QGIS AND OPEN DATA CUBE APPLICATIONS FOR LOCAL CLIMATE ZONES ANALYSIS LEVERAGING PRISMA HYPERSPECTRAL SATELLITE DATA

D. Oxoli[1] , J. R. Cedeno Jimenez[1] *, E. Capizzi[1] , M. A. Brovelli[1] , M. Siciliani de Cumis[2] , P. Sacco[2] , D. Tapete[3]

[1] Dept. of Civil and Environmental Engineering, Politecnico di Milano, Milano Leonardo. Milan, Italy - (daniele.oxoli, jesusrodrigo.cedeno, emanuele.capizzi, maria.brovelli)@polimi.it
[2] Italian Space Agency (ASI), Matera, Italy - (mario.sicilianidecumis, patrizia.sacco)@asi.it
[3] Italian Space Agency (ASI), Roma, Italy - deodato.tapete@asi.it

KEY WORDS: Local Climate Zones, Open Data Cube, QGIS, PRISMA Satellite.

ABSTRACT:

Climate change poses a significant threat to humans and biodiversity, impacting various aspects of livelihoods, infrastructure, and ecosystems. Understanding climate change and its interaction with the environment is crucial for achieving Sustainable Development Goals. Local Climate Zones (LCZ) play a key role in comprehending climate change by categorizing urban areas also based on their thermal characteristics. This study presents prototype open-source software tools developed to integrate ground and satellite data for LCZ analysis in the Metropolitan City of Milan (Northern Italy). These tools consist of a QGIS plugin to access and preprocess ground-based meteorological sensor data and a client-server platform, based on the Open Data Cube and Docker technologies, for the exploitation of multispectral and hyperspectral satellite data in LCZ mapping and analysis. The tools' architecture, data retrieval methods, and analysis capabilities are described in detail. The QGIS plugin facilitates the access and preprocessing of ground-based sensor data within the user-friendly QGIS environment. The platform enables seamless ground-sensor and satellite data management and analysis, using Jupyter Notebooks as an interface to support programmatic operations on the data. The proposed tools provide a framework for studying climate change and its local impacts on urban environments, with the potential of empowering users to effectively analyze and mitigate its effects.

1. INTRODUCTION

Climate change poses an escalating threat to both human population and biodiversity. Its impacts are intensifying and spreading rapidly. On one hand, climate change adversely affects human livelihoods, settlements, infrastructure, and food production. On the other hand, it induces transformations in ecosystems requiring urgent implementation of adaptation strategies. These challenges present substantial barriers to the achievement of the Sustainable Development Goals (SDG) (Zahari et al., 2023) also in urbanized areas, specifically SDG 11 (sustainable cities and communities) (Oxoli et al., 2018) and SDG 13 which aims to take action to cope with climate change and its impacts (UN, 2022). In this context, it is relevant to understand climate change's interaction with urban features and possible negative effects on the urban population. Intensification of heat wave phenomena is among the most widespread climate change-related threats for cities (Founda et al., 2019).

Local Climate Zones (LCZs) are a key method for understanding climate change, especially in urbanized areas. The LCZs propose a classification schema to identify homogenous inner city areas in terms of vegetation cover, impervious surface fractions, building height and texture, etc. (Stewart and Oke, 2012). These characteristics can affect how the urban areas respond to changing climate over time (Puche et al., 2023). LCZs have been primarily used to study the urban heat island effect, that is the phenomenon by which urban areas are warmer than the surrounding rural areas. The urban heat island effect is generally caused by the concentration of impervious surfaces and the lack of vegetation in urban areas (Mohajerani et al., 2017). To that end, LCZs can be used to identify areas where the urban heat island effect is most severe and -in turn - provide valuable inputs to develop effective local strategies for urban heat island mitigation.

Among the possible strategies to compute LCZs, one commonly used approach involves utilizing remote sensing data and geospatial analysis techniques. By leveraging satellite imagery and complementary environmental variables (e.g. topographic databases, meteorological sensors data, etc.) it is possible to classify, contextualise and validate LCZ maps (Zhang et al., 2019). A frontier topic in satellite imagery analysis is the combination of multispectral and hyperspectral data to enhance scene classification through increased spectral resolutions, thus improving the identification of surface materials. These assets, coupled with the application of machine learning algorithms to automate the classification process, have been supported accurate mapping of LCZs (Yuan et al., 2022).

Nevertheless, access and manipulation of the heterogeneous geospatial data required by LCZ mapping (including satellite imagery and meteorological sensor observations) may represent a demanding task, especially for non-expert users. To that end, the present work aims at the development of software tools conceived for the monitoring and analysis of LCZs.

The tools are developed within the project LCZ-ODC, funded by the Italian Space Agency (ASI). The selected study area for the first implementation of the tools is the Metropolitan City of Milan (Lombardy region, Northern Italy). This area is particularly significant to be used as a first approach to the LCZ-ODC project, due to its extensive urbanization and industrialisation. Furthermore, the Metropolitan City of Milan is located in the upper Po River valley where territorial characteristics (including low wind circulation) are additional potential drivers for

* Corresponding author

heat wave events (Cedeno Jimenez et al., 2023).

The main purpose of the proposed software tools is to facilitate access and integration of local meteorological sensors data from the Environmental Protection Agency of Lombardy Region (ARPA Lombardia) and the satellite hyperspectral data provided by the ASI PRISMA (PRecursore IperSpettrale della Missione Applicativa: Hyperspectral Precursor of the Applicative Mission) mission (`https://www.asi.it/scienze-della-terra/prisma`) (Cogliati et al., 2021)) to investigate their support in LCZ mapping and analysis.

The developed tools compose of two main applications as follows. The first is a QGIS (`https://www.qgis.org`) plugin named ARPA Weather plugin. It is designed to facilitate access to openly available data from the ARPA Lombardia sensor networks and their manipulation within QGIS. The plugin simplifies the processes of downloading and preprocessing local meteorological sensor observations. The second component is a client-server platform for data integration named LCZ-ODC. This component uses the Open Data Cube (ODC) software to provide users with a ready-to-use set of tools to exploit analysis-ready satellite data, ground sensors observation and ancillary geospatial layer for LCZ mapping and analysis. A portable ODC implementation correlated by a Jupyter Notebook frontend is created using a Docker container to facilitate its installation on any desktop or server computer. Both tools rely on Free and Open Source Software (FOSS) technologies and are released with an open license.

The remainder of the paper is as follows. Section 2 contains a detailed description of development patterns and the technologies used for the proposed software tools. Section 3 presents an overview of the preliminary results obtained, while conclusions and outlook are discussed in Section 4.

2. SOFTWARE TOOLS DESIGN AND DEVELOPMENT

2.1 ARPA Weather QGIS plugin

Ground-based meteorological sensor networks are essential in monitoring local weather conditions and climate patterns and - in turn - analysing LCZs. Integrating their data into Geographic Information System (GIS) environments is crucial for supporting multiple applications such as urban planning, public health studies, and weather forecasting (Acosta et al., 2021).

Generally, meteorological sensor networks utilize scattered geolocalized sensors that measure multiple atmospheric variables, including air temperature, wind speed, and precipitation. The data collected by these sensors is often made available online by network managers, which can include local or national authorities, private companies, or volunteers. However, the diversity of data providers leads to heterogeneity in data formats and access patterns for meteorological sensor data, resulting in time-consuming preprocessing tasks for users, such as downloading, temporal aggregations and spatial filtering.

To address these challenges and improve the accessibility and usability of ground-based sensor data for end-users, the ARPA Weather QGIS plugin was developed. This plugin aims to simplify the access and preprocessing of openly available data from the meteorological sensor networks, enabling users to directly utilize the data within the QGIS software.

The current version of the ARPA Weather plugin focuses specifically on data provided by the ARPA Lombardia. It leverages the Open Data Lombardia portal (`https://dati.lombardia.it`) to access sensor observations and related information, such as the sensors registry and the observations archive.

For the current month's data, the plugin utilizes the Socrata Open Data API (`https://dev.socrata.com`) to request the observed time series. Archived time series from previous years/months can be downloaded as large CSV files directly from the Open Data Lombardia portal. However, the usability of the API data is limited for non-expert users, while the large CSV files represent a challenge for manual processing due to the data size.

The ARPA Weather plugin addresses these issues by providing access to both the API and CSV files. It automates data retrieval from the API using the Sodapy Python library (`https://github.com/xmunoz/sodapy`), which is a Python client for the Socrata Open Data API. The plugin leverages popular Python libraries such as Pandas (`https://pandas.pydata.org`) and Dask (`https://www.dask.org`) for efficient data handling and parallel computing to speed up I/O operations on large datasets.

With the plugin, users can easily manage different data sources by selecting mandatory parameters such as the year of processing, time period, and sensor type (see Figure 2). To prevent the download of large, unmanageable CSV files, the plugin retrieves data for each year separately. Additionally, users can decide to perform optional pre-processing operations such as the outlier removals and spatial filtering on Lombardy region provinces. Available functionalities can be accessed from an explanatory graphical user interface, shown in Figure 1.

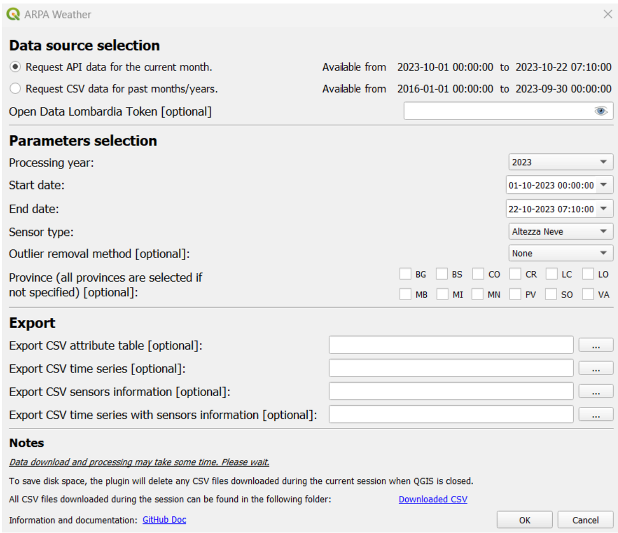

Figure 1. Graphical user interface of the ARPA Weather QGIS plugin.

The ARPA Weather plugin automatically generates a multipoint layer in QGIS as output, containing summary statistics (e.g., mean, maximum, minimum) of the selected sensors' observations within the selected time period (see Figure 3). Optionally, the numerical observations can be merged with the complete sensor registry information. Furthermore, it allows users to export various products in CSV format, including the full sensors' observations time series.

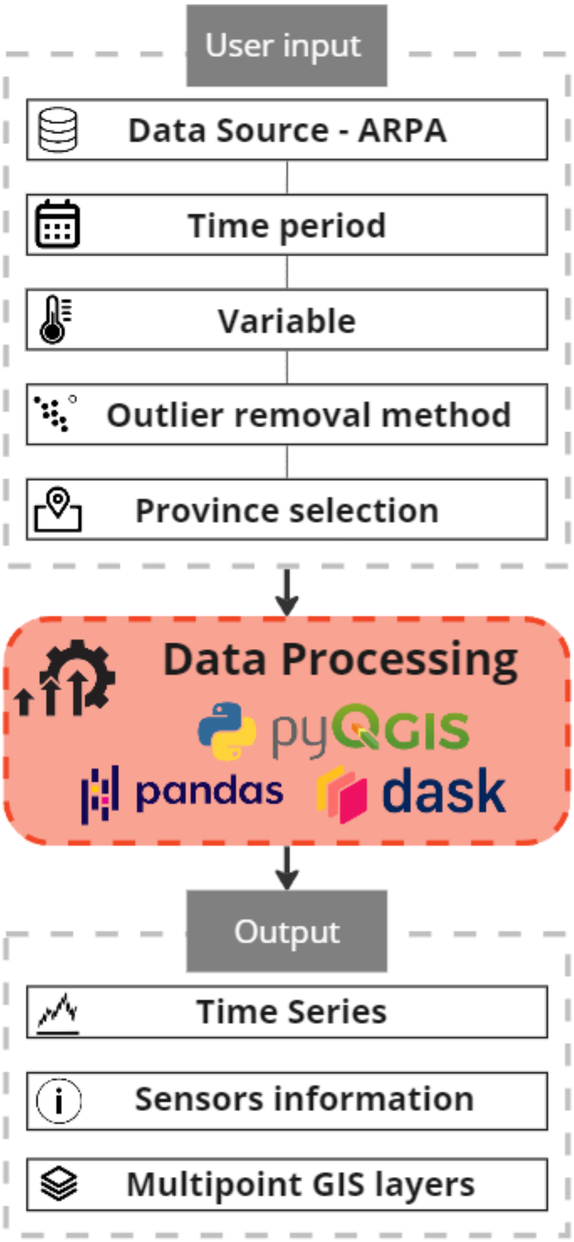

Figure 2. Plugin synthetic workflow: the top part displays user-selectable inputs including data source, time period, variable, outlier removal method, and the province. The output files (time series, sensors information, multipoint GIS layers) are processed using FOSS Python libraries (i.e.pandas, dask, pyQGIS).

2.2 The LCZ-ODC platform

The ODC was selected as the core technology in this project to enhance data management and extraction operations. The ODC (`https://www.opendatacube.org`) is a comprehensive collection of tools designed to manage, process, and analyse satellite data and beyond. One of the principal advantages of the ODC relies upon its data management module which allows for handling multi-layer and multi-temporal gridded data. This multi-layer capability enables a comprehensive understanding of complex geospatial phenomena by overlaying and analyzing various data sources simultaneously. Users can effortlessly

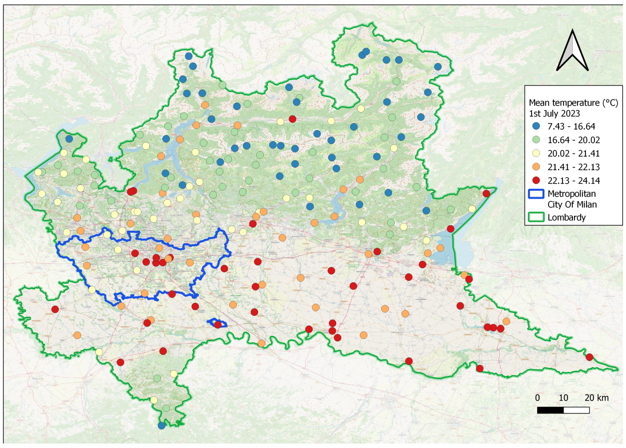

Figure 3. Example of multipoint layer obtained using ARPA Weather plugin - Temperature data 1st July 2023.

access and compare historical and current geospatial data, facilitating detailed temporal analysis and the identification of long-term trends and patterns.

Operating within a Python-based environment, the ODC offers an array of capabilities that streamline the handling of satellite imagery. Its purpose is to centralize and harmonize diverse datasets, allowing users to efficiently discover, access, and utilize geospatial information. By integrating various geospatial sources, the ODC facilitates the creation of comprehensive and consistent data catalogues. These catalogues encompass tasks such as data ingestion, storage optimization, and metadata management. By establishing a standardized framework, the ODC simplifies the otherwise complex process of handling diverse data formats and resolutions. Moreover, the ODC empowers users with advanced processing and analysis capabilities. By utilising Python-based toolsets, practitioners can efficiently carry out geospatial operations such as spatial and temporal queries, data fusion, and change detection.

The ODC environment further enhances its accessibility and extensibility by allowing users to seamlessly integrate the ODC with several data processing and analysis libraries, enabling holistic and customizable geospatial workflow.

First-time users of this software often find it challenging to install the ODC locally or on a server. Moreover, the setup procedure to use the ODC with non-satellite data has been explored in other works (Cedeno Jimenez et al., 2021), but it requires a high level of programming expertise.

To simplify the process, the presented LCZ-ODC platform provides a Docker image that contains everything that is needed to run the LCZ applications developed for this project. Docker (`https://www.docker.com`) is a software platform that allows users to build, run, and deploy applications in containers. Containers are isolated environments that include everything needed to run an application. This means that users can easily run an ODC application without having to worry about installing or configuring the ODC software on their local machine or server. The LCZ-ODC platform embeds a series of interactive processing scripts which are exposed to the platform front-end using Jupyter Notebook, a popular client-server application for interactive programming. The processing notebooks are written in Python, incorporated into the platform architecture and connected with the ODC, thus enabling simplified and programmatic access to the ODC database.

Processing functionalities include reading, re-gridding, georeferencing, plotting, overlaying and classifying functions that can be applied to ODC analysis-ready data, such as the one retrieved from ARPA Lombardia and the PRISMA satellite images.

The LCZ-ODC platform aims not only to provide users with a processing dashboard but also to train them on how effectively handle and analyze the data. Using Jupyter Notebooks provides several advantages. Firstly, Jupyter Notebooks offer an interactive and exploratory coding environment, allowing users to execute code snippets and immediately see the output, facilitating rapid prototyping and iterative development. Secondly, Jupyter Notebooks support the seamless integration of code, visualizations, and explanatory text within a single document, making it an excellent tool for creating and sharing not only numerical analyses but also data-driven narratives and reports.

Figure 4 depicts the architecture and workflow of the LCZ-ODC platform with its main components (ODC, Docker, and Jupyter notebooks.) The "External Volume" component represents a designated location in the user's local file system, housing the data to be utilized. Specifically, this volume contains PRISMA satellite images and ground sensor data from ARPA Lombardia, in their original raw format obtained directly from the provider. Complementary data, including topographic maps and Sentinel-2 multispectral satellite images, are also conveyed into the LCZ-ODC to support LCZ mapping. The output of this component is then directed to the Docker instance, which comprises the Linux Operating System (Ubuntu 20.04) as the foundation, along with ODC and a PostgreSQL database. Within PostgreSQL, ODC manages a collection of tables storing metadata and file locations within the local file system. The third and final component, Jupyter, encompasses a series of pipelines that leverage the Docker structure and ODC capabilities to discover, extract, process, and deliver data to the user. Additionally, within this component, users have access to Jupyter Notebooks enabling them to perform data visualization and analysis from a browser, without the need to install any software locally on their own computers or servers.

Within the aforementioned framework, the ground sensor data from ARPA Lombardia is obtained and seamlessly transferred to the LCZ-ODC External Volume. These data transfer and formatting mechanisms are facilitated by the QGIS Plugin discussed in Section 2.1. The transfer of PRISMA satellite imagery is operated instead by dedicated pre-processing notebooks. The end-user of the LCZ-ODC platform are therefore exempt from any data cleaning and preprocessing tasks.

3. PRELIMINARY RESULTS

The preliminary outcomes of the software tools design and development phase consist of the tools' source code and documentation, published under the MIT open license (`https://mit-license.org`) on GitHub: a) ARPA Weather QGIS plugin (`https://github.com/gisgeolab/ARPA_Weather_plugin`) and b) LCZ-ODC platform (`https://github.com/gisgeolab/LCZ-ODC`). The documentation includes both user and developer's instructions to install, use, replicate and improve the tools.

The beta version of the ARPA Weather plugin can be downloaded and installed in QGIS by any interested users as an alternative tool for retrieval, management, and analysis of ARPA

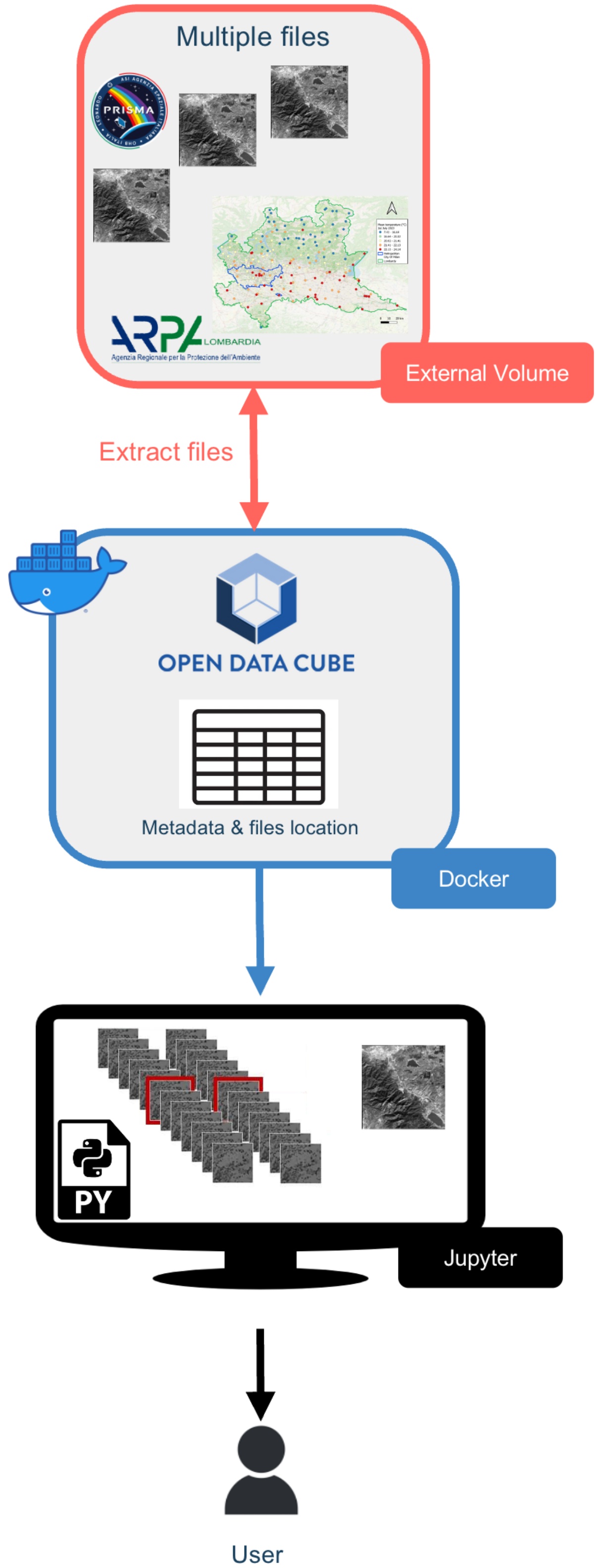

Figure 4. Schematic of the LCZ-ODC platform architecture to ingest, extract and analyse PRISMA and ARPA Lombardia data for generating LCZ maps.

Lombardia meteorological sensor data, supporting diverse applications of local climate monitoring and research. Additionally, the plugin functionalities have been included in Jupyter Notebooks to allow users to access and use them also in standard Python scripting (`https://github.com/gisgeolab/ARPA_Weather_plugin/tree/ARPA_Weather_Notebook`).

The LCZ-ODC platform is presently in the development phase. It is designed as an illustrative example of ODC technologies employed in LCZ mapping through the integration of multi-sources and multi-temporal geospatial and satellite data. A prototype of the platform is scheduled for online release to facilitate testing, involving pilot users participating in the LCZ-ODC project. These users encompass climate experts from regional environmental agencies, private companies and also urban planners. The objective is to promote the utilization of cutting-edge satellite data and LCZ mapping workflow as a practical tool to support real urban-climate mitigation initiatives.

Local installation of the LCZ-ODC platform for adaptation and reuse in different case studies is anyway supported by the release of a portable Docker container of the whole system. LCZ-ODC users' functionalities are included in Jupyter Notebooks and can be accessed through a dedicated branch of the software repository (`https://github.com/gisgeolab/LCZ-ODC/tree/Processing-Notebooks`).

4. CONCLUSIONS AND FUTURE WORK

This work introduced a software tool framework specifically designed for monitoring and analyzing LCZs, with a particular focus on its application on the Metropolitan City of Milan. The framework comprises two key components: a QGIS plugin for accessing and preprocessing meteorological ground sensor data from ARPA Lombardia, and a client-server platform based on ODC technology to integrate various geospatial and satellite data sources for LCZ mapping.

The combination of ground-based sensor data and hyperspectral satellite imagery offers a promising avenue for gaining a comprehensive understanding of climate change dynamics and their implications for urban areas. The software tools presented in this study provide data manipulation, visualization, and analysis capabilities that can be utilized by non-expert end-users as well, enabling them to extract valuable insights into the impact of climate change on cities. The integration of ODC and Docker ensures a replicable infrastructure that facilitates efficient management and analysis of complex geospatial data.

Moving forward, future research endeavours could extend the geographical scope of the study to encompass additional regions and cities, allowing for broader analysis and comparisons. Furthermore, the software and tools developed here can be enhanced by incorporating advanced algorithms and machine learning techniques to further improve data processing and analysis capabilities.

In summary, this study aims to contribute to the growing body of knowledge on the intricate relationship between climate change and the urban environment, by promoting its systematic consideration within urban urban planning practices.

The software framework and tools presented in this work serve as valuable resources for researchers, policymakers, and practitioners involved in climate change mitigation and adaptation efforts. These resources support evidence-based decision-making and foster sustainable development, playing a crucial role in addressing the challenges posed by climate change in cities.

ACKNOWLEDGEMENTS

This work was developed within the LCZ-ODC project, funded by the Italian Space Agency (agreement n. 2022-30-HH.0).

REFERENCES

Acosta, M. P., Vahdatikhaki, F., Santos, J., Hammad, A., Dorée, A. G., 2021. How to bring UHI to the urban planning table? A data-driven modeling approach. *Sustainable Cities and Society*, 71, 102948.

Cedeno Jimenez, J. R., Oxoli, D., Brovelli, M. A., 2021. ENABLING AIR QUALITY MONITORING WITH THE OPEN DATA CUBE: IMPLEMENTATION FOR SENTINEL-5P AND GROUND SENSOR OBSERVATIONS. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVI-4/W2-2021, 31–36. https://isprs-archives.copernicus.org/articles/XLVI-4-W2-2021/31/2021/.

Cedeno Jimenez, J. R., Pugliese Viloria, A. d. J., Brovelli, M. A., 2023. Estimating Daily NO2 Ground Level Concentrations Using Sentinel-5P and Ground Sensor Meteorological Measurements. *ISPRS International Journal of Geo-Information*, 12(3). https://www.mdpi.com/2220-9964/12/3/107.

Cogliati, S., Sarti, F., Chiarantini, L., Cosi, M., Lorusso, R., Lopinto, E., Miglietta, F., Genesio, L., Guanter, L., Damm, A., Pérez-López, S., Scheffler, D., Tagliabue, G., Panigada, C., Rascher, U., Dowling, T., Giardino, C., Colombo, R., 2021. The PRISMA imaging spectroscopy mission: overview and first performance analysis. *Remote Sensing of Environment*, 262, 112499. https://doi.org/10.1016/j.rse.2021.112499.

Founda, D., Pierros, F., Katavoutas, G., Keramitsoglou, I., 2019. Observed trends in thermal stress at European cities with different background climates. *Atmosphere*, 10(8), 436.

Mohajerani, A., Bakaric, J., Jeffrey-Bailey, T., 2017. The urban heat island effect, its causes, and mitigation, with reference to the thermal properties of asphalt concrete. *Journal of environmental management*, 197, 522–538.

Oxoli, D., Ronchetti, G., Minghini, M., Molinari, M., Lotfian, M., Sona, G., Brovelli, M., 2018. Measuring Urban Land Cover Influence on Air Temperature through Multiple Geo-Data—The Case of Milan, Italy. *ISPRS International Journal of Geo-Information*, 7(11), 421. http://dx.doi.org/10.3390/ijgi7110421.

Puche, M., Vavassori, A., Brovelli, M. A., 2023. Insights into the Effect of Urban Morphology and Land Cover on Land Surface and Air Temperatures in the Metropolitan City of Milan (Italy) Using Satellite Imagery and In Situ Measurements. *Remote Sensing*, 15(3), 733.

Stewart, I. D., Oke, T. R., 2012. Local climate zones for urban temperature studies. *Bulletin of the American Meteorological Society*, 93(12), 1879–1900.

UN, 2022. The 17 Goals, Sustainable Development. More at: https://sdgs.un.org/goals. Accessed on: 2023-02-13.

Yuan, J., Wang, S., Wu, C., Xu, Y., 2022. Fine-grained classification of urban functional zones and landscape pattern analysis using hyperspectral satellite imagery: A case study of wuhan. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3972–3991.

Zahari, H. M., Zain, R., Ismail, A., Zainol, N. A. M., Yaacob, S. B., Ali, N. I. R. C., 2023. Climate change and the whole government approach to be resilient towards disaster. *Handbook of Research on Promoting Logistics and Supply Chain Resilience Through Digital Transformation*, IGI Global, 111–120.

Zhang, G., Ghamisi, P., Zhu, X. X., 2019. Fusion of heterogeneous earth observation data for the classification of local climate zones. *IEEE Transactions on Geoscience and Remote Sensing*, 57(10), 7623–7642.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-111-2023