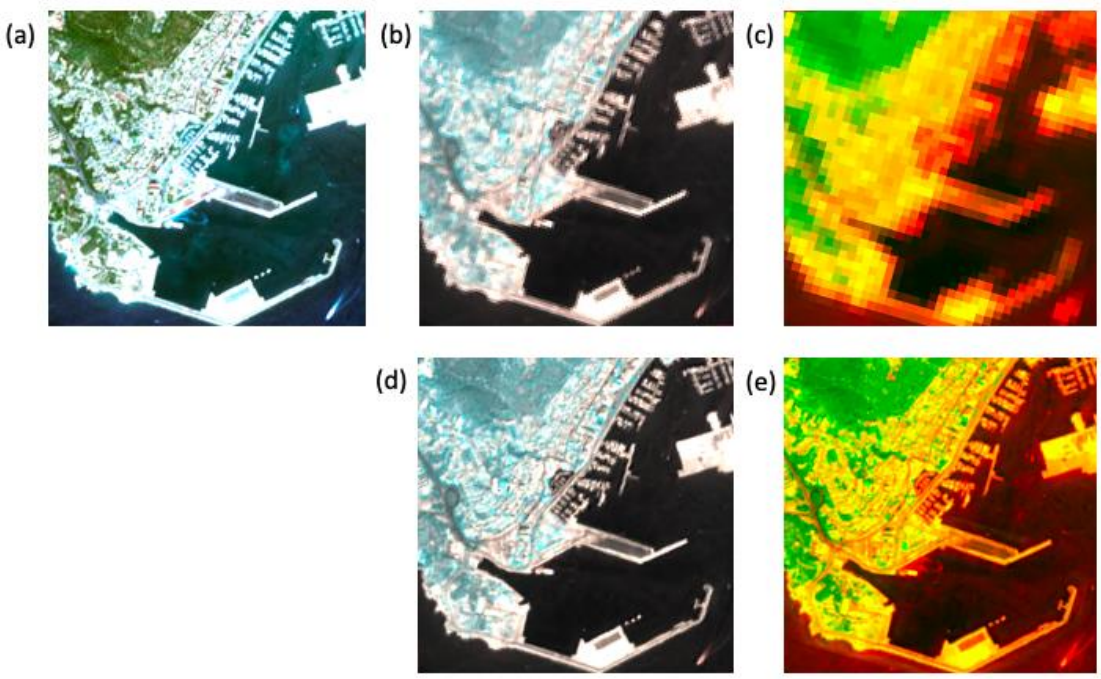

Figure 12. Top: (a) Input Sentinel 2 bands at 10 m (RGB Bands 4,3,2), (b) 20 m (RGB Bands 5,6,7) and (c) 60 m (RGB Bands 1,9) spatial resolution, Bottom: Super-resolved bands at 10 m spatial resolution (d) (RGB Bands 5,6,7) and (e) (RGB Bands 1,9). (Location: Palma Port, Majorca, Spain - Pilot 3)

Funded by the European Union

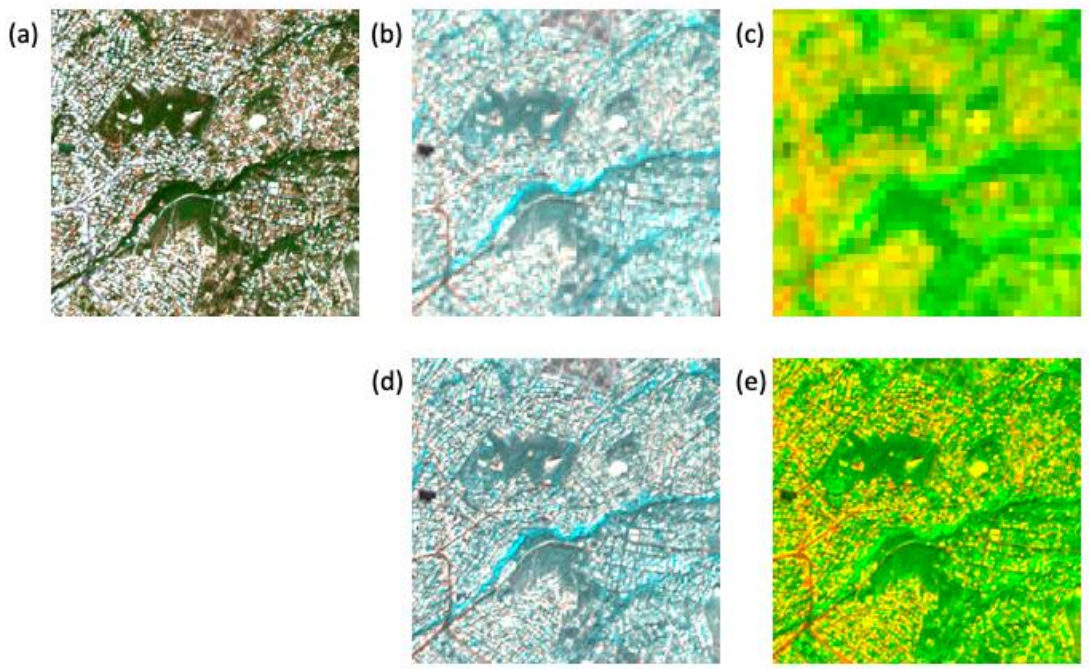

Figure 13. Top: (a) Input Sentinel 2 bands at 10 m (RGB Bands 4,3,2), (b) 20 m (RGB Bands 5,6,7) and (c) 60 m (RGB Bands 1,9) spatial resolution, Bottom: Super-resolved bands at 10 m spatial resolution (d) (RGB Bands 5,6,7) and (e) (RGB Bands 1,9). (Location: Athens, Greece - Pilot 4)

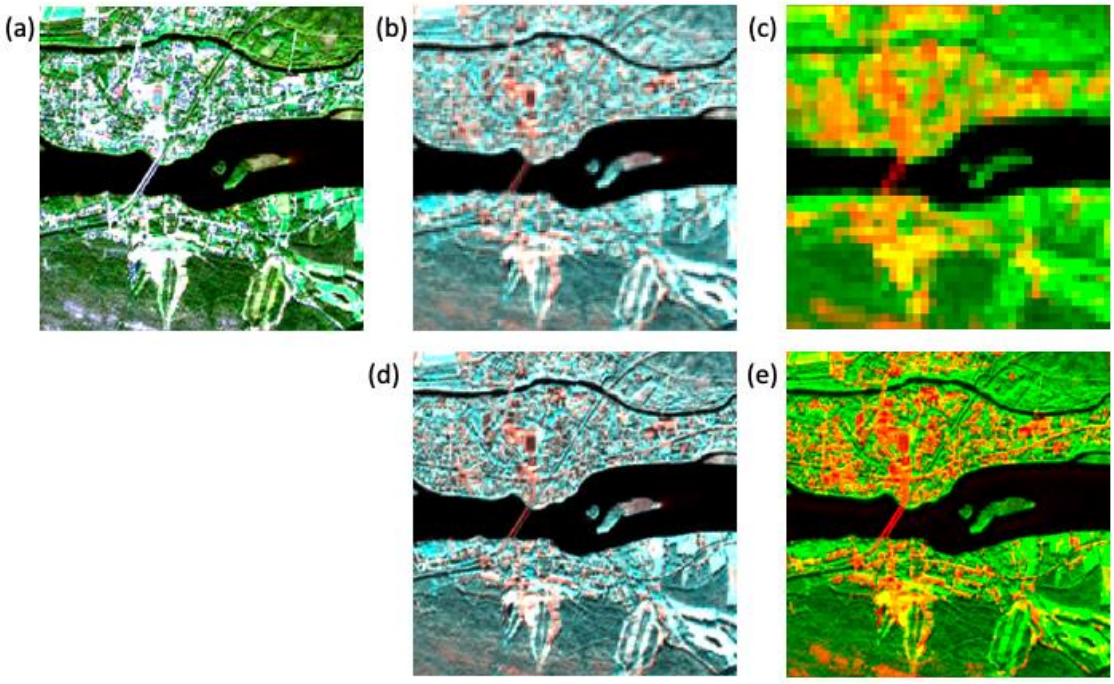

Figure 14. Top: (a) Input Sentinel 2 bands at 10 m (RGB Bands 4,3,2), (b) 20 m (RGB Bands 5,6,7) and (c) 60 m (RGB Bands 1,9) spatial resolution, Bottom: Super-resolved bands at 10 m spatial resolution (d) (RGB Bands 5,6,7) and (e) (RGB Bands 1,9). (Location: Finland - Pilot 5)

Funded by the European Union

2.2 Thermal Sharpening of Sentinel 3 SLSTR using Sentinel 2 Data

The Sentinel satellite constellation is designed to provide earth observation data for many diverse applications. It can be achieved by acquiring frequent observations from a combination of optical, thermal and microwave sensors at various spatial resolutions. Thermal infrared (TIR) remote sensing images can be used to estimate the spatial distribution of land surface temperature and can be used for wildfire detection, mapping land surface energy fluxes and evapotranspiration, monitoring urban heat fluxes and detecting drought. For many of these applications, TIR data are required at a relatively fine resolution. However, the Sentinel constellation contains only one thermal sensor (Sea and Land Surface Temperature Radiometer – SLSTR) with a coarser spatial resolution of 1000 m (on board of the Sentinel-3 satellites) than that of shortwave sensors on the same satellite platform and high-resolution multispectral sensors on the Sentinel 2 satellites. Therefore, an efficient thermal sharpening tool is required to bridge the spatial resolution gap between the currently available coarse resolution Sentinel 3 thermal images and the fine Sentinel 2 multispectral images. In the following sections, we review several thermal sharpening techniques and select the most efficient for the CC adaptation and mitigation applications.

2.2.1 Introduction

Several techniques have been developed for sharpening coarse spatial resolution TIR images with fine spatial resolution shortwave images. The basic assumption of most thermal sharpening approaches is that land surface temperature (LST) has a unique relationship to shortwave band signals across a given imaging scene. Relationships between LST and spectral signals within a scene are determined empirically at the coarse (thermal band) pixel resolution and then applied to the fine (shortwave image) pixel resolution to produce sharpened thermal band imagery. Most of these techniques consist of two general steps. In the first step, a relationship is derived between the TIR data and fine resolution data aggregated to the resolution of the TIR data. In the second step, this relationship is applied to fine resolution data to obtain TIR observations at the same high resolution. A post-processing step is sometimes also included to remove the bias between TIR datasets at the two resolutions.

A classic thermal sharpening technique, TsHARP [25], uses a relationship between land surface temperature (LST) and Normalized Difference Vegetation Index (NDVI) developed empirically at the TIR spatial resolution and applied at the NDVI spatial resolution. However, several studies [26]–[29] have shown that unique relationships between temperature and NDVI may only exist for a limited class of landscapes, with mostly green vegetation and homogeneous air and soil conditions. To address this problem, Merlin et al. [28] developed separate LST-NDVI relationships for photosynthetically and non-photosynthetically active vegetation cover to sharpen imagery over an irrigated cropland in Mexico. Dominguez et al. [26] found that the LST-NDVI relationship was ill-defined over an urban area in Puerto Rico, and included albedo as a predictor in their High-resolution Urban Thermal Sharpener (HUTS), which yielded smaller Mean Absolute Error (MAE) and higher correlation coefficient compared to TsHARP. These refinements to the original TsHARP approach are based also on the same conceptual framework. They define the linear relationship between LST and higher-order products such as albedo, land cover type and emissivity, which are difficult to be applied in an automated way to different landscapes. To extend the application of thermal sharpening

Funded by the European Union

to more complex conditions, a data mining sharpener (DMS) technique was proposed by Gao et al. [30]. The DMS approach builds regression trees between a TIR image and shortwave spectral reflectances based on intrinsic sample characteristics. There is no pre-selection of which shortwave band combination is important for a particular landscape, but it is determined adaptively by the DMS. The method is designed to utilize the full complement of TIR and shortwave data collected by a given satellite platform. Therefore, the Sentinel 3 SLSTR thermal sharpening tool is based on the DMS approach, given its robustness to be applied to various landscapes and conditions without the need for higher order data products.

2.2.2 Methodology

The implemented DMS approach uses reflectances from the fine resolution Sentinel 2 bands as independent variables. The reflectance data are aggregated to match the coarse Sentinel 3 SLSTR spatial resolution. Then, high-quality land surface temperature and fine resolution reflectance sample data are selected from the imagery based on the sub-pixel variation in the coarse resolution pixel, which is computed when the Sentinel 2 reflectance values are aggregated to Sentinel 3 SLSTR resolution. The extracted samples are used to train the thermal-reflectance relationships by using an ensemble decision tree regression method. The ensemble decision tree regression is performed both locally (in a moving-window fashion) and globally (i.e. to the whole study area). The results are then combined based on the residuals computed between the regression outputs and the low-resolution thermal data. Finally, residual analysis and bias correction are performed between the regression output and the low-resolution thermal data to ensure consistency between the sharpened high-resolution pixels and their corresponding low-resolution thermal pixel. Figure 15 illustrates the implemented processing workflow.

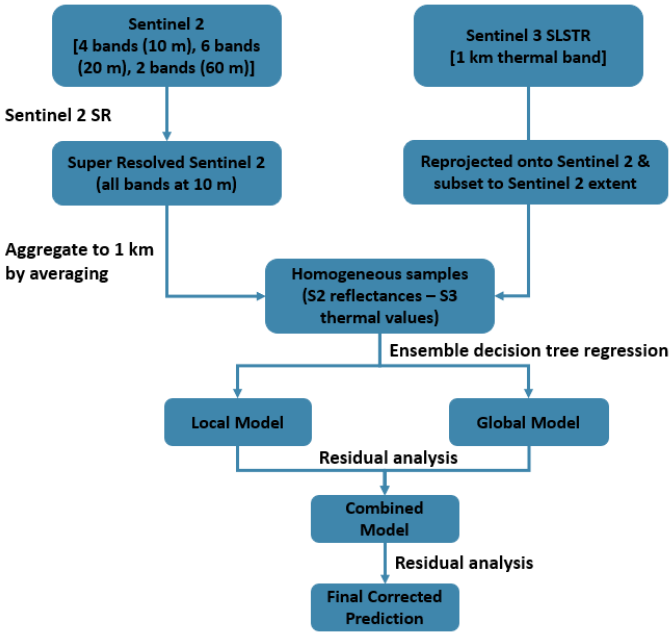

Figure 15. Sentinel 3 SLSTR & Sentinel 2 sharpening workflow

Funded by the European Union

2.2.3 Thermal sharpening Tool

In practice, the thermal sharpening tool requires a coarse resolution Sentinel 3 thermal image from SLSTR sensor and a fine resolution Sentinel 2 atmospherically corrected (Level 2A) image. Both images must be cloud free and temporally close to each other. The Sentinel 2 data are first super resolved to have all its spectral bands at the finer spatial resolution of 10 m with the implemented Sentinel 2 SR tool described in Section 2.1. On the other hand, the input Sentinel 3 SLSTR data are reprojected onto the projection system of Sentinel 2 data with the GDAL Python library. The spatially overlapping part of both images is used for the subsequent processing. For the ensemble decision tree regression, the scikit-learn library is employed, where each decision-tree is trained with a random subset of the training samples drawn with replacement in a method known as Bagging [31]. The final value of the regression is an average of the values produced within the ensemble.

The developed thermal sharpening tool was applied to Sentinel 3 & Sentinel 2 image pairs acquired over the 5 different Pilot's AOIs (Table 4). Table 5 presents the percentage of homogeneous samples extracted from each pair and the root mean square error from the residual analysis of the fused products with the coarse resolution Sentinel 3 thermal image. Figures 16-20 present subsets of the Sentinel 2 and Sentinel 3 image pairs and the predicted image from the developed thermal sharpening tool.

Table 4. The test Sentinel 3 – Sentinel 2 image pairs

Case Area		Datasets
Pilot 1	Sentinel 2	*S2A_MSIL2A_20220420T104631_N0400_R051_T31UFT_20220420T175659*
	Sentinel 3	*S3A_SL_2_LST____20220420T102858_20220420T103158_20220421T193709_0179_084_222_1980_PS1_O_NT_004*
Pilot 2	Sentinel 2	*S2A_MSIL2A_20220827T093601_N0400_R036_T35ULB_20220827T142655*
	Sentinel 3	*S3B_SL_2_LST____20220827T090451_20220827T090751_20220827T112211_0179_069_378_1980_PS2_O_NR_004*
Pilot 3	Sentinel 2	*S2B_MSIL2A_20210815T103629_N0301_R008_T31SDD_20210815T134410*
	Sentinel 3	*S3B_SL_2_LST____20210815T102531_20210815T102831_20210816T221544_0179_056_008_2340_LN2_O_NT_004*
Pilot 4	Sentinel 2	*S2B_MSIL2A_20220525T090559_N0400_R050_T34SGH_20220525T110355*
	Sentinel 3	*S3A_SL_2_LST____20220525T084630_20220525T084930_20220526T175411_0180_085_335_2340_PS1_O_NT_004*
Pilot 5	Sentinel 2	*S2B_MSIL2A_20220828T095549_N0400_R122_T35WMP_20220828T113542*
	Sentinel 3	*S3B_SL_2_LST____20220828T101640_20220828T101940_20220829T130617_0179_070_008_1800_PS2_O_NT_004*

Table 5. Evaluation results of the thermal sharpening product

Case Area	Homogeneous samples	RMSE
Pilot 1	40.6 % (4894/12051)	1.85857
Pilot 2	86.8% (10507/12100)	1.97244
Pilot 3	63.6% (1337/2101)	1.11956
Pilot 4	86.4% (5527/6397)	1.88751
Pilot 5	96.6% (11685/12098)	0.65909

Funded by the European Union

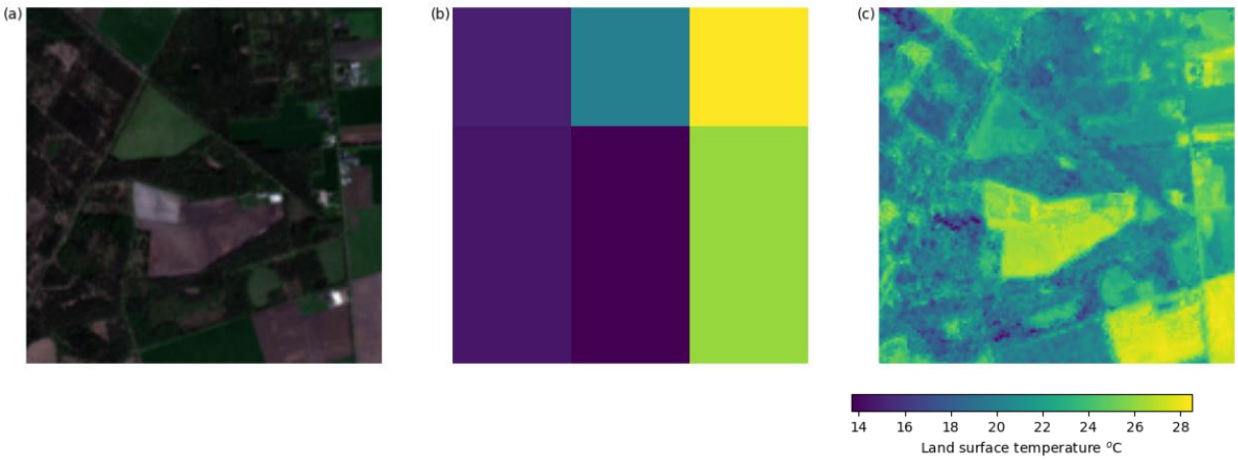

Figure 16 Pilot 1 (a) Sentinel 2 super-resolved image (RGB Bands 4,3,2), (b) Sentinel 3 SLSTR thermal image, (c) Combined prediction

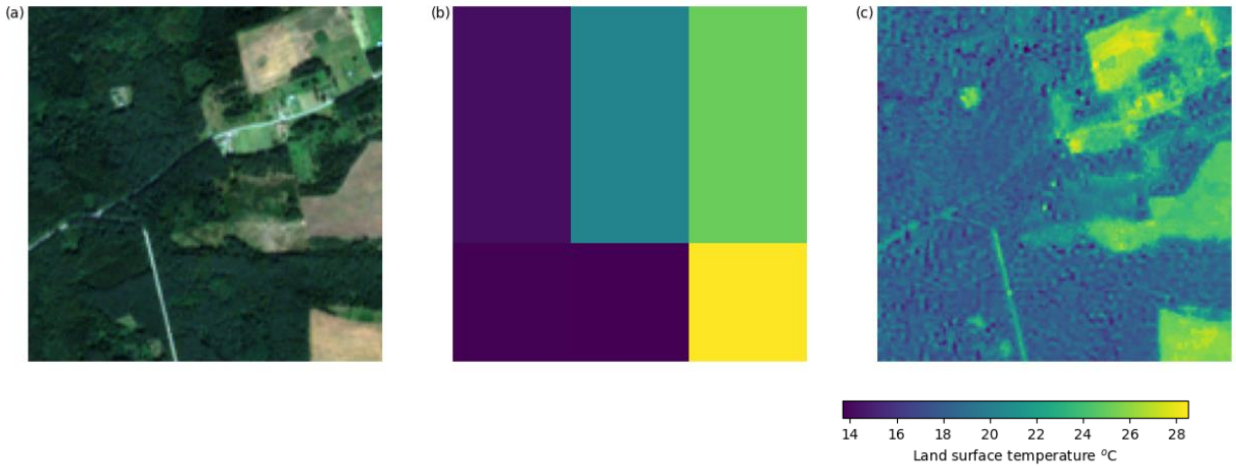

Figure 17 Pilot 2 (a) Sentinel 2 super-resolved image (RGB Bands 4,3,2), (b) Sentinel 3 SLSTR thermal image, (c) Combined prediction

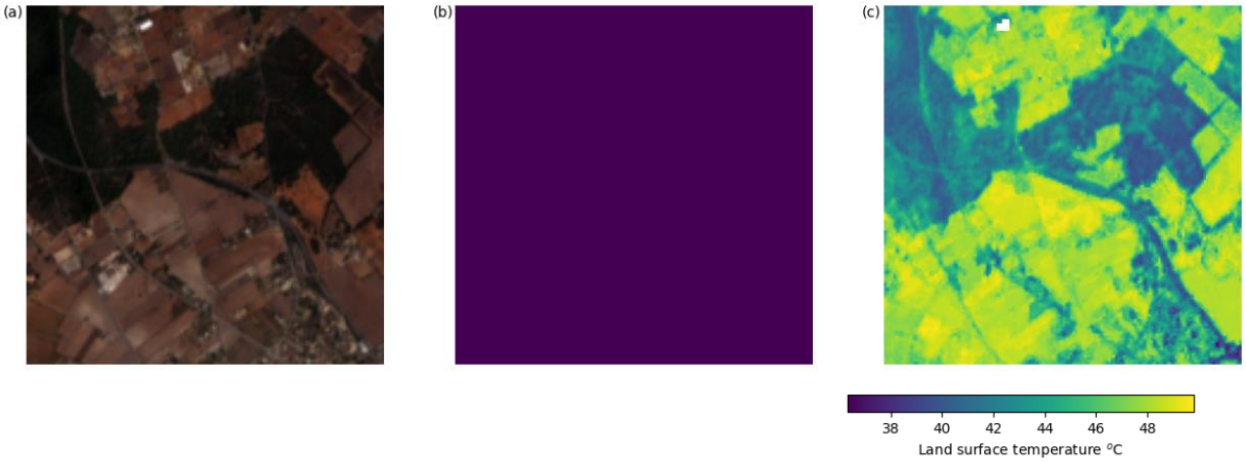

Figure 18 Pilot 3 (a) Sentinel 2 super-resolved image (RGB Bands 4,3,2), (b) Sentinel 3 SLSTR thermal image, (c) Combined prediction

Funded by the European Union

Figure 19 Pilot 4 (a) Sentinel 2 super-resolved image (RGB Bands 4,3,2), (b) Sentinel 3 SLSTR thermal image, (c) Combined prediction

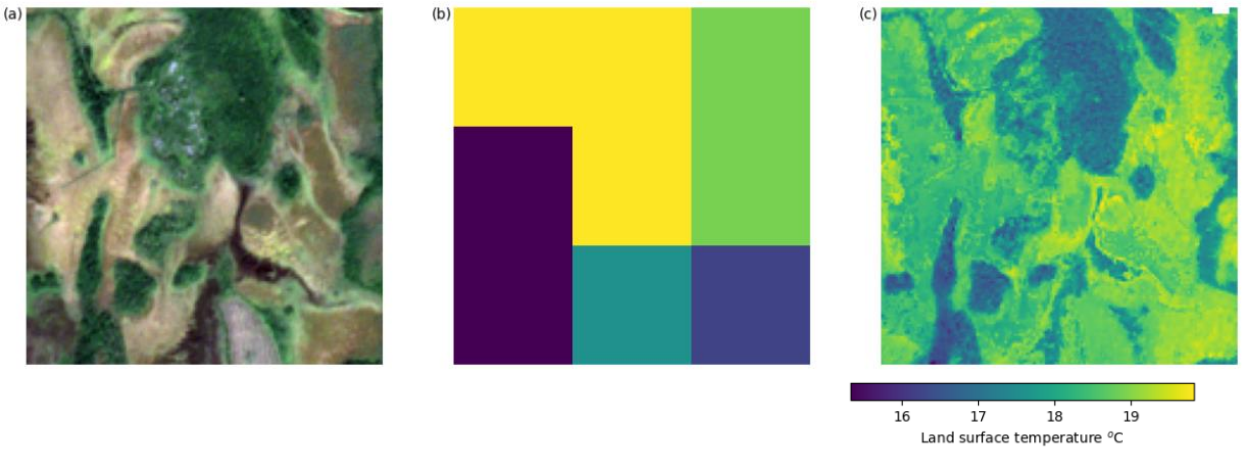

Figure 20 Pilot 5 (a) Sentinel 2 super-resolved image (RGB Bands 4,3,2), (b) Sentinel 3 SLSTR thermal image, (c) Combined prediction

Funded by the European Union

3 Future Work

Three tools for spatial augmentation and data fusion are currently under development: (i) a DL-based SR tool for augmenting the spatial resolution of Sentinel 3 OLCI and SYNERGY data. (ii) A spatio-temporal data fusion of Sentinel 2 and Sentinel 3 OLCI and SYNERGY data to create synthetic daily data that have the fine spatial resolution of Sentinel 2. And (iii) tools to enhance the spatial resolution of Sentinel 5p data products by fusing them with Sentinel 2 images, proxy and in-situ data, if available.

3.1 Sentinel 3 Augmentation & Fusion Tools

The first two components for spatial resolution augmentation and spatio-temporal fusion of Sentinel 3 medium spatial resolution (i.e. 300 m) multispectral images are suitable for areas that require the frequent acquisition of fine spatial resolution images, which no single sensor can achieve. Sentinel 2 satellites carry fine spatial resolution multispectral sensors and can revisit the same area every 5 days. Due to cloud and shadow contamination, however, it generally requires more than 5 days (e.g. probably several months) to acquire a cloud-free Sentinel 2 image for specific areas. On the other hand, Sentinel 3 also comprises a pair of satellites that carry the ocean and land color instrument (OLCI), which captures multispectral images with a spatial resolution of 300 m and a temporal resolution of 1.4 days. Therefore, intersensory schemes are considered to refine the available data products offered by Sentinel's operational chain. Specifically, the two Sentinel 3 OLCI related tools will try to spatially enhance Sentinel 3 OLCI products by taking advantage of the higher spatial resolution of the Sentinel 2 images and the higher temporal resolution.

To address these issues, two different tools are going to be implemented and tested for enhancing the four similar spectral bands that are acquired by both Sentinel 2 and Sentinel 3 sensors: (i) a DL-based super resolution for enhancing the Sentinel 3 OLCI spatial resolution, where Sentinel 2 fine resolution images are only accessible for training. And (ii) a spatio-temporal data fusion tool, which aims at creating synthetic images that combine the fine spatial resolution of Sentinel 2 imagery and the substantially shorter Sentinel 3 revisit time.

Effectively exploiting such intersensor synergies raises important challenges in terms of operational data availability, sensor alignment and substantial spatial and spectral resolution changes, among others [32], [33]. Each pair of Sentinel 2 – Sentinel 3 images must be atmospherically corrected, projected to a common coordinate system and co-registered to remove important spatial differences between the MSI and OLCI sensors. Concerning that atmospherically corrected Sentinel 3 OLCI products are still not available, it is necessary to predefine which atmospheric correction algorithm is suitable for each Pilot application. In M17, we discussed this issue with Pilot 5 users, which are interested in using Sentinel 3 related tools. And we decided to use the Sentinel 3 SYNERGY instead of OLCI products, which include atmospherically corrected bands acquired from both OLCI and SLSTR sensors. Therefore, a new training dataset with Sentinel 3 SYNERGY and Sentinel 2 image pairs was collected and processed to train the Sentinel 3 SR CNNs [8], [32]. Additionally, a spatio-temporal data fusion method [33] that can predict both gradual changes and land cover type changes will be implemented, which cannot be addressed with most spatiotemporal data fusion methods [34], [35].

Funded by the European Union

3.2 Sentinel 5p Augmentation & Fusion Tools

Spatial resolution augmentation of Sentinel 5p data is of particular interest for Pilot 3 and 4 applications. In Pilot 3, the measurements from a dense network of air quality ground stations can be potentially acquired real-time and used to infer ground level concentrations of various air pollutants from Sentinel 5p. On the other hand, Pilot 4 does not have in-situ data available but requires only the inference of ground level NO_2 concentrations at the spatial resolution of 1 km. The development of this tool runs in parallel with the activities of the Pilots (WP5) in order to ensure the proper integration with the available in-situ data.

Several types of methods are going to be developed to fulfil the requirements of both pilots. ML-based methods [36] can combine in-situ ground measurements with regression models to derive detailed information about the spatial distribution of air pollutants. DL-based algorithms [37], [38] trained with a large number of in-situ measurements collected from AirBASE (Air Quality e-Reporting data repositories maintained by the European Environmental Agency - EEA) can also be applied when in-situ measurements are not available. Additionally, model-based methods [39], [40] will also be tested that can infer ground-level concentrations from satellite data by using a chemical transport model instead of in-situ measurements.

Funded by the European Union

4 References

[1] C. Lanaras, J. Bioucas-Dias, E. Baltsavias, and K. Schindler, "Super-Resolution of Multispectral Multiresolution Images from a Single Sensor," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Aug. 2017, vol. 2017-July, pp. 1505–1513. doi: 10.1109/CVPRW.2017.194.

[2] A. D. Vaiopoulos and K. Karantzalos, "Pansharpening on the narrow VNIR and SWIR spectral bands of Sentinel-2," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2016, vol. 41, pp. 723–730. doi: 10.5194/isprsarchives-XLI-B7-723-2016.

[3] H. Park, J. Choi, N. Park, and S. Choi, "Sharpening the VNIR and SWIR bands of Sentinel-2A imagery through modified selected and synthesized band schemes," *Remote Sens (Basel)*, vol. 9, no. 10, Oct. 2017, doi: 10.3390/rs9101080.

[4] Q. Wang, W. Shi, Z. Li, and P. M. Atkinson, "Fusion of Sentinel-2 images," *Remote Sens Environ*, vol. 187, pp. 241–252, Dec. 2016, doi: 10.1016/j.rse.2016.10.030.

[5] C. Paris, J. Bioucas-Dias, and L. Bruzzone, "A HIERARCHICAL APPROACH TO SUPERRESOLUTION OF MULTISPECTRAL IMAGES WITH DIFFERENT SPATIAL RESOLUTIONS."

[6] M. O. Ulfarsson and M. Dalla Mura, "A low-rank method for sentinel-2 sharpening using cyclic descent," in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Oct. 2018, vol. 2018-July, pp. 8857–8860. doi: 10.1109/IGARSS.2018.8519256.

[7] N. Brodu, "Super-Resolving Multiresolution Images With Band-Independent Geometry of Multispectral Pixels," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 8, pp. 4610–4617, Aug. 2017, doi: 10.1109/TGRS.2017.2694881.

[8] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," *IEEE Trans Pattern Anal Mach Intell*, vol. 38, no. 2, pp. 295–307, Feb. 2016, doi: 10.1109/TPAMI.2015.2439281.

[9] J. Kim, J. K. Lee, and K. M. Lee, "Deeply-Recursive Convolutional Network for Image Super-Resolution," Nov. 2015, [Online]. Available: http://arxiv.org/abs/1511.04491

[10] J. Kim, J. K. Lee, and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," Nov. 2015, [Online]. Available: http://arxiv.org/abs/1511.04587

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, [Online]. Available: http://arxiv.org/abs/1512.03385

[12] X.-J. Mao, C. Shen, and Y.-B. Yang, "Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections," Mar. 2016, [Online]. Available: http://arxiv.org/abs/1603.09056

[13] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," Jul. 2017, doi: 10.48550/arxiv.1707.02921.

[14] B. Ayhan and C. Kwan, "Mastcam image resolution enhancement with application to disparity map generation for stereo images with different resolutions," *Sensors (Switzerland)*, vol. 19, no. 16, Aug. 2019, doi: 10.3390/s19163526.

[15] C. Lanaras, J. Bioucas-Dias, S. Galliani, E. Baltsavias, and K. Schindler, "Super-resolution of Sentinel-2 images: Learning a globally applicable deep neural network," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 305–319, Dec. 2018, doi: 10.1016/j.isprsjprs.2018.09.018.

Funded by the European Union

[16] M. Gargiulo, “Advances on CNN-based super-resolution of Sentinel-2 images,” Feb. 2019, doi: 10.48550/arxiv.1902.02513.

[17] F. Palsson, J. R. Sveinsson, and M. O. Ulfarsson, “Sentinel-2 image fusion using a deep residual network,” *Remote Sens (Basel)*, vol. 10, no. 8, Aug. 2018, doi: 10.3390/rs10081290.

[18] H. v. Nguyen, M. O. Ulfarsson, J. R. Sveinsson, and J. Sigurdsson, “Zero-Shot Sentinel-2 Sharpening Using a Symmetric Skipped Connection Convolutional Neural Network,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Sep. 2020, pp. 613–616. doi: 10.1109/IGARSS39084.2020.9323614.

[19] A. Panagiotopoulou, L. Grammatikopoulos, E. Charou, E. Bratsolis, N. Madamopoulos, and J. Petrogonas, “Very Deep Super-Resolution of Remotely Sensed Images with Mean Square Error and Var-norm Estimators as Loss Functions.”

[20] J. Wu, Z. He, and J. Hu, “Sentinel-2 sharpening via parallel residual network,” *Remote Sens (Basel)*, vol. 12, no. 2, Jan. 2020, doi: 10.3390/rs12020279.

[21] R. Zhang, G. Cavallaro, and J. Jitsev, “Super-Resolution of Large Volumes of Sentinel-2 Images with High Performance Distributed Deep Learning,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, Sep. 2020, pp. 617–620. doi: 10.1109/IGARSS39084.2020.9323734.

[22] L. Wald, T. Ranchin, and M. Mangolini, “Fusion of Satellite Images of Different Spatial Resolutions: Assessing the Quality of Resulting Images.”

[23] K. He, X. Zhang, S. Ren, and J. Sun, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” Feb. 2015, doi: 10.48550/arxiv.1502.01852.

[24] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” Dec. 2014, doi: 10.48550/arxiv.1412.6980.

[25] N. Agam, W. P. Kustas, M. C. Anderson, F. Li, and C. M. U. Neale, “A vegetation index based technique for spatial sharpening of thermal imagery,” *Remote Sens Environ*, vol. 107, no. 4, pp. 545–558, Apr. 2007, doi: 10.1016/j.rse.2006.10.006.

[26] A. Dominguez, J. Kleissl, J. C. Luvall, and D. L. Rickman, “High-resolution urban thermal sharpener (HUTS),” *Remote Sens Environ*, vol. 115, no. 7, pp. 1772–1780, Jul. 2011, doi: 10.1016/j.rse.2011.03.008.

[27] A. K. Inamdar, A. French, S. Hook, G. Vaughan, and W. Luckett, “Land surface temperature retrieval at high spatial and temporal resolutions over the southwestern United States,” *Journal of Geophysical Research Atmospheres*, vol. 113, no. 7, Apr. 2008, doi: 10.1029/2007JD009048.

[28] O. Merlin *et al.*, “Disaggregation of MODIS surface temperature over an agricultural area using a time series of Formosat-2 images,” *Remote Sens Environ*, vol. 114, no. 11, pp. 2500–2512, Nov. 2010, doi: 10.1016/j.rse.2010.05.025.

[29] K. Zakšek and K. Oštir, “Downscaling land surface temperature for urban heat island diurnal cycle analysis,” *Remote Sens Environ*, vol. 117, pp. 114–124, Feb. 2012, doi: 10.1016/j.rse.2011.05.027.

[30] F. Gao, W. P. Kustas, and M. C. Anderson, “A data mining approach for sharpening thermal satellite imagery over land,” *Remote Sens (Basel)*, vol. 4, no. 11, pp. 3287–3319, Nov. 2012, doi: 10.3390/rs4113287.

[31] L. Breiman, “Bagging Predictors,” 1996.

Funded by the European Union

[32] R. Fernandez, R. Fernandez-Beltran, J. Kang, and F. Pla, "Sentinel-3 Super-Resolution Based on Dense Multireceptive Channel Attention," *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, pp. 7359–7372, 2021, doi: 10.1109/JSTARS.2021.3097410.

[33] X. Zhu, E. H. Helmer, F. Gao, D. Liu, J. Chen, and M. A. Lefsky, "A flexible spatiotemporal method for fusing satellite images with different resolutions," *Remote Sens Environ*, vol. 172, pp. 165–177, Jan. 2016, doi: 10.1016/j.rse.2015.11.016.

[34] F. Gao, J. Masek, M. Schwaller, and F. Hall, "On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2207–2218, Aug. 2006, doi: 10.1109/TGRS.2006.872081.

[35] Q. Wang and P. M. Atkinson, "Spatio-temporal fusion for daily Sentinel-2 images," *Remote Sens Environ*, vol. 204, pp. 31–42, Jan. 2018, doi: 10.1016/j.rse.2017.10.046.

[36] K. Stebel *et al.*, "Samira-satellite based monitoring initiative for regional air quality," *Remote Sens (Basel)*, vol. 13, no. 11, Jun. 2021, doi: 10.3390/rs13112219.

[37] M. Kim, D. Brunner, and G. Kuhlmann, "Importance of satellite observations for high-resolution mapping of near-surface NO2 by machine learning," *Remote Sens Environ*, vol. 264, Oct. 2021, doi: 10.1016/j.rse.2021.112573.

[38] L. Scheibenreif, M. Mommert, and D. Borth, "Estimation of Air Pollution with Remote Sensing Data: Revealing Greenhouse Gas Emissions from Space," 2021.

[39] M. J. Cooper, R. v. Martin, C. A. McLinden, and J. R. Brook, "Inferring ground-level nitrogen dioxide concentrations at fine spatial resolution applied to the TROPOMI satellite instrument," *Environmental Research Letters*, vol. 15, no. 10, Oct. 2020, doi: 10.1088/1748-9326/aba3a5.

[40] L. N. Lamsal *et al.*, "Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument," *Journal of Geophysical Research Atmospheres*, vol. 113, no. 16, Aug. 2008, doi: 10.1029/2007JD009235.

Funded by the European Union

Appendix A. Sentinel 2 SR Training Dataset

Training Dataset

1. S2A_MSIL2A_20210314T032541_N0214_R018_T47NQF_20210314T071359.zip
2. S2A_MSIL2A_20210419T052641_N0300_R105_T43QDV_20210422T110554.zip
3. S2A_MSIL2A_20210725T021351_N0301_R060_T51NYH_20210725T062917.zip
4. S2A_MSIL2A_20210729T031541_N0301_R118_T48NTF_20210730T164044.zip
5. S2A_MSIL2A_20210930T082741_N0301_R021_T37TCN_20210930T114832.zip
6. S2A_MSIL2A_20211001T013721_N0301_R031_T52KEG_20211001T034715.zip
7. S2A_MSIL2A_20211001T162111_N0301_R040_T17TLH_20211001T210728.zip
8. S2A_MSIL2A_20211005T023551_N0301_R089_T51RTP_20211005T053933.zip
9. S2A_MSIL2A_20211005T092031_N0301_R093_T34SEG_20211005T122518.zip
10. S2A_MSIL2A_20211005T092031_N0301_R093_T35TLM_20211005T122518.zip
11. S2A_MSIL2A_20211005T143731_N0301_R096_T19HCC_20211005T183922.zip
12. S2A_MSIL2A_20211007T151711_N0301_R125_T19PCN_20211007T171210.zip
13. S2A_MSIL2A_20211008T060741_N0301_R134_T42SVG_20211008T081424.zip
14. S2A_MSIL2A_20211009T053731_N0301_R005_T44UPG_20211009T081123.zip
15. S2A_MSIL2A_20211009T085901_N0301_R007_T36UWE_20211009T120451.zip
16. S2A_MSIL2A_20211009T103941_N0301_R008_T31UGU_20211009T134536.zip
17. S2A_MSIL2A_20211010T071211_N0301_R020_T38KPG_20211010T092728.zip
18. S2A_MSIL2A_20211010T100941_N0301_R022_T33UVR_20211010T115015.zip
19. S2A_MSIL2A_20211013T083921_N0301_R064_T36RTV_20211013T113537.zip
20. S2A_MSIL2A_20211013T101951_N0301_R065_T32TNR_20211013T132717.zip
21. S2A_MSIL2A_20211013T184331_N0301_R070_T12UVF_20211013T213944.zip
22. S2B_MSIL2A_20210727T032539_N0301_R018_T48QWH_20210727T064635.zip
23. S2B_MSIL2A_20210731T050609_N0301_R076_T44NNP_20210731T070112.zip
24. S2B_MSIL2A_20210901T141049_N0301_R110_T19FCC_20210901T183101.zip
25. S2B_MSIL2A_20210902T221559_N0301_R129_T60HUB_20210902T235008.zip
26. S2B_MSIL2A_20210907T040539_N0301_R047_T47SNB_20210907T073543.zip
27. S2B_MSIL2A_20210924T190029_N0301_R013_T10TEQ_20210924T215332.zip
28. S2B_MSIL2A_20211001T102739_N0301_R108_T32TMT_20211001T132709.zip
29. S2B_MSIL2A_20211003T161019_N0301_R140_T18UUC_20211003T210214.zip
30. S2B_MSIL2A_20211004T172059_N0301_R012_T14TPM_20211006T162515.zip
31. S2B_MSIL2A_20211006T000239_N0301_R030_T56HKH_20211006T014333.zip
32. S2B_MSIL2A_20211006T112119_N0301_R037_T29SNC_20211006T132214.zip
33. S2B_MSIL2A_20211008T003709_N0301_R059_T54HUG_20211008T023756.zip
34. S2B_MSIL2A_20211008T140059_N0301_R067_T21LYJ_20211008T163719.zip
35. S2B_MSIL2A_20211008T170119_N0301_R069_T14RNN_20211008T211421.zip
36. S2B_MSIL2A_20211008T170119_N0301_R069_T14RQV_20211008T211421.zip

Validation Dataset

1. S2B_MSIL2A_20211009T095029_N0301_R079_T34UCV_20211009T120840.zip
2. S2B_MSIL2A_20211010T092029_N0301_R093_T33QYE_20211010T112716.zip
3. S2B_MSIL2A_20211010T105859_N0301_R094_T29PLM_20211010T141305.zip
4. S2B_MSIL2A_20211010T105859_N0301_R094_T30TXQ_20211010T123844.zip
5. S2B_MSIL2A_20211011T070819_N0301_R106_T39RVJ_20211011T095944.zip
6. S2B_MSIL2A_20211011T171139_N0301_R112_T14SNB_20211011T210903.zip
7. S2B_MSIL2A_20211011T185319_N0301_R113_T10SFG_20211011T215018.zip
8. S2B_MSIL2A_20211013T074849_N0301_R135_T37QED_20211013T101008.zip
9. S2B_MSIL2A_20211013T110919_N0301_R137_T30SUJ_20211013T125854.zip

Funded by the European Union

Test Dataset

1. S2A_MSIL2A_20210910T101031_N0301_R022_T35WMP_20210910T114151.zip
2. S2B_MSIL2A_20210815T103629_N0301_R008_T31SDD_20210815T134410.zip
3. S2B_MSIL2A_20210907T104619_N0301_R051_T31UFT_20210907T144922.zip
4. S2B_MSIL2A_20210926T094029_N0301_R036_T34UFG_20210926T110446.zip
5. S2B_MSIL2A_20210927T090659_N0301_R050_T35SKC_20210928T150645.zip

Funded by the European Union