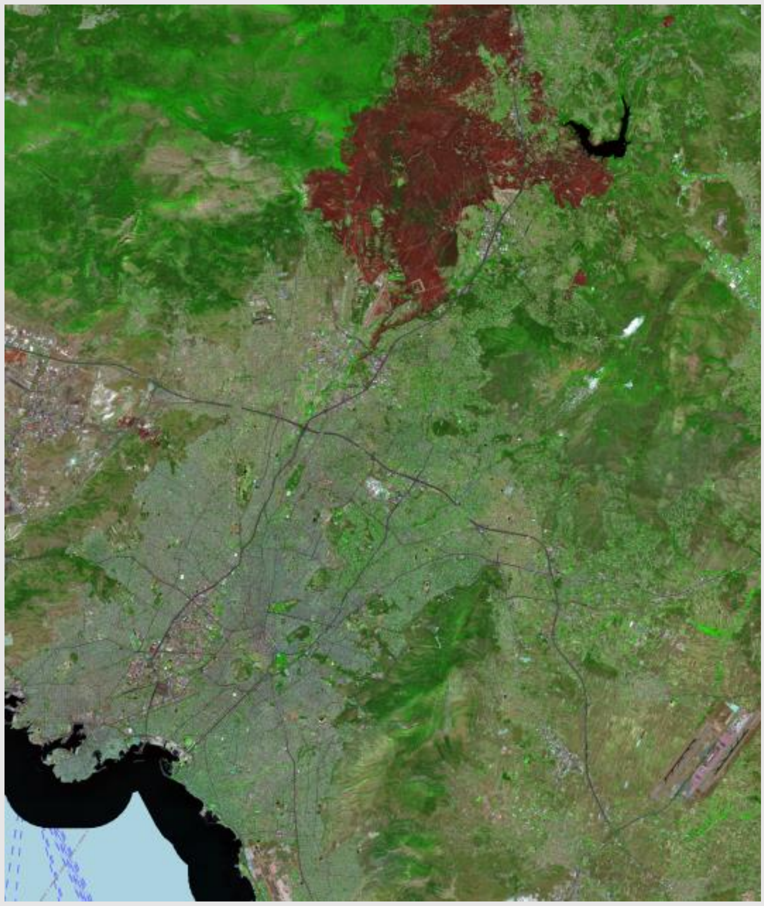

Assimila ClearSky Maps / Copernicus Sentinel data

Requirements and availability on data acquisition, transformation & intelligence extraction

Deliverable 2.3

Work package: WP2

Dissemination level: PU

Lead partner: Assimila

Authors: Gerardo Lopez Saldana, Jon Styles

Due date: 31.12.2021

Submission date: 24.12.2021

Deliverable	**Requirements and availability on data acquisition transformation & intelligence extraction**
Deliverable No.	D2.3
Work Package	2
Dissemination Level	PU
Author(s)	Jon Styles (Assimila), Gerardo Lopez Saldana (Assimila),
Co-Author(s)	Andy Shaw (Assimila), Stefano Salvi (INGV), Ikonen Jaako (FMI), Zina Mitraka (FORTH)
Date	24/12/2021
File Name	WP2_D2_3_Requirements_availability_data_acquisition_transformation_ASSIMILA_v1.0_final
Status	Final
Revision	1.0
Reviewed by (if applicable)	GSH, KTH
Information to be used for citations of this report	Gerardo Lopez Saldana, Jon Styles. (2021): *Requirements and availability on data acquisition transformation & intelligence extraction*, D2.3, HARMONIA. Horizon 2020 Grant Agreement No 101003517, European Commission, 84 pp.

The HARMONIA project has received funding from the EU Horizon 2020 research and innovation programme under agreement No. 101003517.

CONTACT

Email: contact@harmonia-project.eu
Website: http://harmonia-project.eu/

Summary

Figures

Tables

Document revision history

Version	Date	Modification reason	Modified by
1	24/12/20201	First Edition	Jon Styles, Gerardo Lopez Saldana
2			
3			

Partners organisations

No.	Name	Short name	Country
1	Assimila Limited	ASSIMILA	UK
2	Foundation for Research and Technology Hellas	FORTH	Greece
3	Ilmatieteen Laitos	FMI	Finland
4	Istituto Nazionale Di Geofisica E Vulcanologia	INGV	Italy
5	Institute Of Communication And Computer Systems	ICCS	Greece
6	Geosystems Hellas	GSH	Greece
7	Kungliga Tekniska Högskolan	KTH	Sweden

Abbreviations and definitions

AI	Artificial Intelligence
API	Application Programming Interface
BOA	Bottom Of the Atmosphere
BRDF	Bidirectional Reflectance Distribution Function
CAMS	Copernicus Atmosphere Monitoring Service
C3S	Copernicus Climate Change Services
CH	Chapter
CMIP	Coupled Model Intercomparison Project
CORDEX	Coordinated Regional Climate Downscaling Experiment
D	Deliverable
DIAS	Data and Information Access Services
DEM	Digital Elevation Model
DSS	Decision Support System
ECMWF	European Centre for Medium-Range Weather Forecasts
EO	Earth Observation
ERA	European Reanalysis
EU	European Union
FAIR	Findability, Accessibility, Interoperability, and Reusability
FWI	Fire Weather Index
GI	Green Infrastructure
IPCC	Intergovernmental Panel on Climate Change
LAI	Leaf Area Index
LP DAAC	Land Processes Distributed Active Archive Center
LST	Land Surface Temperature
ML	Machine Learning
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration

NBS	Nature-Based Solutions
OLCI	Ocean and Land Colour Instrument
PROSAIL	Combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model, also referred to as PROSAIL
RA	ReAnalysis. Reanalysis data provide the most complete picture currently possible of past weather and climate. They are a blend of observations with past short-range weather forecasts rerun with modern weather forecasting models.
RCM	Regional Climate Model
RTM	Radiative Transfer Model
S1	Sentinel-1
S2	Sentinel-2
SAR	Synthetic Aperture Radar
SDG	Sustainable Development Goal
SIAC	Sensor Invariant Atmospheric Correction
SLSTR	Sea and Land Surface Temperature Radiometer
SR	Surface Reflectance
TIP	Two-stream Inverse Package
TOA	Top of the Atmosphere
VIIRS	Visible Infrared Imaging Radiometer Suite
VHR	Very High Resolution. EO data with spatial resolution below 10m pixel size.
WP	Work Package

Executive Summary

This document describes the outputs of HARMONIA Task 2.2 which is concerned with analysis of the data and processing models and algorithms that are needed to fulfil the HARMONIA applications.

During the HARMONIA project a wide range of applications and services are being developed to help manage the risks to urban populations from changes in climate including atmospheric (thermal and pollution), hydrological (flooding) and geophysical (subsidence, landslides, and land deformation). These applications require a complex array of algorithms and models to transform Earth observation and other data into risk and ultimately into actionable information. The main objective of this task is to analyse the different applications data and model requirements, identify the areas of relative maturity where we are confident that the applications can be fulfilled, and to highlight those gaps, or areas of less certainty where the research and development activities are needed. In summary the objectives of the task are to:

- Undertake an analysis of the data sources required for the HARMONIA applications
- Survey the analysis needs of the applications, including the processing models that are needed to go from input data to impacts indicators, including an analysis of the adequacy of the models and data to deliver results at the required spatial scale
- Undertake a gap analysis to establish areas where the currently identified data and models require further development to meet the needs of users.

The five thematic Applications in WP5 are therefore the key points in which data and model requirements can be elaborated. For each of the five Applications, data and model requirements have been structured into a number of types. Data are classified into three types: EO data, in situ observations, and pre-processed or modelled datasets (for example reanalyses). For each application, the intended data and models to be used have been documented in a summary graphical form, and in a set of detailed descriptive tables which are included as an annex in section 7.

A summary of all these data types and models has then been made in order to establish those that are used by a large number of applications, so that these commonalities can be exploited to ensure the efficient design and operation of the HARMONIA platform.

Data Commonalities: The most commonly used data come from the Copernicus programme and include Sentinel-1 SAR data, Sentinel-2 optical data and Sentinel-3 thermal data along with products from the Copernicus Climate Change and Atmospheric Monitoring Services The most commonly used non-Copernicus source is NASA MODIS data, which while having lower spatial resolution than the other cited products, has a very long-time data record and high temporal granularity, making it useful for tracking urban environments through time.

There are no particular issues highlighted in terms of access to EO data. All EO data identified for the applications is readily accessible now. In terms of future availability, all of the key datasets are secured in the medium term (5-10 years) through funded programmes that extend or enhance the current provision of data.

Gap analysis – Data: The primary gaps in data sources identified for the HARMONIA applications are pertinent to: i) fitness for purpose of some EO and modelled datasets for use at fine spatial scales in urban environments, ii) the need for high resolution EO data for validation, and iii) for in situ data of all kinds. Almost every application has uncertainties related to in situ data either for calibration / validation or for operational processes.

Nearly all the data issues are related to in situ observations – either for model calibration / validation or for operational use. There are several areas of uncertainty here and close interaction with the cities will be needed as the data availability differs from place to place. All the EO data that is required by the applications can easily be accessed on-line from a range of different sources. A strong focus is therefore needed in WP3 on trying to harmonise the in situ data needs of the applications so that they, and subsequently the DSS, can operate seamlessly across all the pilots.

Gap Analysis – Models

Most of the models needed by the applications have been identified but there are gaps in linking EO to in situ data, in downscaling results to city neighbourhood scale and in relating physical parameters to human and economic risks. The requirements for downscaling are critical to the project and are being addressed in WP4. Most of the applications are dealing more with hazard than risk and so the interface between WP5 and W6 to translate the output of the applications into decision support information will be critical.

With respect to the impact models and downscaling issues, it will remain important to work with city authorities to understand the indicators that they use to manage hazards and identify exposure and vulnerability and ultimately, derive urban risks. The activities in WP4 that lead into the Applications in WP5 may rely very heavily on good quality training data if AI / ML techniques are going to be applied. This puts a further demand on the need for high quality in situ data.

Opportunities

All of the EO data that has been proposed to fulfil the HARMONIA applications comes from already existing EO missions. However, there are many new satellite missions being developed, either as enhancements of existing constellations, or as entirely new concepts that could enhance the HARMONIA applications in the future. There is therefore a strong opportunity for the HARMONIA applications to be sustained and improved in the future.

1. Introduction

This document summarises the outputs of HARMONIA Task 2.2 which is concerned with analysis of the data and processing models and algorithms that are needed to fulfil the HARMONIA applications.

1.1. Context

During the HARMONIA project, a wide range of ambitious applications and services are being developed to help manage the risks of urban populations from changes in climate. These applications span a wide range of hazard domains including atmospheric (thermal and pollution), hydrological (flooding) and geophysical (subsidence and earthquakes). To deliver meaningful information to citizens and planners on the changing landscape risk requires a complex array of algorithms and models, transforming a wide range of different environmental and socio-economic data types into actionable information.

The HARMONIA applications are currently at a range of different maturity levels. Some parts of them have already been demonstrated in other projects, while others are more experimental. Through the four years' lifespan of HARMONIA the aim is bring the different applications to sufficient maturity levels to deliver actionable information. This requires the identification, testing and validation of all of the different data types and models that the applications need to operate, taking into consideration issues of system integration.

In this context the Task 2.2 analyses the different applications data and model requirements, identifies the areas of relative maturity where we are confident that the applications can be fulfilled, and highlights those gaps, or areas of less certainty where the research and development activities in HARMONIA need to be concentrated.

1.2. Objectives and Scope

The objectives of Task 2.2 are as follows:

1. Undertake an analysis of the data sources required for the HARMONIA applications and to specify the:
    a. Data types
    b. Data sources
    c. Access routes
    d. Licensing issues/constraints
2. Survey the analysis needs of the applications, including the processing models that are needed to go from input data to impacts indicators, including an analysis of the adequacy of the models and data to deliver results at the required spatial scale
3. Undertake a gap analysis to establish areas where the currently identified data and models require further development to meet the needs of users, focussing on:
    a. The data landscape including the data characteristics available now, and in the future
    b. The fitness for purposes of algorithms and models to deliver required information especially concerning the relevant spatial scales

In this task we cover all types of data required by the applications, including data required for calibration / validation as well as those that are needed operationally. There are many kinds of processing steps, models and information extraction required by the HARMONIA applications and throughout this document we use the term "model" to describe all of them, whether they are primary retrieval of parameters from EO data, numerical modelling, or the use of AI tools.

1.3. This deliverable in the context of other HARMONIA tasks and deliverables

There are several other tasks in HARMONIA that are concerned with analysing data types and sources. The purpose of this section is to highlight the specific objectives and outputs of Task 2.2 compared to other similar tasks. Many types of data are used across the project, but the different actors have different needs regarding them, and therefore need to take different actions to ensure their roles in the project can be fulfilled. The relevant related tasks are briefly described below, along with the distinguishing features of the current task 2.2

T1.5 (D1.1) Data management plan & data ethics is required to produce a plan to demonstrate that the project data are well managed for consistency with the FAIR principles. While this task and deliverable has to catalogue the data that are to be stored and generated by the project, it is more concerned with the data management principles than with the technicalities of the data quality, performance and formats.

T1.6 (D1.4) Data and services Taxonomy, legal and ethical use of EO data. This task is concerned with the legal and ethical impacts of the use of data and applications in HARMONIA.

T2.1 (D2.2) Concerns the collection of user requirements from the urban stakeholders involved in the project. In the end, these user requirements will drive the development of the applications and services in the project. However, user requirements are to be expressed in output information and functional terms. They are in principle data agnostic as they specify what the user wants to achieve, not the methods and data sources by which these achievements should be made.

This Task 2.2 (D 2.3) is concerned with both data **and models / algorithms**. The main focus is the **fitness for purposes** of these two elements, in combination. It is concerned not so much to make a long list or catalogue of all the technical characteristics and formats of data, but rather to identify the maturity of the data and models to serve user needs and requirements. It also seeks proactively to search for and to highlight gaps in availability and therefore to help direct research and activity efforts to fill and mitigate those gaps within the project. The task also has a future focus – identifying data sources that can sustain, and also improve, the HAROINIA applications in the future.

Finally, WP3: Data acquisition, integration and handling is mainly concerned with the mechanics of data access and storage, ensuring the relevant data can be efficiently accessed stored and provided to other tasks in the project.

2. Methodology

The methodology adopted for this task revolves around the identification of combined data and model function requirements suitable for the HARMONIA applications. As data and processing models are used widely across the project, we need a focus point from which to build the data and model landscape.

In parallel to the execution of this task within WP2, another task is the collecting of user requirements. However, in system terms, user requirements should be data and method agnostic. They express what a user wants to achieve – not exactly how that achievement should be fulfilled, so they cannot be used as a direct source of information for data and models. We have therefore elected to use the specifications of the HARMONIA applications themselves in WP5 as the focus for information collection in this task.

In HARMONIA, a number of different tasks build towards the ultimate demonstration and validation of services aimed at fulfilling user's needs, as shown in Figure 1.

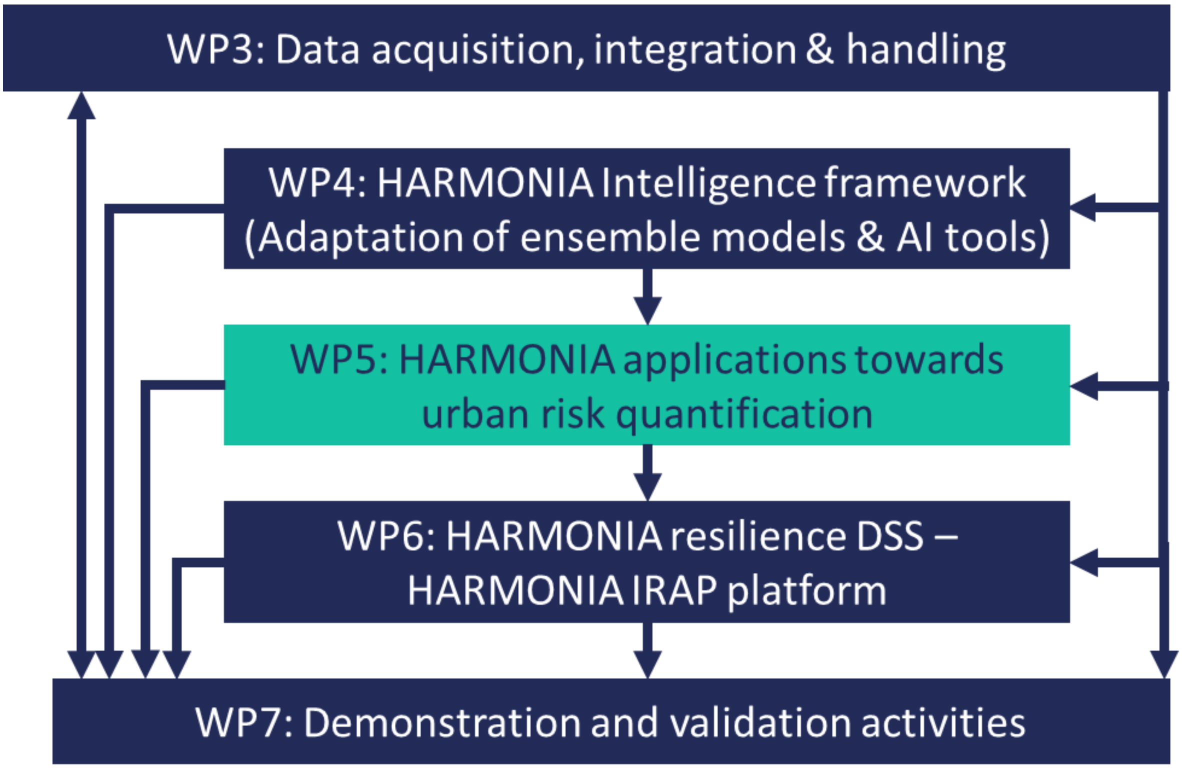

Figure 1 – HARMONIA data, applications and services logic flow, showing the interaction between the work-packages and WP5 as the focus point for the specification of applications data and model requirements

At the heart of this process is WP5 where the HARMONIA Applications are elaborated. WP5 services will rely on R&D undertaken in WP4 to develop and refine the processing models needed in the Applications. Subsequently the services developed in WP5 will be integrated into the IRAP decision support systems, and ultimately demonstrated to users in WP7. The five thematic Applications in WP5 are therefore the key point in which data and model requirements can be elaborated.

Each of the five applications, data and model requirements have been structured based on the segmentation shown in Figure 2. In this framework the data and models have been divided into a number of types.

1. **EO data** meaning all types of remotely sensed data from space or airborne platforms
2. **In situ data** meaning all types of data collected from the ground, including refence data, socio-economic data and data from ground sensors
3. **Modelled data** meaning data that has been pre-computed or modelled by another organisation external to HARMONIA – this typically means the outputs of reanalysis systems such as ERA5 or atmospheric modelling such as the Copernicus Atmosphere Monitoring Service (CAMS).

The models and processing algorithms have been divided into two types:

- **Retrieval models** which act primarily on raw data to derive physical parameters;
- **Biophysical or impact models** which are primarily concerned with calculation of parameters relevant to end users).

In reality, some data and models may fall between these categories, but this division is used as a convenient way of structuring the landscape. The purpose of the framework diagrams is not to define the architecture or flow of data, but simply to categorise the different data and model types used in each application.

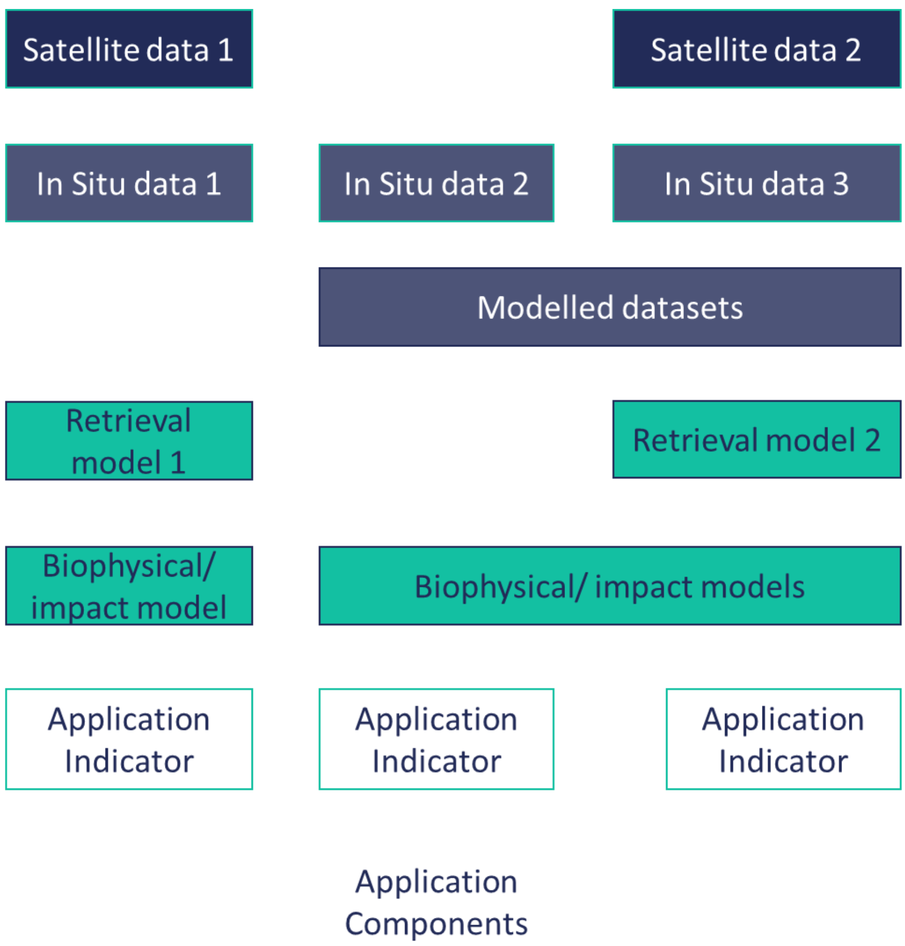

Figure 2 – Example of an application information collection framework demonstrating the categorisation of data and model types needed to fulfil the HARMONIA applications.

For each component of the five HARMONIA Applications the data and models needed to fulfil it have been specified by the WP5 task leader responsible for the application using the above framework. For each Application there is a one-page summary to describe the overall flow, from application indicators to data, giving the required processing steps. This is repeated for each of the sub-components of the Application. The Applications differ significantly in scope and structure, so not every component will necessarily be present in each case. Given the diversity of the Applications, the following should be borne in mind:

- Some cells in the diagram might be empty because a component is not needed – for example if some modelled dataset is available already and can be used in its native form.
- There may be some cases where a specific model or dataset has not yet been identified. In this case, to avoid ambiguity between an element that is not needed, and one which is currently unknown, when an element is blank this indicates it is not needed. However when an element is needed, but an exact solution has not been decided upon, then a generic description is included and the item is marked in red to indicate a gap.
- Where more than one kind of data or model is needed, all of them are included, but if there several *alternatives* that do the same job, only the primary one is included

Based on this overview structure, each data and model element has then been defined in more detail in a specification table which includes for **data**:

- Summary of the dataset
- Technical characteristics of the data from the perspective of its fitness for purpose in the proposed application
- Availability of the dataset in terms of the spatial and temporal coverage
- Technical Access mechanisms or the dataset – how it can physically be accessed
- QA – quality assurance status of the dataset
- Legal issues – any licensing or usage constraints

For **models** a similar table is complied with the following specifications:

- Summary of the model or algorithm, with its main inputs and outputs
- Key References
- Brief description of the technical fitness for purpose with respect to is proposed usage in the application
- Development Level of the model – if it is in basic algorithm form or has been elaborated into a software package
- Model Validation – specification of the extent to which the model and its outputs have been validated
- Software QA status of the model (if it is in software form) indicating the level of reliability and maintenance
- Legal issues – any licensing or usage constraints

For each of the categories, the Application leader has assigned a "traffic light" flag to indicate the status of the item with respect to its suitability and reliability for use in the Application. The status lights are assigned as follows:

Green: Everything is OK, the model/data is available for use in HARMONIA and meets the needs of the application, scientifically, technically or legal as applicable. Some minor modification to software or pre-processing of data might be needed.

Amber: The element is available and appears to be fit for purpose. There may be some uncertainties over performance, and moderate effort may be needed in further validation, processing or development.

Red: Default if the element is completely missing. Otherwise, there is doubt about the feasibility of the model or data because it is not clear if the performance is adequate, or a significant amount of development, validation.

In order to help the readability of this document, all of these detailed data and model specification tables have been included in an Appendix in section 7, with a brief summary of them included in the body of the document.

Based on this data collection and structuring, a cross-cutting analysis has been undertaken to analyse the commonalities and gaps in data and model requirements for the project, in order to support subsequent tasks in the efficient access to data, and to help focus the development activities on the project towards the critical missing or uncertain items needed to fulfil the Applications.

3. Application data and model requirements

Each subsection in this chapter describes one HARMONIA Application with their corresponding data and models required. For each application a summary table is presented with:

- Data types (e.g. EO data, in situ, socio-economic, citizen observatories)
- Data sources (existing open services such as GEOSS, Copernicus services, ESA TEPs; local/regional/national statistical and geospatial data; one-off campaigns, commercial; research);
- Access routes (e.g. online open access, proprietary, commercial)
- Licensing issues/constraints.

The complete tables of data and models requirements can be found in section 7.

3.1. Urban air quality services

Key variables influencing urban air quality in the long term are i) green areas, ii) carbon-powered traffic volume and iii) weather (wind, humidity and rain). The HARMONIA urban air quality service will combine these important aspects of air quality research to estimate the air quality impacts of urban development activities. This service will have four components as shown in Figure 3.

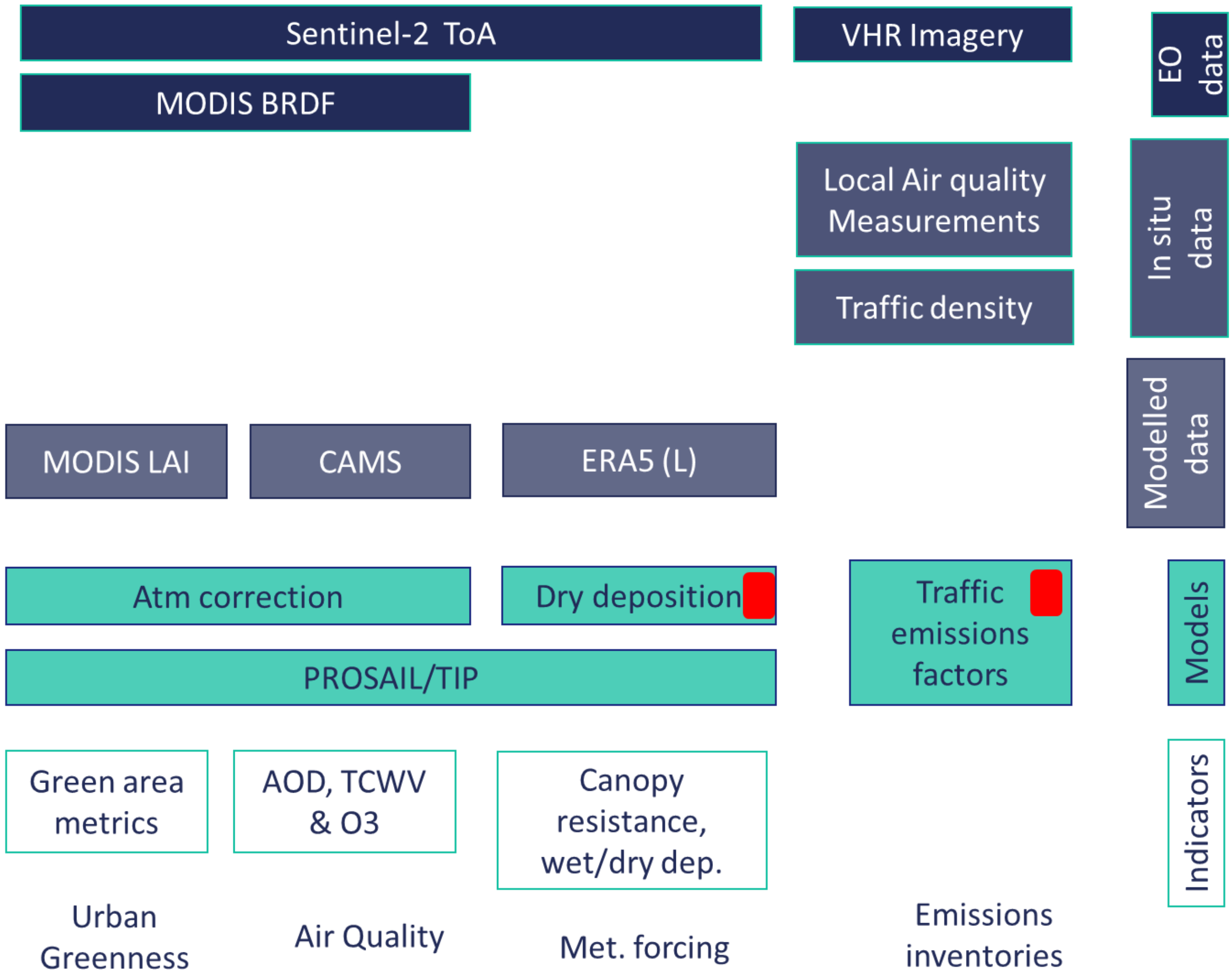

Figure 3 – Data and model framework for the urban air quality application showing the elements needed to fulfil the four components of the application. The red flags indicate gaps or uncertainties in the dry deposition and traffic emissions factors models,

Air quality indices – Atmospheric pollution will be investigated using Sentinel-2 TOA and CAMS atmospheric composition datasets to derive AOD, TCWV and O3 at high resolution (10m). The Sensor Invariant Atmospheric Correction (SIAC) atmospheric correction (Feng et al., 2021) scheme will be used to generate the aforementioned atmospheric composition elements. As a by-product of this processing the surface reflectance is produced and will be used on other components of the air quality service.

The effects of rain for cleaning air pollution through deposition and wind for diluting pollution into less harmful concentrations are well known. To be able to use seasonal and climate predictions a service will provide the outlooks for changes in precipitation and wind to give an index for improved or worsened air quality conditions.

Urban greenness – Using the Sentinel-2 surface reflectance generated using SIAC, vegetation biophysical parameters (LAI, brown matter, carotenoids concentration, etc.) will be retrieved using a radiative transfer model such as PROSAIL. In particular, the Leaf Area Index (LAI) parameter will be used to derive indicators such as the percentage of area with LAI over certain threshold, the size of patches categorised by LAI and an analysis of distribution and/or clumping. These parameters will be related to known dry deposition potential for current air quality information available from CAMS and local measurements.

Emissions inventory – Through assessment of satellite imagery we will analyse vehicle and pollution activity in the target cities and compare this to both CAMS and local air quality data sources. Very high resolution data such as Planet, Pleiades and SPOT will be used to attempt to relate limited number of traffic measurements to traffic conditions across the city.

Meteorological Forcing – Since urban vegetation affects air quality through influencing pollutant deposition and dispersion (Janhäll, 2014) having baseline datasets that will help the understanding of the parameters influencing the effect of vegetation on air pollution is necessary. Using a synergistic approach of EO at different spatial resolutions and CAMS reanalysis data will be used to link vegetation biophysical parameters and air pollutants over the last 20+ years, through the relationship of dry deposition and LAI, deposition velocity and the air concentration of the pollutant.

The data that have been identified for fulfilling this application are summarised in Table 1 and the models in Table 2. More details of both can be found in section 7.1 (Appendix A).

Data Requirements Summary					
Name	Type	Source	Access	License	Flag
Sentinel-2 TOA	EO	Copernicus	Online open access	Open access	
MODIS BRDF	EO – modelled	NASA-LP DAAC	Online open access	Open access	
MODIS LAI	EO – modelled	NASA-LP DAAC	Online open access	Open access	
CAMS	EO - modelled	Copernicus	Online open access	Open access	
VHR EO data	EO	Commercial	Online proprietary access	Commercial – per vendor licence	

In situ (AQ, traffic measurements in some cities)	In situ	Regional statistics	Government/Research	TBD	

Table 1 -Data requirements summary for urban air quality application

Model Requirements Summary					
Name	Function	Input Data	Application (output) Data	License	Flag
SIAC	Sentinel-2 atmospheric correction	Sentinel-2 TOA	Sentinel-2 BOA and atmospheric composition	Open	
ProSAIL and TIP (RTMs)	Retrieval of vegetation biophysical parameters	Sentinel-2 BOA	Vegetation biophysical parameters (LAI, FAPAR)	Open	
Dry deposition model	Dry deposition modelling	Atmospheric composition, pollutants concentration, LAI.	Dry deposition rate per unit area	TBD	
Traffic emissions factor model (for some cities)	Relationship between quantity of pollutants and associated activities	TBD	TBD	TBD	

Table 2 -Model requirements summary for urban air quality application

3.2. Urban Heat Fluxes & urban heat islands

Key variables influencing the urban heat fluxes are the surface cover and morphology, as well as the current meteorological conditions. The HARMONIA urban heat fluxes estimation is using the well established Aerodynamic Resistance Method (ARM) and the Objective Hysteresis Model (OHM) for the estimation of sensible heat flux and heat storage flux respectively, as shown in Figure 4.

The turbulent sensible heat flux refers to the heat exchange between the urban surface and the atmosphere and defines the amount of energy that is available for heating the urban atmosphere and it is thus closely related to air temperature. This amount of energy is strongly modified by the properties of the surface (land cover/ land use, 3D geometry) and the input of heat by human activities (traffic, buildings, industry).

The heat storage flux refers to the change of energy stored in the urban canopy (especially in buildings), which is approximately 2-6 times larger than in non-urban canopies. The slow release of this energy, stored mainly in the buildings and paved surfaces of cities, causes the UHI effect and it is therefore related to the energy efficiency and consumption in cities. Therefore, the urban land cover, geometry, radiation and air/surface temperature are affecting the heat storage flux.

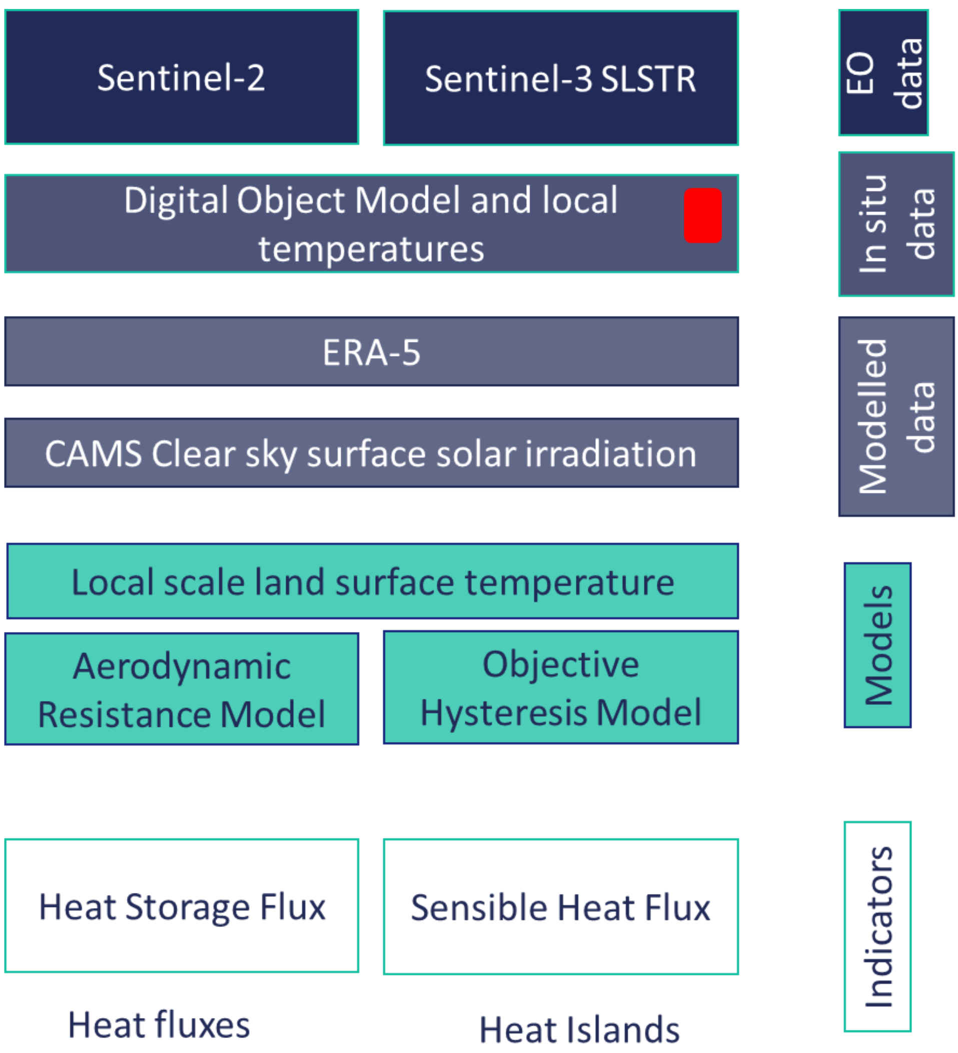

Figure 4 – Data and model framework for the urban heat fluxes and urban heat islands application. The red flag indicates some gaps and uncertainties in the provision of local data on city morphology and temperatures.

The data that have been identified for fulfilling this application are summarised in Table 3 and the models in Table 4. More details of both of these can be found in section 7.2 (Appendix A).

Data Requirements Summary					
Name	Type	Source	Access	License	Flag
ERA-5	Modelled	Copernicus	Open	Open	
CAMS	Modelled	Copernicus	Open	Open	
Sentinel-2	EO	Copernicus	Open	Open	
Sentinel-3 SLSTR Thermal	EO	Copernicus	Open	Open	
Local in situ surface model and temperature data	In situ	TBD	TBD	TBD	

Table 3 -Data requirements summary for the urban heat flux and urban heat islands application

Model Requirements Summary					
Name	Function	Input Data	Application (output) Data	License	Flag
Local scale land surface temperature model	Downscales	Sentinel-3 SLSTR brightness temperatures	100M resolution LST	Model published in literature. Software IPR belongs to FORTH	
Aerodynamic Resistance Method (ARM)	Estimates momentum and scalar fluxes in the atmospheric surface	Digital Object Model, various met parameters	Urban sensible and latent heat fluxes	Model published in literature. Software IPR belongs to FORTH	
Objective Hysteresis Model	Established model using net radiation to calculate storage heat flux in urban areas based on land cover fractions.	Landcover, net radiation	Storage heat flux	Model published in literature. Software IPR belongs to FORTH	

Table 4 -Model requirements summary for the urban heat flux and urban heat islands application

3.3. Humidity and Temperature Anomalies & Complex Scenarios

One of the most important key variables that affect wildfire occurrence, intensity and spread rate are the fuel moisture content, air temperature and relative humidity. Fires can produce a change in albedo, hence a change in the Earth's energy balance system that can be measured in terms of radiative forcing, which may lead to hot spots of climate change where surface temperature can rise and affect natural ecosystems surrounding urban environments. A reliable indicator of changes in fire regimes, understood as the pattern, frequency, and intensity of the wildfires that prevail in an area over long periods of time, is known as fire danger. Hence this application will focus on two core components as shown in Figure 5.

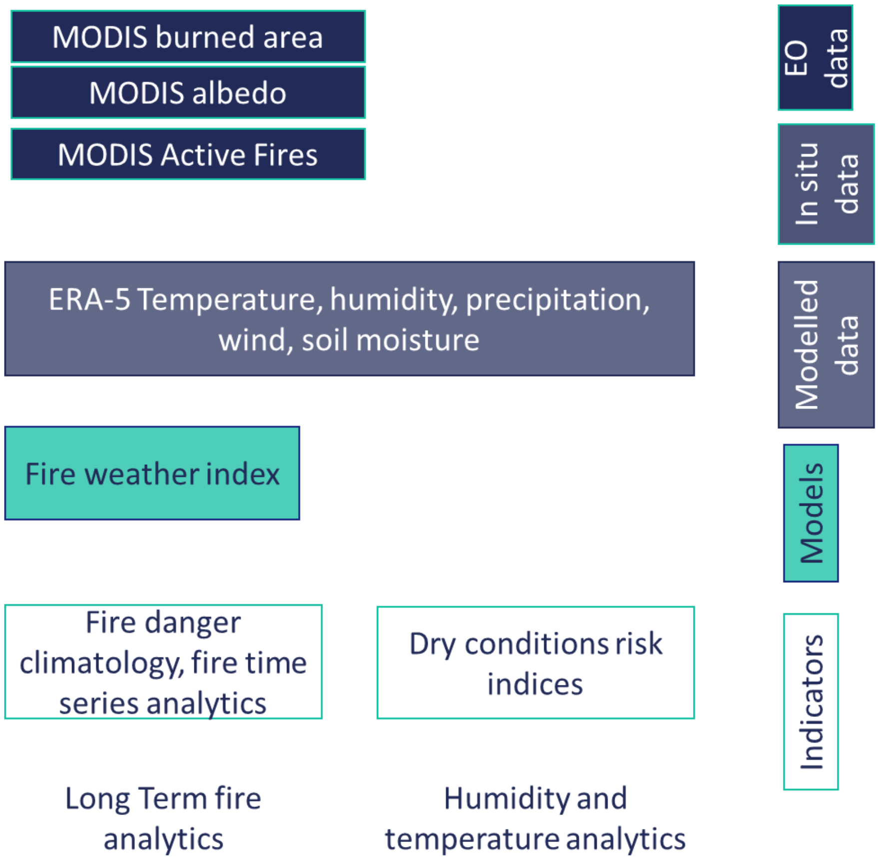

Figure 5 – Data and model framework for the humidity, temperature anomalies and complex scenarios application

Humidity and temperature analytics – This application will perform a humidity and temperature analysis to identify and describe risks related to dry conditions, particularly in urban-forest interfaces where wildfires can affect not only natural vegetation but because of the proximity to urban areas, represent a high risk to human life. Seasonal forecasts together with ERA5 will be used to perform bias correction methods to identify anomalies and describe related risks. Ancillary information from C3S, like the soil wetness, will be used for risk quantification.

Fire danger long term analytics – Using satellite data it is possible identify fire occurrences, however they are difficult to predict since most of the ignitions are related to human activity. Fire danger models do not predict where and when an ignition might occur but can identify the meteorological conditions that eventually will allow a fire to spread out of control. Using a long-term data record of fire danger, it will be possible to assess how 'fire weather' variables (accumulated precipitation, relative humidity, temperature and wind speed) have changed over time.

In this task application, a fire danger long term data record will be created using ERA5-Land 9km data based on the methodology developed by Vitolo et al. (2020) that generated a global fire danger index using ERA5 (about 28 km grid cell size). Using ERA5-Land will allow for the better characterisation of local conditions. Some fire datasets (active fires, burned area and albedo from MODIS) will be included in the analysis. The fire danger index and fire variables and associated products will include: Fire danger climatology from 1981 to present; Identification of extreme conditions; Wildfire time series analytics (fire occurrence, fire radiative power and burned area). Analysis of radiative forcing from fire-induced shortwave albedo change will be produced following the methodology developed by Lopez et al. (2015).

The data that have been identified for fulfilling this application are summarised in Table 5 and the models in Table 6. More details of both of these can be found in section 7.3 (Appendix A).

Data Requirements Summary					
Name	Type	Source	Access	License	Flag
MODIS Burned area	EO – modelled	NASA-LP DAAC	Online open access	Open access	
MODIS Active Fires	EO – modelled	NASA-LP DAAC	Online open access	Open access	
MODIS BRDF	EO – modelled	NASA-LP DAAC	Online open access	Open access	
ERA5	RA	Copernicus	Online open access	Open access	

Table 5 -Data requirements summary for the humidity, temperature anomalies and complex scenarios application

Model Requirements Summary					
Name	Function	Input Data	Application (output) Data	License	Flag
FWI	Model fuel moisture response to atmospheric forcing	ERA5	Fire danger	Open	

Table 6 -Model requirements summary for the humidity, temperature anomalies and complex scenarios application

3.4. Urban Flash Flooding

Meteorological reanalysis, in situ observation data sets as well as climate model prediction data will be used to derive a statistical urban flash flood impact model to estimate the risk of extreme precipitation events in three urban environments: Ixelles (Belgium), Milan (Italy) and Sofia (Bulgaria). Based on thresholds for strong and extreme precipitation events and the probability of these conditions to occur, the service will provide information to the end users to prepare for the impacts of climate change as well as provide an operational forecasting service for estimating/visualizing short-term impacts. Precipitation forcing data sets will be downscaled from meteorologic nowcasts, short-term forecasts, climate scenarios and seasonal forecasts to provide a short-term forecast and an up to 7-month seasonal outlook for operational duties and long-term scenarios to prepare for future worst-case scenarios. Downscaled precipitation fields will be converted to synthetic precipitation radar images (with ecPoint-Calibrate) to be used as input for an urban hydrologic/hydraulic model (Itzï) to create urban flood impact nowcasts/forecasts. Both the precipitation and impact nowcasts/forecasts will be incorporated into an existing FMI/Cityzer platform to serve as a data feed and visualization tool.

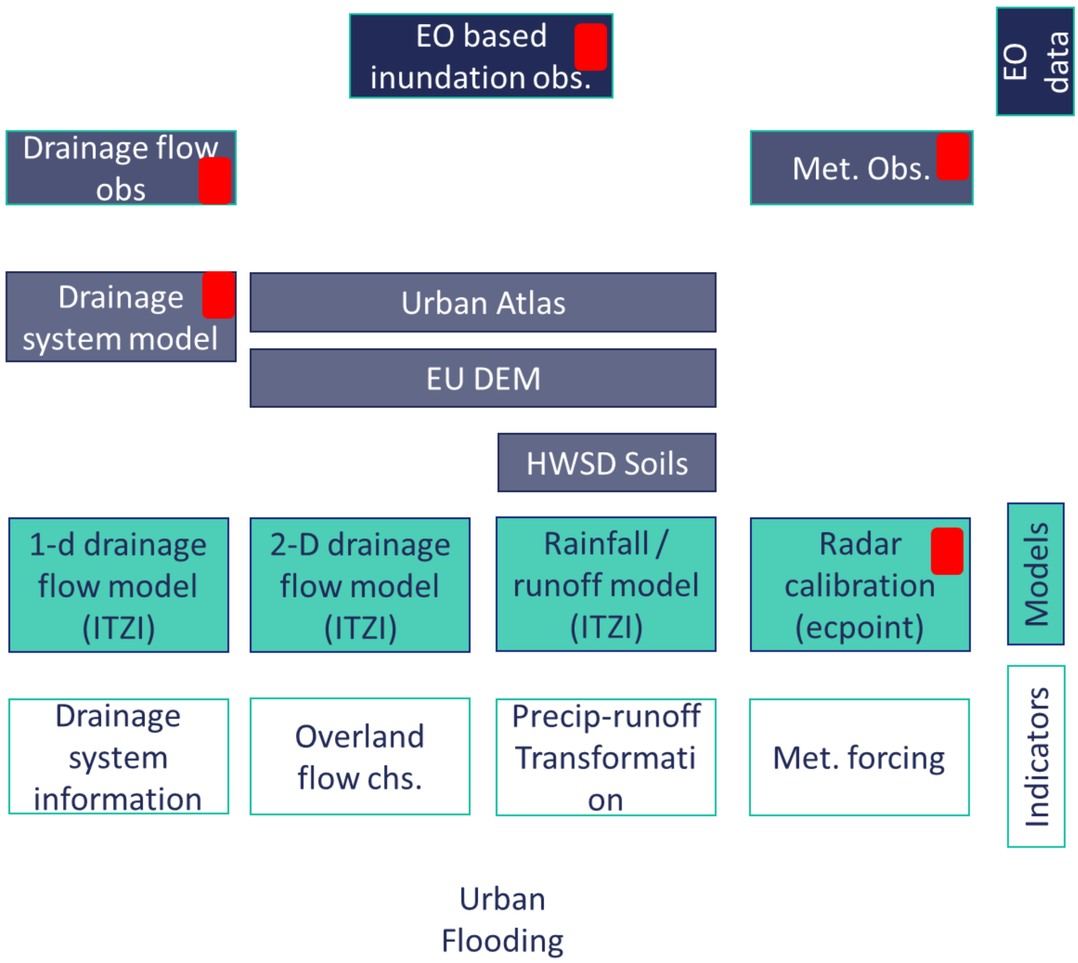

Figure 6 – Data and model framework for the urban flooding scenarios application. The red flags indicate potential gaps in availability of local system drainage model, and uncertainty over the applicability of the ecpoint calibration model.

In the high-level service elements table, there are four static elements; "EU DEM", "Urban Atlas", "HWSD Soils" and "Drainage System GIS Model", one forcing data generator element; "Weather Radar Data Model" and three dynamic validation elements; “Drainage Flow Obs.”, “Met. Obs.”, “EO-Based Flood Inundation Obs.”. The first three static elements (DEM, Urban Atlas and Soils) will be used to calculate Itzï model rainfall-runoff model parameter values. The “Drainage System GIS” data will be used to build a conceptual model of a stormwater drainage system and to estimate drainage system flow parameter values for the Itzï - distributed hydrologic and hydraulic modelling system. The urban atlas element together with the DEM will also be used to derive a 2-dimnesional surface flow routing model, defining flow directions and flow friction parametrization. In essence the static system elements are used to define the hydrologic model and its parameter values.

The three dynamic validation elements, albeit not necessary for successful project implementation, are important in assessing the predictive power of the modelling system and as such are an integral part of the project. The “Drainage Flow Obs.” element provides information on subsurface flow in a city’s drainage system. This data is used to both validate and calibrate the Itzï modelling system’s 1D drainage flow SWMM model. Any simulated past or future flood outcomes could be validated with “EO-Based Flood Inundation Observations”. A prerequisite for developing this element is knowledge of past flooding events and the availability of corresponding satellite imagery. At the time of writing, it is envisioned that the ecPoint-calibrate element will be used to downscale precipitation fields from lower resolution meteorological datasets to develop experimental synthetic weather radar images, which are used as dynamic forcing data input for the hydrological model. Using this method/software is however still somewhat questionable and other algorithms/models may either be employed or developed in-house at FMI during the project. Any data produced with this element is validated/compared against in situ meteorological observations (“Met. Obs.”).

The data that have been identified for fulfilling this application are summarised in Table 7 and the models in Table 8. More details of both of these can be found in section 7.4 (Appendix A).

Data Requirements Summary					
Name	Type	Source	Access	License	Flag
EO Inundation observations	EO	TBD	TBD	TBD	
Drainage flow observations	In Situ	TBD	TBD	TBD	
Meteorological observations	In Situ	TBD	TBD	TBD	
Drainage system GIS model	Modelled	TBD	TBD	TBD	
EU DEM	Modelled	Copernicus		Open	
Urban Atlas	Modelled	Copernicus		Open	
HWSD Soils	Modelled	FAO		Only for non-commercial use	

Table 7 -Data requirements summary for the urban flash flooding application

Model Requirements Summary					
Name	Function	Input Data	Application (output) Data	License	Flag
ITZI	Urban hydraulic and hydrological modelling	Meteorological forcing	Surface and drainage flows	Open	
ePoint Calibrate	Tool to calibrate NWP outputs against point observations	NWP outputs and point observations	Downscaled precipitation estimations	Open	

Table 8 -Model requirements summary for the urban flash flooding application

3.5. Geohazards

This section describes the data and algorithms relative to the geohazard applications carried out using satellite SAR Interferometry by INGV. During the last 25 years the Interferometric Synthetic Aperture Radar (InSAR) image processing technique has become a widely used method for the detection of small ground deformation signals. The most important geophysical observable, measured by InSAR, is the displacement of a ground scatterer occurring within the time interval defined by two or more image acquisitions. The InSAR technique can measure the projection of the deformation vector onto the Line of Sight (LoS) direction, defined as the shortest path from a given point on the ground to the SAR antenna phase centre. InSAR can provide surface deformation maps with high spatial resolutions (5–80 m) over large areas.

The basic InSAR measurement is an interferogram, and represents the per-pixel phase difference between two SAR acquisitions. The interferogram contains both topographic and surface motion information; surface motion can be isolated by removing the topographic component either using an external DEM or using a second SAR image pair.

The recovery of the integer multiples of 2p, and thus the determination of the phase gradient between any two interferogram pixels, represents a 2D phase unwrapping problem, for which only approximate solutions can be found by automated algorithms. The measured ground displacements are affected by unmodelled effects, the most common being the water vapor content in the troposphere, the Total Electron Content in the ionosphere, and uncertainties in the satellite position.

In HARMONIA we will use multi-temporal InSAR techniques, which exploit the redundancy offered by tens or hundreds of image pairs acquired over a given temporal frame, to reduce the aforementioned limitations. The output of these techniques is a ground displacement time-series, temporally referred to the first acquisition date of the image stack, and spatially referred to a reference point or area within the radar coverage. The existing algorithms fall into two broad categories, namely the Persistent Scatterer (PS) and the Small Baseline (SB) approaches, although algorithms exploiting both methodologies have also been proposed. The aim of the PS methods is to identify coherent radar targets exhibiting a high phase stability over the whole temporal span of the observations. These targets are only slightly affected by temporal and geometrical decorrelation, and often correspond to man-made structures or bare rock. In contrast, in the SB approach, interferometric pairs are chosen to minimize temporal and geometric decorrelation, allowing displacement time-series to be retrieved for distributed scatterers, i.e., neighbouring radar resolution cells, which are not dominated by a single scatterer, and share the same backscattering properties.

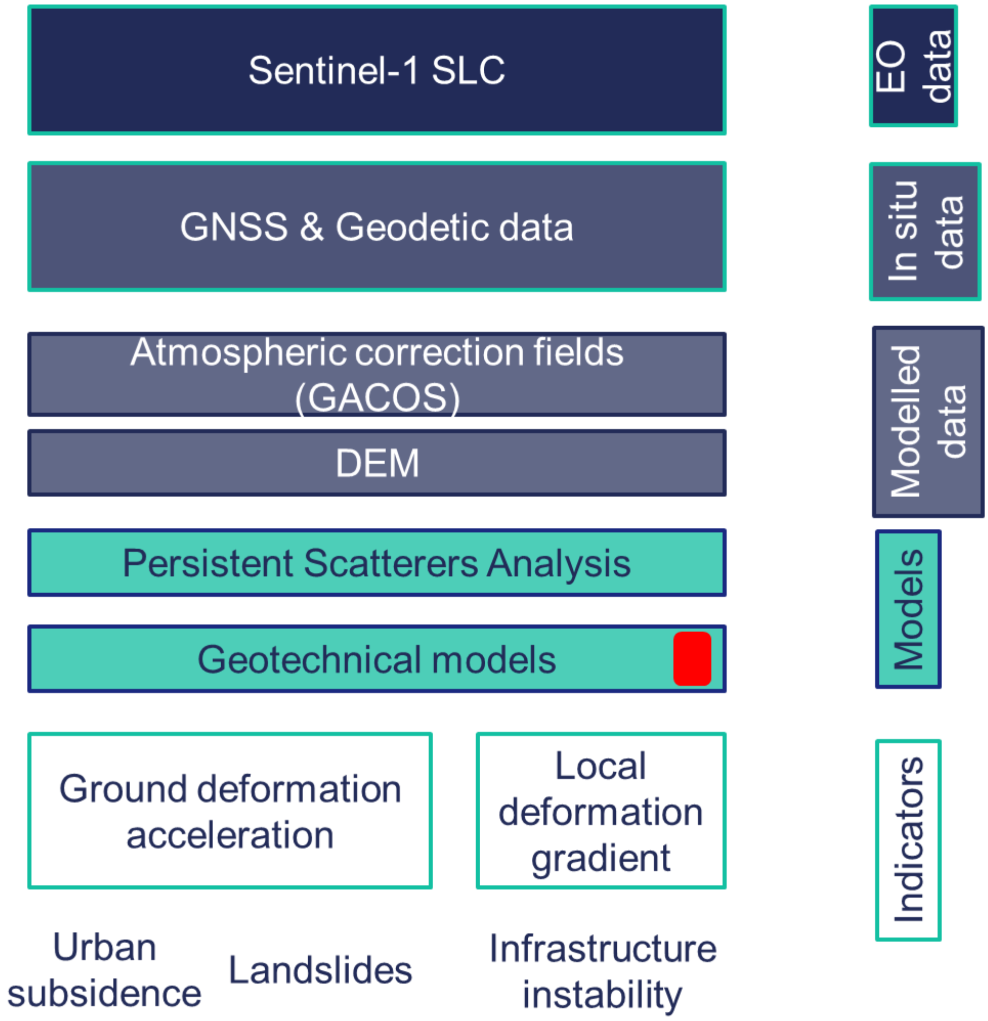

Figure 7 – Data and model framework for the geohazards application. The red flag indicates that further work is needed to develop geotechnical models to interpret ground movement measured from SAR data.

Considering the specific applications of HARMONIA (geohazard monitoring and assessment in urban areas) we will use the PS technique, which allows a higher ground resolution than the SB method, allowing, in favourable conditions, to investigate differential motion and internal deformation of buildings and infrastructures as well as extended areas.

With large datasets and good measurement conditions, the PS method can measure long term velocities of ground objects with an accuracy of 1-2 mm/yr. We will use Sentinel 1 satellite SAR data for the PS applications, the SRTM DEM for topographic corrections, and pre-calculated GACOS tropospheric delay maps for the correction of atmospheric artefacts.

The data that have been identified for fulfilling this application are summarised in Table 9 and the models in Table 10. More details of both of these can be found in section 7.5 (Appendix A).

Data Requirements Summary					
Name	Type	Source	Access	License	Flag
Sentinel-1 SLC	EO	Copernicus	Multiple API sources	Open	
SRTM DEM	EO	USGS	USGS Earth Explorer	Open	
Generic Atmospheric Correction Online Service for InSAR (GACOS)	Modelled	GACOS, University of Newcastle	Website / API	Open, with attribution	

Table 9 -Data requirements summary for the geohazards application

Model Requirements Summary					
Name	Function	Input Data	Application (output) Data	License	Flag
Persistent scatterers analysis	Uses InSAR analysis to measure ground movement from persistent scatterers (PS)	SAR Images	Mean velocity of each PS, time series of deformation	Proprietary (ENVI)	
Geotechnical analysis	Various methods to interpret the observed ground displacements from the processing of SAR	Output of PS analysis	Interpretation of the causes and consequences of ground displacement	TBD	

Table 10 -Model requirements summary for the geohazards application

4. Cross-cutting analysis

4.1. Data requirements overview

A summary of the data requirements from the HARMONIA applications is shown on Table 11. The colour indicates the availability of the data. Green is related to open and freely available data whilst amber is only commercially availability or data where there are limitations associated to its acquisition. Overall, the applications will be built upon EO and RA datasets whit full open access. The main constrain in here can be found in the in situ observations and VHR EO datasets that can be obtained only from a commercial entity.

Data Requirements Summary											
Application	S1	S2	S3	MODIS	VHR EO	DEM	ERA5	CAMS	Cartography	Modelled/EO+RA	In situ
Urban Air Quality Services											
Urban Heat Fluxes & urban heat islands											
Humidity and Temperature Anomalies											
Urban Flash Flooding											
Geohazards											

Table 11 -Summary of the data requirements for the HARMONIA applications

Strictly speaking, “in situ” describes data collected adjacent to the measuring instrument, like temperature readings by a thermometer. In the HARMONIA scope, the concept is much wider, covering all data not collected by EO satellites. It includes physical observations, such as air temperature or air quality, reference data (e.g. maps), and economic, environmental and social information such as population characteristics. The value of EO data, in a sense, depends on their effective integration with other technical resources and observations such as in situ ones. HARMONIA WP3 has allocated vast resources to integrate EO data with a wide range of data sources such as citizen-bases and urban context. These efforts must include ensuring the integration of essential in situ observations for the applications. For instance, the Urban Flash Flooding depends heavily on each municipality requesting an urban flood modelling service to provide or point to a place where city wide drainage system flow can be obtained. Another example is the Urban Air Quality Services, in situ measurements of air pollutants are required to validate the results of the atmospheric composition models.

In terms of commercial datasets, some data providers such as Planet have ongoing education and research programmes where any university-affiliated researcher can apply for non-commercial access to the data, this can be a valuable way for HARMONIA researchers to access these kind of datasets and assess the feasibility of using them for urban studies and potentially integrated them into future operational applications.

4.2. Access mechanisms

A summary of the data access mechanisms from the HARMONIA applications is shown on Table 12. The colour indicates the online availability. Green indicates open access whilst amber, restricted access either for commercial purposes or proprietary in situ data. For EO data HARMONIA datasets will be stored and managed in the HARMONIA Data Cubes. In this context, A Data Cube is a multi-dimensional array with two spatial dimensions, the x-y coordinates, and additional dimensions that describe either the temporal dimension of the data and/or other spectral properties.

Data Access Mechanisms											
Mechanism	S1	S2	S3	MODIS	VHR EO	DEM	ERA5	CAMS	Cartography	Modelled/EO+RA	In situ
Online – open access											
Online – commercial access											
TBD											

Table 12 -Summary of the data access mechanisms for the HARMONIA applications

The EO datasets identified in this task and, and task 2.2 will be acquired and stored into the HARMONIA Data Cubes. In some cases, particularly for the Copernicus EO data and services, the potential to use a DIAS platform will be ideal since all data is already accessible there. For instance, CREODIAS contains online most of Copernicus Sentinel satellites data and Services, Envisat and Landsat and the participation of Cretech as leader of HARMONIA WP3 should facilitate the availability of Sentinel data. The challenge here will be the integration of all non-Copernicus data, e.g. MODIS and VIIRS data seamlessly in a DIAS environment.

All EO data sources and cartography required provided excellent APIs, which will make easier any machine-to-machine data requests and will facilitate a transition to operational status for potential future services. Meanwhile, the in situ data will rely on bespoke acquisition methods that will have to be addressed in WP3.

4.3. Data Landscape

Almost all applications rely on the use of Sentinel data. Given the operational nature of the ESA Sentinel programme, the nominal mission lifespan of all platforms and the planned replacements once the platforms reach the end of their operational lifespans, the HARMONIA applications can be built upon Sentinel data.

In terms of weather and atmospheric composition, the HARMONIA applications heavily use data from ERA5 and CAMS. One of the main characteristics of ERA5 data is the availability within 5 days of real time which will serve users of HARMONIA applications that need recent meteorological information that will be combined with a long and consistent climate record. ECMWF's vision for C3S post-2020 continues to allocate a very high priority to reanalysis. A centennial global reanalysis back to 1900 or earlier is proposed and was due to start in 2021. The next full-observing system reanalysis, ERA6, will be based on a coupled Earth system modelling and data assimilation approach and will assimilate the new datasets mentioned above (Hersbach et al., 2018), it will be started by 2023.

Four out of five HARMONIA applications will require a DEM. Ensuring the DEM is representative of the areas of interest where it will be used is key to ensure consistency amongst all applications. Two main potential sources are the NASA Shuttle Radar Topography Mission (SRTM) and the ASTER Global Digital Elevation Model. Both datasets provide elevation every 1 arc-second (approximately 30 meters).

The applications where characterising land surface processes is relevant will use data derived from the MODIS sensor on board the Terra & Aqua platforms. Whilst MODIS has been a very successful programme, the life expectancy of the Terra platform has far exceeded its design life and has a strong chance of having issues e.g. Terra began drifting to an earlier Equatorial Crossing Time in February 2020 and that could affect the geolocation and radiometry of the instrument. Some mitigation measures have been taken such as drag make up manoeuvres, meaning Terra's instruments will continue to be operational during this entire process. The Aqua platform has not reported any issues but it is reaching its 20th year in orbit.

In order to capture changes in climate that are affecting the areas of interest selected in HARMONIA, a long-term data record is required. The temporal availability of the EO & reanalysis datasets will be discussed in terms of the length of the data record and the potential to keep using the datasets in the future.

Reanalysis

All reanalysis datasets provide more than 30 years of data. Current ERA5 data can be obtained from 1979 to present. Eventually, ERA5 will cover the period from 1950 to the present. A pre-release quality control however revealed some problems affecting the performance in the tropics. The strategy to deal with this issue has delayed the publication of the dataset. Data should gradually become available in 2022. Additionally, ERA5-Land is a global land-surface dataset at 9 km resolution that would be much better for urban applications. It is consistent with atmospheric data from the ERA5 reanalysis but currently it has a lag of approximately 3 months. Daily updates 5 days behind real time are expected to be made available in 2022.

Coarse resolution EO

MODIS data provided one of the longest consistent coarse resolution data records with 20+ years of land surface data that can be used within the HARMONIA applications. Besides MODIS, no other coarse resolution sensor has consistent operational datasets that span more than 20 years (2000 – present day) that can deliver consistent, validated and easily available land surface parameters. Several of the HARMONIA applications use MODIS-derived datasets such as albedo and LAI. Part of the success of the MODIS programme is the continuous improvement of the retrieval algorithms, this is achieved by reprocessing the whole MODIS archive from Level-0 (raw data) to Level-1B (calibrated radiances) all the way to high-level processing (Level-2, 3 and 4). For instance, latest version 6.1 Level-1B products have been improved by undergoing various calibration changes that include: changes to the response-versus-scan angle (RVS) approach that affects reflectance bands for Aqua and Terra MODIS, corrections to adjust for the optical crosstalk in Terra MODIS infrared (IR) bands, and corrections to the Terra MODIS forward look-up table (LUT) update for the period 2012 - 2017. This has an impact on all high-level products; hence it will be compulsory for the HARMONIA applications to use only MODIS products version 6.1.

All HARMONIA applications should start considering the use of different sources of coarse resolution optical and thermal data from instruments such as VIIRS, OLCI and SLSTR. These instruments provide similar capabilities in terms of radiometric and spatial resolution as MODIS and particularly in the case of VIIRS, several of the science algorithms to derive Level-2 and Level-3 MODIS products have been adapted to work with VIIRS data, for instance the surface reflectance, albedo and BRDF model parameters, and active fires.

Sentinel Series and related missions

Sentinel-1, Sentinel-2 and Sentinel-3 provide excellent data on an operation basis. However, the data is not necessarily useful for climate analysis. Sentinel-1 Ground Range Detected (GRD) scenes availability ranges from October 2014 to present day from Sentinel-1A and late 2016 for Sentinel-1B. Sentinel-2A TOA from June 2015 and Sentinel 2B from June 2017. Sentinel-3A provides data from mid-2016 and Sentinel-3B from mid-2018. Any HARMONIA application aiming to use high-resolution surface reflectance, thermal observations and potentially high-resolution (10-30m) instantaneous atmospheric composition (as a by-product of the atmospheric correction scheme) might need to use Landsat data which long term data record (earliest images from 1972) could prove highly useful to characterise urban environments.

Looking to the future the Sentinel series of satellites is well supported programmatically by Copernicus. Current funding is in place for the next 6 years and in cooperation with ESA a long term scenario building continuation and improvement to the current missions is in place. As shown in Figure 8 the domain or C-Band SAR the current Sentinels 1A and 1B should last well into the current decade, based on historic performance of ESA missions, and beyond that the C and D models of the current series will ensure operations beyond 2030.

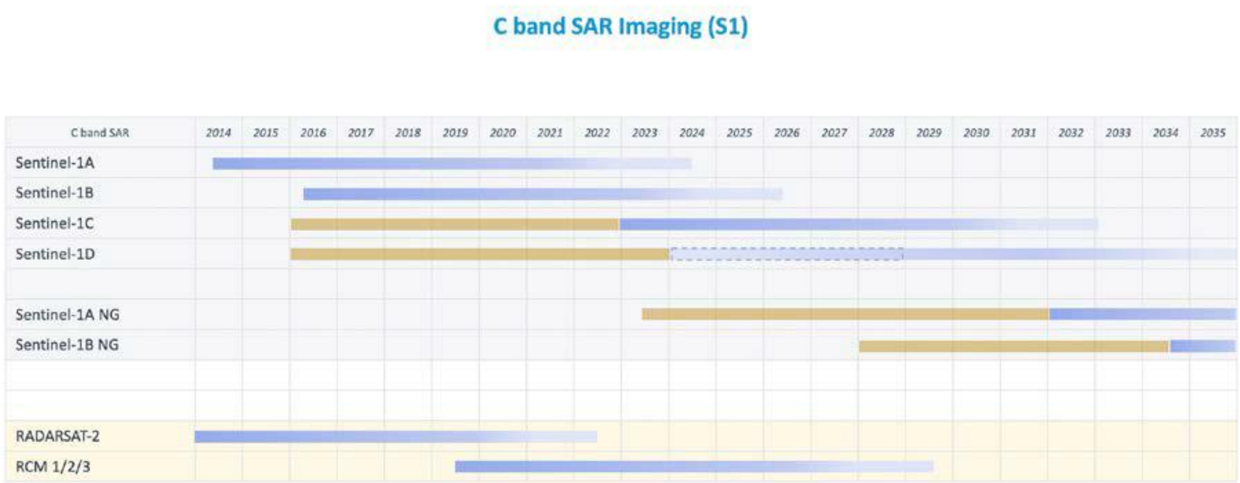

Figure 8 –C-band SAR timeline: Shwing good coverage of observations from Sentinel-1, Radarsat and RCM for the next 10-15 years (Source: ESA)

This situation is similar in the high resolution multispectral optical domain with the current Sentnel-2 satellites and their immediate replacements scheduled to continue operations for the foreseeable future, as shown in Figure 9. In addition to the Copernicus missions, the newly launched Landsat 9 will provide additional high-quality observations with the additional benefit of including thermal channels.

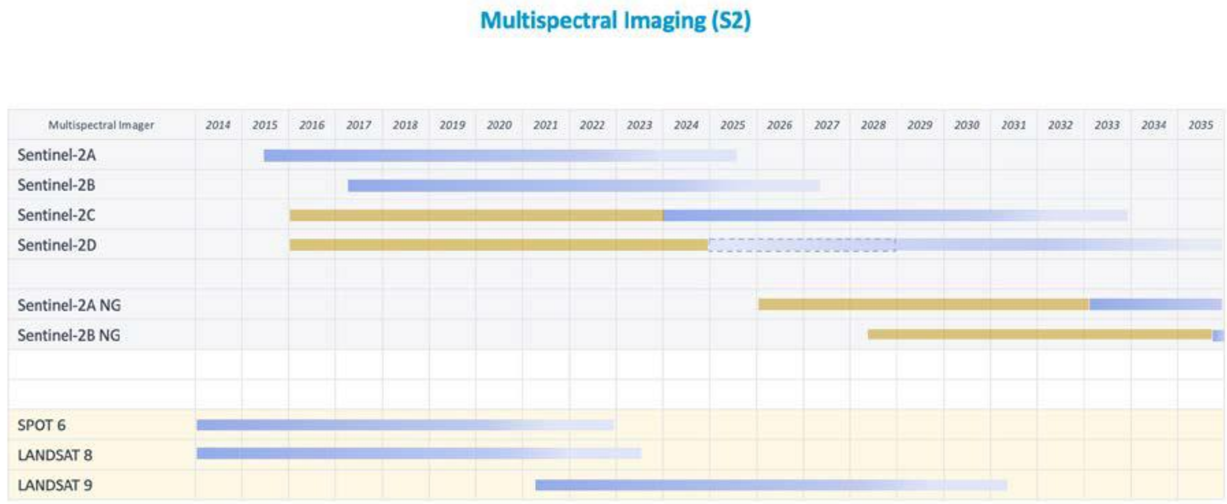

Figure 9 – High-resolution optical data platforms timeline: Current Sentinel-2 satallites wull be followed by two new models in the current series, before the next generation (NG) models. Together with SPOT-6, Landsat-8 & Landsat-9 there is a good supply of high resolution optical data for the foreseeable future. (Source: ESA)

Digital Elevation Models

The DEMs required by the HARMONIA applications are i) the NASA Shuttle Radar Topography Mission (SRTM) collection and the ASTER GDEM2 (Global Digital Elevation Model Version 3). SRTM was the primary payload on the STS-99 mission of the Space Shuttle Endeavour, which launched February 11th, 2000, and flew for 11 days. The first version of the ASTER GDEM, released in June 2009, was generated using stereo-pair images collected by the ASTER instrument on-board Terra. ASTER GDEM coverage spans from 83 degrees north latitude to 83 degrees south, encompassing 99 percent of Earth's landmass. Version 3 shows significant improvements over the previous release. However, users are advised that the data contains anomalies and artifacts that will impede effectiveness for use in certain applications. NASA SRTM V3.0 has eliminated all voids by filling primarily from ASTER GDEM2, and secondarily from the USGS Global Multi-resolution Terrain Elevation Data (GMTED2010). Since the HARMONIA applications are going to focus on regional to local scale applications, assessing potential sources of contemporaneous DEMs would be advisable.

4.3.1. Spatial resolution

The main trade off when analysing, particularly urban areas, using EO data is the length of the data record against the spatial resolution and the associated temporal acquisitions. Figure 10 depicts six different optical datasets going from MODIS-Aqua coarse resolution at 250m to Worldview-4 at 0.3m very-high spatial resolution. It is evident that for very local applications such as urban heat islands higher spatial resolution for LST is required. Typical LST products are produced at 500-1000m. A synergistic approach combining polar orbiters such as MODIS/VIIRS together with high-res data from Landsat and geostationary data, such as MSG could provide LST at a better spatial resolution (~100m) while still capturing the temperature diurnal cycle.

Typically, albedo has been produced using coarse resolution data. Within the HARMONIA applications albedo products are required to analyse the energy balance in urban areas. However in order to capture potential hot-spots of change in terms of radiative forcing, having better spatial resolution would be ideal. Deriving land surface broadband albedo from Sentinel-2 could be a good alternative to using only coarse resolution. Bonafoni et al. (2020) derived land surface broadband albedo using Sentinel-2 data. Validations were carried out on six field stations of the Surface Radiation Budget Network (SURFRAD), and an urban area (Perugia, Central Italy), to better evaluate the impact of the spatial resolution on the albedo retrieval. They obtained root mean square error (RMSE) for the two tests that presented very good values for both algorithms (around 0.02)

Figure 10 – Comparison of different optical datasets and their corresponding spatial resolutions. Image taken from http://radiant.earth

When analysing the climate and weather datasets required within the HARMONIA applications a combination of RA and in situ data is required. RA data presents a consistent, well characterised, long-term data record where the spatial resolution varies from approximately 9-30km. Even for very large urban conglomerates such as Milan, it is possible to see in Figure 11 that a single ERA5 grid cell covers a vast area of Milan city centre and the south-southwest suburbs. HARMONIA applications will indeed explore the downscaling of RA datasets to improve the spatial representations of weather and climate variables. Downscaling could focus on ERA5-Land that provides the best spatial resolution but it should not be limited to this alone.

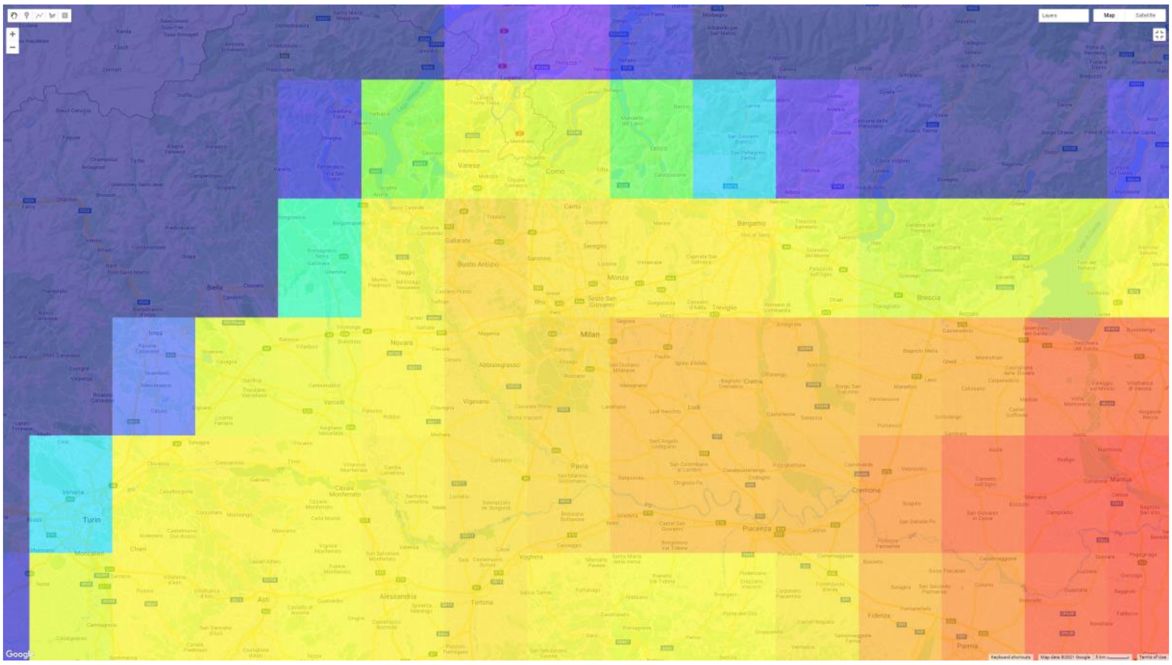

Figure 11 – Daily ERA5 2m air temperature for July 15, 2019, showing the relatively coarse resolution of the data compared to the city scales considered in HARMONIA.

4.3.2. Geographic coverage

No limitations are foreseen in terms of geographic coverage. All EO & RA datasets mentioned in the HARMONIA applications provide full coverage of the AOIs.

4.4. Model and Algorithm Landscape

4.4.1. Gap summary

This section summarises the models coming from the application model specifications provided by the WP4 task leaders for each of the applications, where there is a an amber or red flag concerning the model fitness for purpose or development level

Air Quality

Dry deposition is the transport and removal of pollutants from the atmosphere onto a surface (e.g., building materials, vegetation, soil) in the absence of precipitation. Dry deposition is the most important physical removal mechanism for air pollutants during the dry season. Although there are several dry deposition models in the literature, an assessment of the one most adequate should be carried out and then decide how to adapt it to the HARMONIA sites local conditions.

Traffic emissions factors Traffic emission factors allow users to calculate road vehicle pollutant emission rates for oxides of nitrogen (NOx) and particulate matter (PM - PM10 and PM2.5), based on road type, vehicle speed and vehicle fleet composition. Carbon dioxide (CO2) exhaust emission rates can also be calculated. These outputs would be used to validate and interpret EO based air quality estimates.

Various models are available, but they depend upon local information on road types, speeds and vehicle fleet composition (age, fuel types etc..)

Urban Heat Fluxes

Aerodynamic Resistance Method (ARM). ARM uses the Monin-Obukhov similarity theory (MOST) as the theoretical basis to estimate momentum and scalar fluxes in the atmospheric surface layer. The model is fully developed but needs to be adapted for the cities/areas in question. Parameterization and adjustment are necessary for a specific city/area since several local conditions alter the behaviour of the model.

Objective Hysteresis Model: The Objective Hysteresis Model (OHM) is a well-established and robust method using net radiation and a set of Land Use/Land Cover dependent coefficients for the calculation of storage heat flux. The model is fully developed but needs to be adapted for the cities/areas in question. Parameterization and adjustment are necessary for a specific city/area since several local conditions alter the behaviour of the model.

Urban Flash Flooding

Itzï - distributed hydrologic and hydraulic model: Itzï is a dynamic hydrologic and hydraulic model that simulates hydrological processes like rainfall, infiltration surface and drainage flows. Urban flooding models in general are challenging to validate often due to the lack of availability of fit-for-purpose validation datasets. There are however numerous case studies where the modelling system has seemingly been utilised effectively to simulate urban flooding scenarios.

EcPoint-Calibrate ecPoint-Calibrate is a conditional verification tool to compare numerical weather prediction (NWP) model outputs against point observations. The ecPoint-calibrate element is intended to be used to downscale precipitation fields from lower resolution meteorological datasets to develop experimental synthetic weather radar images used as dynamic forcing data input for the hydrological model. Using this method/software is however still somewhat experimental and other algorithms/models may either be employed or developed during the project.

Geohazards

Geotechnical analysis: Several geotechnical approaches can be applied to interpret the observed ground displacements from the processing of SAR (Synthetic Aperture Radar) images. These approaches range from simplified analytical methods aimed at assessing the origin of urban subsidence or landslide movements, to complex numerical methods. The precise details of these approaches have not yet been finalised and need further analysis in WP4.

Across all of the applications, there is also an issue related to data for validating models in the urban context. Several of the applications need data either for calibration and validation of models, or for training algorithms in the context of AI/ML approaches. So the model fitness for purpose will be constrained by the availability of relevant high resolution data.

4.4.2. Atmospheric correction

The standard atmospheric correction scheme used on Sentinel-2 products is not necessarily fit for purpose in regions with high atmospheric loadings of aerosols such as urban areas. Obregon et al. (2019) validated aerosol optical thickness (AOT) and integrated water vapour (IWV) Sentinel-2 products generated using Sen2Cor. For that purpose, data from 94 Aerosol Robotic Network (AERONET) stations over Europe and adjacent regions, covering a wide geographical region with a variety of climate and environmental conditions and during the period between March 2017 and December 2018 were used. The comparison between Sen2Cor atmospheric composition and AERONET shows a better agreement for IWV than the AOT, with normalized root mean square errors (NRMSE) of 5.33% and 9.04%, respectively. Whilst the R^2 were 0.99 and 0.65 for IWV and AOT, respectively. Meanwhile, CAM AOD at 550nm against AERONET showed a R^2 of 0.8 for 547 samples. This clearly indicates that the RA product performance was better than Sen2Cor. HARMONIA applications are strongly encouraged to assess different atmospheric correction schemes such the Sensor Invariant Atmospheric Correction (SIAC) which has shown strong performance compared against AERONET data (Yin et al., 2021). The methodology was demonstrated on Sentinel 2 and Landsat 8 data. Aerosol Optical Thickness (AOT) retrieval for both S2 and L8 shows a very high correlation to AERONET estimates (R^2>0.9, RMSE<0.025 for both sensors), although with a small underestimation of AOT. Total Column Water Vapour is accurately retrieved from both sensors R^2>0.95, RMSE<0.02).

4.4.3. Downscaling

In order to mitigate the foregoing shortcomings for both RA (atmospheric composition and climate/weather) and EO-based datasets, two potential measures could be implemented: (a) statistical correction of the incorporated datasets, and (b) downscaling lower resolution products with an extensive temporal coverage to finer spatial scales. There are various techniques of parametric and non-parametric (purely empirical) statistical downscaling of, for instance, precipitation estimates (Mamalakis et al., 2017 ; Tang et al. , 2020). Emmanouil et al. (2021) derived 4km precipitation datasets based on the downscaling of ERA5 rainfall estimation using a parametric quantile mapping framework. They used NOAA's rain gauge measurements as a benchmark, while findings reveal that the developed product benefits from the strengths of the calibration datasets, demonstrating good performance and robust behaviour over all studied time periods and Köppen climate classification zones.

Another alternative is to use regional climate models (RCMs) that provide finer spatial resolutions (~10km) through initiatives like CORDEX, although with a smaller number of general circulation models (GCMs) and several RCMs to evaluate model uncertainties. ECMWF are producing a dataset of high-resolution CMIP6 climate projections through statistical downscaling using the latest C3S reanalysis products: ERA5 and ERA5-Land. Six essential climate land surface variables from all the CMIP6 projections available on the CDS are being downscaled globally (~25x25km) using ERA5 and over Europe (~9x9km) using ERA5-Land. Those six variables are: temperature (daily, minimum, and maximum), surface precipitation, wind speed, and solar radiation. The data will be quality assured and methodological uncertainties will be quantified.

Additionally, higher spatial resolution for EO optical data might be needed for some of the HARMONIA applications. Several downscaling methodologies have been developed to achieve super-resolutions of sub 5m with Sentinel-2 data. Tao et al. (2020) developed a methodology to derive an averaged effective resolution enhancement factor of about 2.91 times (equivalent to ~3.44 m/pixel for the 10 m/pixel Sentinel-2 bands). This just shows the potential of using very-high resolution commercial EO datasets together with open distribution ones and derived super-resolution datasets that can characterise urban landscape at neighbourhood level as required by some HARMONIA applications.

Higher spatial and temporal resolution EO thermal data are also necessary for the urban heat fluxes estimations. Available thermal imagery of high temporal resolution (daily coverage) are currently available in 1 km spatial resolution from existing satellite sensors (Sentinel-3 SLSTR and MODIS). Several downscaling methodologies have been developed to achieve higher spatial resolution from these sensors (Zhan et al., 2013). A physically-based downscaling algorithm is employed in HARMONIA for achieving high spatial resolution spectral emissivity and Land Surface Temperature (LST) from Sentinel-3 imagery (Mitraka et al., 2015).

Several of the applications need this kind of downscaling and the issues involved have been flagged by the relevant task leaders. Furthermore, there is considerable resource available in WP4 to deal with this specific issue, so while it is a significant challenge it is also one that has been identified by the project and has resources allocated to meet it.

4.4.4. Impact Models

The HARMONIA applications will derive different environmental measurements associated to Urban environments from statistics, metrics, indices, and indicators. WP2 will cover extensively the documentation of the appropriate standards regarding the use of the Risk Metrics Evaluation & Impact Assessment Methodology deliverable D2.6. All HARMONIA applications described within this document aim towards the quantification of urban risk development, however the applications specified in WP4 are deriving a component of risk more associated with the probability of a hazard rather than the risk itself.

The IPCC online glossary defines risk as: “The potential for adverse consequences for human or ecological systems, recognising the diversity of values and objectives associated with such systems”. Risk management is defined as “plans, actions, strategies or policies to reduce the likelihood and/or magnitude of adverse potential consequences, based on assessed or perceived risks”.

In the context of climate change, risks can arise from potential impacts of climate change as well as human responses to climate change. Within the HARMONIA applications it is clear that the applications focused on air quality, temperature and natural hazards such as wildfires, flooding and land displacement (subsidence, landslides, etc.) will not be explicitly linked to these adverse consequences including those on lives, livelihoods, health and wellbeing, economic, social and cultural assets and investments, infrastructure, services (including ecosystem services), ecosystems and species.

When aiming to quantify urban risk in the context of climate change impacts, the HARMONIA applications being developed in WP4 will need to be integrated into the Decision Support System (DSS) to quantify risk as a result of dynamic interactions between climate-related hazards with the exposure and vulnerability of the affected human or ecological system to the hazards.

Hazards, exposure and vulnerability may each be subject to uncertainty in terms of magnitude and probability of occurrence, and each may change over time and space due to socio-economic changes and human decision-making (see also risk management, adaptation, mitigation). The HARMONIA applications should aim to quantify the uncertainties for all products, metrics, indices, etc. in a way that they can be used in a probabilistic framework.

In the context of climate change responses, HARMONIA DSS might consider that risks result from the potential for such responses not achieving the intended objective. Risks can arise, for example, from uncertainty in implementation, effectiveness or outcomes of climate policy, climate-related investments, technology development or adoption, and system adoptions. The applications could consider scenarios where no risk management will take place and assess the outcomes of that action.

Uncertainty quantification should be applied along the whole applications development and particularly to the three core components of risk. The applications will be able to quantify magnitude and frequency of hazards but also the exposure and vulnerability to any given hazard. If a study or conclusion assumes no change or no uncertainty in one of those dimensions (i.e. risk due to flooding considering only changes in flood frequency and magnitude, but not societal conditions) it should be stated explicitly since not only the magnitude but even direction of change in risk could depend on past or future socio-economic developments.

5. Opportunities and Conclusions

In conclusion, this section highlights opportunities based on data that have not been identified so far for use in the project and summarises the main gaps and issues found.

5.1. Opportunities in new datasets

The applications are relying primarily for their EO and modelled data on products from the Copernicus programme. This reflects the quality of the data available from the Sentinels and from the Copernicus services, the open data policy, and the assurance of long-term provision of the data and services. There are some potentially relevant data types that have not yet been identified by the applications. Figure 12 identifies the missions dedicated to atmospheric remote sensing that are part of the Copernicus programme, along with other related missions.

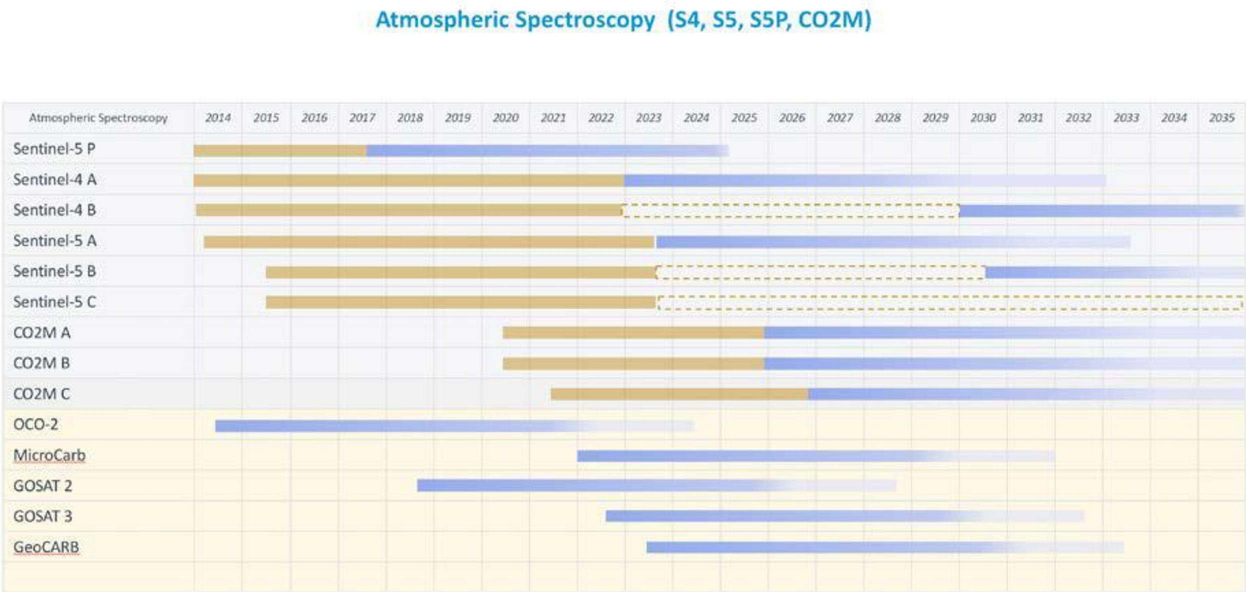

Figure 12 – Atmospheric spectroscopy platforms timeline including Sentinel-4, Sentinel-5, Sentinel-5P. Timeline includes CO_2 dedicated instruments such as, CO2M, OCO-2, MicroCarb, GOSAT and GeoCARB (Source: ESA)

Although the resolution of the current sensors may not be directly applicable to the urban environment (at approximately 50KM for Sentinel 5P, for example) the data from these missions may be valuable for providing context, boundary conditions and diagnosis of effects at urban scales.

A significant number of hyperspectral and thermal sensors are also becoming available in the future, as shown in Figure 13. Already, Landsat 9 has begun to re-instate the relatively high-resolution thermal data that was lost by the failure of Landsat 8. This provides thermal data at 100m resolution and although the temporal sampling will be insufficient for operational applications, it may provide a valuable source information for validation, for downscaling, and for helping to interpret the impacts of thermal hazards. In the future the proposed Sentinel LSTM will improve and extend this capacity.

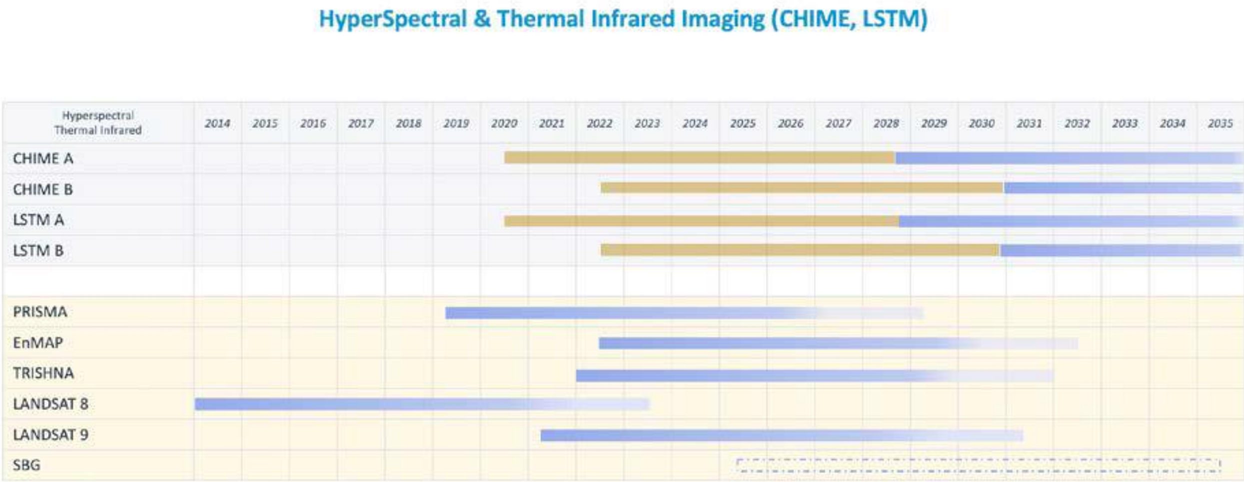

Figure 13 – Hyper-spectral and thermal infrared imaging platforms timeline (Source: ESA)

5.2. Newspace mapping constellations

Newspace, or Space 2.0, missions are loosely defined as space programmes that exploit recent advances in miniaturisation of space components (bus, payload, sub-systems, etc.) aligned with lowered cost of access to space (cheaper launch opportunities) and lower operational costs through services such as ground segment-as-a-service concepts. The Earth observation domain is witnessing a proliferation of constellation concepts promising to deliver a wide variety of measurements at a varying levels of spectral, radiometric, temporal and spatial resolution. We list here a selection of the Newspace missions already in orbit and those under development and have raised capital to finance the constellation. A summary is provided including details of payload, constellation size and links for further information. We have chosen those missions with relevance to HARMONIA and application such as monitoring green spaces, thermal sensing and air quality/atmospheric composition.

Company	Main application	Payload type	Constellation size	URL
Planet	General purpose	Optical, NIR, Pan	~200	www.planet.com
BlackSky	General purpose	Optical, Pan	6-60	www.blacksky.com
21AT	General purpose	Optical, NIR, Pan	3	www.21at.sg
Satellogic	General purpose	Optical, hyperspectral	17-300	satellogic.com
GHGsat	Methane emission monitoring	Imaging spectrometer	3-10	www.ghgsat.com
SPIRE Global	Weather data	GNSS radio occultation		spire.com
ICEYE	General purpose; persistent monitoring	X-band SAR	4-18+	www.iceye.com
Capella Space	General purpose	X-band SAR	up to 36	www.capellaspace.com

Table 13 - Selection of Newspace constellations with assets in orbit

As Table 13 illustrates, there is a growing number of companies offering commercial EO data of different types. The most well known is Planet, a company that recently floated on the New York Stock Exchange with a market capitalisation of approximately $2.8 billion. The Planet data offering is basic optical imaging at high temporal frequency and at the 2-3m spatial resolution. This type of data would be useful for identifying, mapping and monitoring the extent of urban green spaces. Other Newspace companies offering this type of data include BlackSky, Satellogic and 21AT.

Other data types are available. For instance, ICEYE and Capella Space now offer high resolution X-band SAR imagery. The all day/night capability means that that such data is useful for so-called 'persistent monitoring' applications – the ability to revisit the same location frequently with no cloud cover issues. Such capability is useful for regular monitoring of assets such as ports, buildings and other valuable infrastructure. The limited area coverage means it is not so well suited to wide area monitoring, such as agriculture. However, green spaces of limited extent could be monitored on a frequent basis.

Emerging on the Newspace market are two forms of atmospheric data. First is GHGsat and the evolution of small, highly performant imaging spectrometers that can be used for point source methane and CO_2 detection. Second is the technique called GNSS reflectometry or radio occultation (RO). Sensors on board a 'small sat' can pick up the reflected signals emitted by the GPS or other GNSS constellations. The attenuation of GNSS signals in the atmosphere and on the ground surface, generally linked with moisture, can be converted into data sets such as soil moisture, atmospheric density, wind speeds and other products. SPIRE Global are one of the world leaders in GNSS RO with a small constellation of operational satellites.

5.3. New constellation concepts

A large number of additional constellations are under development, largely funded by private finance. Many cover similar spectral domains (mainly optical and X-band SAR) addressed by existing Newspace players. Others offer additional wavelengths for applications in thermal imaging such as building heat emissions and wildfire detection and monitoring. As before, this is a selection of missions for which information is available and many other missions are under development.

Company	Main application	Payload type	Constellation size	URL
Resilience Constellation	General purpose	Optical, NIR	5-10	www.resilienceconstellation.com
Wyvern	General purpose	Hyperspectral	tbc	https://wyvern.space/
Synspectiver	General purpose	X-band SAR	tbc	synspective.com
MethaneSat	Methane monitoring	Imaging spectrometer	tbc	www.methanesat.org
Carbon Mapper	Methane monitoring	Imaging spectrometer	tbc	carbonmapper.org
PlanetIQ	Weather data	GNSS Radio occultation	tbc	planetiq.com
Satellite Vu	Thermal imaging	Infrared spectrometer	5	www.satellitevu.com
Ororatech	Thermal imaging	Multispectral thermal	tbc	ororatech.com

Table 14 - Selection of proposed Newspace constellations

Novel thermal imaging constellations are proposed by companies such as Satellite Vu and Ororatech. Satellite Vu have signed a supplier contract with Surrey Satellite Technology Limited to build five satellites carrying infrared spectrometers with GSD of approximately 3-5m[1]. This will allow thermal emission monitoring at individual building level. Ororatech seem to be aiming at the wildfire detection and monitoring market but with similar imaging technology. Both could contribute to city-scale thermal imaging and monitoring.

Further methane detection constellations are proposed with Carbon Mapper emerging as a partnership between NASA JPL and Planet. MethaneSat is another system aiming at what looks to become a crowded market place. Additional X-band SAR mission are proposed by Synspective, a Japanese firm and Umbra, a US firm that launched its initial test mission in June 2021. Both offer sub-metre scale imagery. Wyvern is a Canadian start up offering band-selectable hyperspectral imaging.

Among the organisations planning to launch high resolution optical payloads is the Resilience Constellation (RC). This is a new approach seeking to develop joint mission proposals with government clients around the world. The data will initially be made available to partners developing national reporting and mapping systems designed to support climate resilience in developing countries. RC has a variety of mission concepts under development but the initial offering will be a 5-8m resolution RGB/NIR imaging constellation with additional red-edge bands for better vegetation discrimination and plant health analysis capability.

5.4. GEO and EuroGEOSS

The Group on Earth Observation (GEO) is a partnership of 104 national governments plus the European Commission and 118 Participating Organizations. GEO implements a Global Earth Observation System of Systems (GEOSS) and envisions a future where decisions and actions for the benefit of humankind are informed by coordinated, comprehensive and sustained Earth observations

The European caucus is the largest caucus of GEO comprising several European governments plus the European Commission and many participating organisations which have an EO mandate within Europe. Since 2017 there has been a European GEO initiative, called EuroGEOSS. EuroGEOSS is the European component of GEOSS

Harmonia aims to develop, deploy and validate several applications to enhance the resilience of urban areas. In order to achieve that Harmonia scientists will make the most of the Copernicus EO data, additionally, EO and ancillary data available in EuroGEOSS, and other operational EU and international services like DIAS and urban Thematic Exploitation Platform (TEP), in order to quantify climate change effects on urban areas.

According to the analysis undertaken here, the applications are well served by Copernicus data and services. The main gaps are to be found in quite specific models and in in situ data related to specific locations, rather than generic capabilities and data that could be provided on a consistent European or Global basis by GEOS

The new services & applications should be then integrated in EuroGEOSS as new applications. Each Harmonia application can be registered as a EuroGEOSS application. This will support combining and strengthening the Harmonia applications by adding other potential data sources, adapt the applications to larger geographical domains and overall use additional GEOSS resources or relevant activities from the GEO work programme.

[1] https://www.sstl.co.uk/media-hub/latest-news/2021/sstl-signs-contract-with-satellite-vu-for-mid-wave

5.5. Concluding remarks

Data Commonalities

The most commonly used data come from the Copernicus programme and include Sentinel-1 SAR data, Sentinel-2 optical data, Sentinel-3 thermal data from the space components, and a range of different products from the ERA5 and ERA-5 land reanalysis products produces by the Copernicus Climate Change Service. In addition, several products from CAMS are also used by several applications. The most commonly used non-Copernicus data source is NASA MODIS data, which while lower resolution than the other cited, has a very long time data record and high temporal sampling making it useful for tracking urban environments through time.

There are no particular issues highlighted in terms of access to EO data. All the EO data identified for the applications is readily accessible now. In terms of future availability, all of the key datasets are secured in the medium term (5-10 years) through funded programmes that extend or enhance the current provision of data.

Gap analysis – Data

The primary gaps in data sources identified for the applications are in the fitness for purpose of some EO and modelled datasets for use at fine spatial scales in urban environments, for high resolution EO data for validation, and for in situ data of all kinds. Almost every application has an unknown, or uncertainty, related to in situ data either for calibration / validation or for operational processes. These data are extremely diverse in character and the details, sources and formats of each dataset are likely to differ from city to city involved in the project.

Gap Analysis – Models

Most of the models have been identified but there are gaps, in linking EO and in situ data, in downscaling results to city neighbourhood scale and in relating physical parameters to human and economic risks. The requirements for downscaling are critical to the project, but this is a known challenge, and substantial resources are dedicated in WP4 to address it. As specified, most of the applications are dealing more with hazard than risk and so the interface between WP5 and W6 to translate the output of the applications into decision support information will be critical.

Nearly all the data issues are related to in situ observations – either for model calibration / validation or for operational use. There is a lot of uncertainty in this area, even down to what data can be made available to the applications by the cities. Iterative steps will be needed here because uncertainty over what can be provided affects the cities or models that are applicable. By contrast, the situation for EO and modelled datasets is more straightforward. All the EO data that is required by the applications can easily be accessed on-line from a range of different sources. A strong focus is therefore needed in WP3 on trying to harmonise the in situ data needs of the applications so that they, and subsequently the DSS, can operate seamlessly across all the pilots.

With respect to the impact models and downscaling issues, it will remain important to work with city authorities to understand the indicators that they use to manage hazards and identify exposure and vulnerability and in the end derive urban risks. The activities in WP4 that lead into the Applications in WP5 may rely very heavily on good quality training data if AI / ML techniques are going to be applied. This puts a further demand on the need for high quality in situ data.

Opportunities

All of the EO data that has been proposed to fulfil the HARMONIA applications comes from already existing EO missions. This is not surprising because the applications need to be demonstrated within the 4-year timeframe of the project. However, there are many new satellite missions being developed, either as enhancements of existing constellations, or as entirely new concepts that could enhance the HARMONIA applications in the future. There is therefore a strong opportunity for the HARMONIA applications to be sustained and improved in the future.

6. References

P. Alpert, O. Shvainshtein and P. Kishcha (2012). AOD Trends over Megacities Based on Space Monitoring Using MODIS and MISR. American Journal of Climate Change, Vol. 1 No. 3, 2012, pp. 117-131. https://doi.org/10.4236/ajcc.2012.13010

Bonafoni, S., Ieee, M. & Sekertekin, A (2020). Albedo Retrieval From Sentinel-2 by New Narrow-to-Broadband Conversion Coefficients. 17, 1618–1622. https://doi.org/10.1109/LGRS.2020.2967085

Chrysoulakis, N., Grimmond, S., Feigenwinter, C., Lindberg, F., Gastellu-Etchegorry, J.-P., Marconcini, M., Mitraka, Z., Stagakis, S., Crawford, B., Olofson, F., Landier, L., Morrison, W., Parlow, E., 2018. Urban energy exchanges monitoring from space. Sci. Rep. 8, 11498. https://doi.org/10.1038/s41598-018-29873-x

Courty, L. G., Soriano-Monzalvo, J. C., & Pedrozo-Acuña, A. (2019). Evaluation of open-access global digital elevation models (AW3D30, SRTM and ASTER) for flood modelling purposes. Journal of Flood Risk Management, 1–14. https://doi.org/10.1111/jfr3.12550

Courty, L. G., Rico-Ramirez, M. Á., & Pedrozo-Acuña, A. (2018). The Significance of the Spatial Variability of Rainfall on the Numerical Simulation of Urban Floods. Water, 10(2), 207. https://doi.org/10.3390/W10020207

Courty, L. G., Pedrozo-Acuña, A., & Bates, P. D. (2017). Itzï (version 17.1): an open-source, distributed GIS model for dynamic flood simulation. Geoscientific Model Development, 10(4), 1835–1847. http://doi.org/10.5194/gmd-10-1835-2017

Emmanouil, S. & Langousis, A. An ERA-5 Derived CONUS-Wide High-Resolution Precipitation Dataset Based on a Refined Parametric Statistical Downscaling Framework Water Resources Research. 1–17 (2021). https://doi.org/10.1029/2020WR029548

Feigenwinter, C., Vogt, R., Christen, A., 2012. Eddy Covariance Measurements Over Urban Areas, in: Eddy Covariance. Springer Netherlands, Dordrecht, pp. 377–397. https://doi.org/10.1007/978-94-007-2351-1_16

Feng Yin, Phillip E Lewis, Jose L Gomez-Dans, Quinling Wu (2021). A Sensor Invariant Atmospheric Correction: Sentinel-2/MSI and Landsat-8/OLI. Preprint. https://doi.org/10.31223/osf.io/ps957

S. Grimmond, T.R. Oke 1991. Heat Storage in Urban Areas: Local-Scale Observations and Evaluation of a Simple Model. Joyrnal of Applied Meteorology, vol 38,pp 922-940. doi: https://doi.org/10.1175/1520-0450(1999)038<0922:HSIUAL>2.0.CO;2

Hersbach et al. (2018). Operational global reanalysis: progress, future directions and synergies with NWP. ERA Report Series, 27. http://dx.doi.org/10.21957/tkic6g3wm

Hewson, T.D., Pillosu, F.M. A low-cost post-processing technique improves weather forecasts around the world. Commun Earth Environ 2, 132 (2021). https://doi.org/10.1038/s43247-021-00185-9

Jacquemoud S., Verhoef W., Baret F., Bacour C., Zarco-Tejada P.J., Asner G.P., François C., & Ustin S.L. (2009), PROSPECT + SAIL models: a review of use for vegetation characterization. Remote Sensing of Environment, 113, S56-S66. https://doi.org/10.1016/j.rse.2008.01.026

López-Saldaña, G., Bistinas, I. & Pereira, J. M. C. Global analysis of radiative forcing from fire-induced shortwave albedo change (2015). Biogeosciences 12, 557–565. https://doi.org/10.5194/bg-12-557-2015

Mamalakis, A., Langousis, A., Deidda, R., & Marrocu, M. (2017). A parametric approach for simultaneous bias correction and high-resolution downscaling of climate model rainfall. Water Resources Research, 53(3), 2149– 2170. https://doi.org/10.1002/2016WR019578

Mitraka, Z. *et al.* (2015) 'Urban Surface Temperature Time Series Estimation at the Local Scale by Spatial-Spectral Unmixing of Satellite Observations', *Remote Sensing*. Multidisciplinary Digital Publishing Institute, 7(4), pp. 4139–4156. doi: 10.3390/rs70404139.

Obregon, M.A., Rodrigues, G., Costa, M. J., Potes, M. & Silva, A. M . Validation of ESA Sentinel-2 L2A Aerosol Optical Thickness and Columnar Water Vapour during 2017 – 2018. (2019).

Pinty, B., T. Lavergne, M. Voßbeck, T. Kaminski, O. Aussedat, R. Giering, N. Gobron, M. Taberner, M. M. Verstraete, and J.-L. Widlowski (2007). Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS) - Multiangle Imaging Spectroradiometer (MISR) albedo products, J. Geophys. Res., 112, D10116. https://doi.org/10.1029/2006JD008105

Sara Janhäll. (2015). Review on urban vegetation and particle air pollution e Deposition and dispersion. Atmospheric Environment, 105, 130–137. https://doi.org/10.1016/j.atmosenv.2015.01.052

Tang, G., Clark, M. P., Newman, A. J., Wood, A. W., Papalexiou, S. M., Vionnet, V., & Whitfield, P. H. (2020). SCDNA: A serially complete precipitation and temperature dataset for North America from 1979 to 2018. Earth System Science Data, 12(4), 2381– 2409. https://doi.org/10.5194/essd-12-2381-2020

Tao, Yu, Siting Xiong, Rui Song, and Jan-Peter Muller (2021). Towards Streamlined Single-Image Super-Resolution: Demonstration with 10 m Sentinel-2 Colour and 10–60 m Multi-Spectral VNIR and SWIR Bands. Remote Sensing 13, no. 13: 2614. https://doi.org/10.3390/rs13132614

Vitolo, Claudia; Di Giuseppe, Francesca; Barnard, Christopher; Coughlan, Ruth; San-Miguel-Ayanz, Jesus ; Libertá, Giorgio ; Krzeminski, B. ERA5-based global meteorological wildfire danger maps. Nat. Sci. Data 1–11 (2020). https://doi.org/10.1038/s41597-020-0554-z

Van Wagner, C. E. et al (1974). Structure of the Canadian forest fire weather index, vol. 1333. Environment Canada, Forestry Service

H.C. Ward, S. Kotthaus, L. Järvi, C.S.B. Grimmond. Surface Urban Energy and Water Balance Scheme (SUEWS): development and evaluation at two UK sites Urban Climate, 18 (2016), pp. 1-32, 10.1016/j.uclim.2016.05.001

Yin F, Gomez-Dans J, Lewis P, Lopez-Saldana G. (2019). The Sensor Independent Atmospheric Correction (SIAC) approach applied to Sentinel-2 and Landsat-8 data. In Geophysical Research Abstracts; vol. 21. https://meetingorganizer.copernicus.org/EGU2019/EGU2019-11269.pdf

Zhan, W. *et al.* (2013) Disaggregation of remotely sensed land surface temperature: Literature survey, taxonomy, issues, and caveats, *Remote Sensing of Environment*. Elsevier Inc., 131(19), pp. 119–139. doi: 10.1016/j.rse.2012.12.014.

7. Appendix A: Data and Model tables

7.1. Urban Air Quality

Urban Air Quality Data

	Sentinel-2	R	A	G
Summary	The Copernicus Sentinel-2 mission comprises a constellation of two platforms in the same sun-synchronous orbit with a swath width of 290km with a revisit time of ~5 days with both satellites. S2 carries a sensor delivering optical / near infra-red imagery.			
Technical characteristics	- 4 x 10m bands: blue (~493nm), green (560nm), and red (~665nm), near infra-red (~833nm); - 6 x 20m bands: 4 narrow Bands in the VNIR vegetation red edge spectral domain (~704nm,~740nm, ~783nm and ~865nm) and 2 wider SWIR bands (~1610nm and ~2190nm); - 3 x 60m bands mainly focused towards cloud screening and atmospheric correction (~443nm for aerosols and ~945nm for water vapour) and cirrus detection (~1374nm). - Radiometric resolution is 12bit The characteristics of the data are well suited for identification of urban green areas, with even small parks down to 0.2ha identifiable. For input into high resolution air quality, the main limitation is temporal sampling at ~4 days and also the fact that the time of sample is always the same.			G
Availab-ility	Sentinel-2A: data available from late 2015 Sentinel-2B: data available from mid 2017 Good support for long term sustainability but historical record is short		A	
Technical Access	Sentinel-2 products available for users (either generated by the ground segment or by the Sentinel-2 Toolbox) are Level-1C and Level-2A. The Copernicus Open Access Hub provides complete, free and open access to Sentinel-2 data. Additionally, Sentinel-2 data products are available in the Copernicus Data and Information Access Service (DIAS) cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services. Extensive M2M discovery and download options.			G
QA	The radiometric accuracy is less than 5% (goal 3%). Radiometric resolution is also dependent upon the Signal to Noise Ratio (SNR) of the detector. The system is under constant quality monitoring by ESA			G
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			G

MODIS BRDF Data Specification

	MODIS BRDF	R	A	G
Summary	The Moderate Resolution Imaging Spectroradiometer (MODIS) MCD43A1 Version 6.1 Bidirectional Reflectance Distribution Function and Albedo (BRDF/Albedo) Model Parameters dataset is produced daily using 16 days of Terra and Aqua MODIS data at 500 metre resolution. The BRDF model parameters are used to derive surface reflectance under any feasible acquisition and illuminations conditions for all MODIOS spectral bands in the visible, near-infrared and shortwave infrared.			
Technical characteristics	Daily MODIS surface reflectance data are used to invert a BRDF model. The daily observations are temporally weighted to the ninth day of the retrieval period. The product coded as MCD43A1, it provides the three model weighting parameters (isotropic, volumetric, and geometric) used to derive the Albedo (MCD43A3) and Nadir BRDF-Adjusted Reflectance (NBAR) (MCD43A4) products. - 2 x 250m bands: red (620-670nm) and near-infrared (841-876nm) - 4 x 500m bands: blue (459-479nm), green (545-565nm), shortwave near-infrared (1230-1250nm) and shortwave infrared (1230-1250nm). - Radiometric resolution is 12bit MODIS BRDF model parameters are suited for characterising the land surface on a daily basis and help creating the surface reflectance expectation required by the atmospheric correction scheme used within the Urban Air Quality application.			
Availab-ility	Combined products, Terra + Aqua, are available from July 2001 making them ideal for long term analysis. However, in this application only data from 2017 will be used to match the Sentinel-2 era. The methodology to derive the BRDF model parameters has been adapted to create a similar product but using VIIRS data. This will ensure continuity once the Terra and Aqua platforms are deactivated.			
Technical Access	MODIS products are freely available from a wide range of data providers. The Land Processes Distributed Active Archive Center (LP DAAC) is the most common data portal for MODIS data and provides access using different interfaces from automated APIs to data servers. Additionally, MODIS data can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	MODIS data has been validated and the BRDF product seats at stage 3 on the CEOS LPV validation hierarchy. Part of the validation involves comparison with *in situ* albedo observations such as AmeriFlux, NEON and SURFRAD. Liu et al. (2017) used this technique to validate the MODIS BRDF C006 snow-free product (using only full inversion retrievals) and found an overall 0.0195 RMSE with a -0.0038 bias.			
Legal	Open access. NASA/USGS give freedom to use, modify and distribute basic and adapted products.			

MODIS LAI Data Specification

ID	MODIS Leaf Area Index	R	A	G
Summary	The MCD15A2H Version 6 Moderate Resolution Imaging Spectroradiometer (MODIS) Level 4, Combined Fraction of Photosynthetically Active Radiation (FPAR), and Leaf Area Index (LAI) product is an 8-day composite dataset with 500 meter pixel size.			
Technical characteristics	LAI is defined as the one-sided green leaf area per unit ground area in broadleaf canopies and as one-half of the total needle surface area per unit ground area in coniferous canopies. FPAR is defined as the fraction of incident photosynthetically active radiation (400-700 nm) absorbed by the green elements of a vegetation canopy. The MODIS LPAR/LAI algorithm chooses the best pixel available from all the acquisitions of both MODIS sensors located on NASA's Terra and Aqua satellites from within the 8-day period. LAI derived from MODIS data provides a long-term data record that is suitable for long term data analysis. However, given the 500m spatial resolution is not necessarily adequate for city-level characterisation.			
Availab-ility	Combined products, Terra + Aqua, are available from July 2001 making them ideal for long term analysis. However, in this application only data from 2017 will be used to match the Sentinel-2 era. The methodology to derive the FPAR/LAI has been adapted to create a similar product but using VIIRS data (VNP15A2H). This will ensure continuity once the Terra and Aqua platforms are deactivated.			
Technical Access	MODIS products are freely available from a wide range of data providers. The Land Processes Distributed Active Archive Center (LP DAAC) is the most common data portal for MODIS data and provides access using different interfaces from automated APIs to data servers. Additionally, MODIS data can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	MODIS data has been validated and the LAI product seats at stage 2 on the CEOS LPV validation hierarchy. With expanded and updated ground-measured LAI validation datasets, the validation practice has been conducted for C6 LAI product. When all biomes are taken in account, Yan et al. (2016) found that the accuracy of LAI is 0.69 LAI units RMSEs.			
Legal	Open access. NASA/USGS give freedom to use, modify and distribute basic and adapted products.			

CAMS Data Specification

ID	Copernicus Atmsophere Monitoring Service (CAMS)	R	A	G
Summary	The Copernicus Atmosphere Monitoring Service (CAMS) provides operational, consistent, and quality-controlled information related to air pollution and health, greenhouse gases and climate forcing at global scale.			
Technical characteristics	CAMS produces global operational forecasts for atmospheric composition twice a day. The forecasts consist of more than 50 chemical species (e.g. ozone, nitrogen dioxide, carbon dioxide) and seven different types of aerosol. Additionally, the CAMS global reanalysis ECMWF Atmospheric Composition Reanalysis 4 (EAC4) is the fourth generation ECMWF global reanalysis of atmospheric composition. Reanalysis combines model data with observations from across the world into a globally complete and consistent dataset using a model of the atmosphere based on the laws of physics and chemistry. Datasets are generated at surface, total column, model levels and pressure levels. For the operational ones at 0.4°x0.4°, 1-hourly (single-level), 3-hourly (multi-level), reanalysis at 0.75°x0.75° and 3-hourly temporal resolution. The characteristics of the data are well suited to be used as a prior within the atmospheric correction data assimilation scheme. However, they are not fit for purpose to monitor atmospheric composition at local scale, particularly where topography would influence heavily pollutant distribution.			
Availab-ility	The observing system has changed drastically over time, and although the assimilation system can resolve data holes, the initially much sparser networks will lead to less accurate estimates. For this reason, EAC4 is only available from 2003 onwards. The operational forecast is available from 2015 to present day. Good support for long term sustainability but historical record is short			
Technical Access	The Copernicus Atmosphere Data Store (ADS) provides free access to all CAMS products. The service includes interactive and automated downloads via an API. Some single-level (surface) parameters like the total aerosol optical depth at 550nm are available from the Google Earth Engine.			
QA	The CAMS reanalysis has smaller biases compared with most of the independent ozone, carbon monoxide, nitrogen dioxide and aerosol optical depth observations used for validation in Inness et al. (2019) this paper than the previous two reanalyses and is much improved and more consistent in time, especially compared to the Monitoring Atmospheric Composition and Climate (MACC) reanalysis.			
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

Very High-Resolution EO Data Specification

	Planet, SPOT, Pleiades, etc	R	A	G
Summary	Multispectral optical Earth observation data with a spatial resolution of 1-10m such as Planet, SPOT and Pleiades. These datasets are often acquired in multispectral and panchromatic mode to allow a so-called pan-sharpening to improve spatial resolution. Swath width goes from 20km on Pleiades and 24.6km on Planet to 60km on SPOT6. Revisit time varies from 26 days on SPOT 6 to daily on Pleiades and sub-daily with Planet.			
Technical characteristics	All three aforementioned platforms carry instruments sensing the Earth's surface on visible spectral bands (red, green and blue) plus the near-infrared. SPOR and Pleiades provide an additional panchromatic band covering the whole visible spectrum. All datasets have been orthorectified. The spatial resolution of the datasets is relevant to improve the Urban Climate Zones and together with local AQ measurements and traffic density understand the relationships with emissions. Revisit time is not an issue since the emissions inventory can be updated on an annual basis, hence an update very-high resolution composite would be suitable.			
Availab-ility	All data are available only from about 5 years from the start of the project. No historical data records will necessarily be used, hence available data is adequate. Good support for long term sustainability but being commercial datasets, a potential funding mechanism should be designed to keep an up-to-data data record.			
Technical Access	All products provide data access via dedicated APIs via different vendors.			
QA	Main QA should be performed on the geolocation accuracy. All data providers claim very good accuracy and geometric correction that reduce distortions derived from acquisitions geometry, pointing errors and terrain variability. No actual figures are available but from Assimila's experience all three datasets are adequate for these types of applications. However, an assessment should be carried out to ensure no geolocation issues are present before using any of the datasets.			
Legal	All datasets are commercial. This does not limit their use on an operational system but potential alternatives should be explored.			

Urban Air Quality Models

Sensor Invariant Atmospheric Correction (SIAC)

	Sensor Invariant Atmospheric Correction (SIAC)	R	A	G
Summary, Inputs, outputs	SIAC is a data assimilation based atmospheric correction system that can simultaneously retrieve surface reflectance and atmospheric composition from a combination of coarse resolution (MODIS) and high resolution (Sentinel -2, Landsat etc..), and CAMS atmospheric fields. It uses the expectation of surface reflectance from MODIS Bidirectional Reflectance Distribution Function (BRDF), plus a prior of atmospheric constituents from CAMS to optimally retrieve the surface reflectance for Sentinel-2 pixels as well as a posterior update of the atmospheric constituents with associated uncertainties.			
Key Reference(s)	- Feng Yin, Phillip E Lewis, Jose L Gomez-Dans, Quinling Wu (2021). A Sensor Invariant Atmospheric Correction: Sentinel-2/MSI and Landsat-8/OLI. Preprint. https://doi.org/10.31223/osf.io/ps957 - Yin F, Gomez-Dans J, Lewis P, Lopez-Saldana G. (2019). The Sensor Independent Atmospheric Correction (SIAC) approach applied to Sentinel-2 and Landsat-8 data. In Geophysical Research Abstracts; vol. 21. https://meetingorganizer.copernicus.org/EGU2019/EGU2019-11269.pdf			
Technical fitness for purpose	The results demonstrated are outstanding across a very wide range of atmospheric conditions, including those with very high aerosol loads (eg in China). For this reason it is considered well suited for use over urban areas. The method is demonstrated on Sentinel 2 and Landsat 8 data. Aerosol Optical Thickness (AOT) retrieval for both S2 and L8 shows a very high correlation to AERONET estimates (r^2>0.9, RMSE<0.025 for both sensors), although with a small underestimate of AOT. Total Column Water Vapour is accurately retrieved from both sensors r^2>0.95, RMSE<0.02).			G
Development Level	Code library available on GitHub (https://github.com/MarcYin/SIAC) and as applications on Google Earth Engine. Code is written in Python 3		A	
Model Validation	Good validation results have been demonstrated but not yet published in peer-reviewed literature. Such publication should occur within the next 12 months.		A	
Software QA	Software is publicly available but not formally maintained. Limited user base.	R		
Legal	IPR belongs to University College London, but code is open source and can be used and modified. Fully available for the HARMONIA applications			G

Radiative Transfer Models – ProSAIL and TIP

	ProSAIL and TIP	R	A	G
Summary, Inputs, outputs	The combined PROSPECT leaf optical properties model and SAIL canopy bidirectional reflectance model, also referred to as PROSAIL, have been used to study plant canopy spectral and directional reflectance in the solar domain. PROSAIL has also been used to develop new methods for retrieval of vegetation biophysical properties. The Two-stream Inverse Package (TIP) is built around a one-dimensional two-stream scheme of the radiative transfer within the vegetation canopy. TIP enable the assimilation of operational remote-sensing flux products into a state-of-the-art two-stream radiation transfer scheme suitable for climate models.			
Key Reference(s)	- Jacquemoud S., Verhoef W., Baret F., Bacour C., Zarco-Tejada P.J., Asner G.P., François C., & Ustin S.L. (2009), PROSPECT + SAIL models: a review of use for vegetation characterization. Remote Sensing of Environment, 113, S56-S66. https://doi.org/10.1016/j.rse.2008.01.026 - Pinty, B., T. Lavergne, M. Voßbeck, T. Kaminski, O. Aussedat, R. Giering, N. Gobron, M. Taberner, M. M. Verstraete, and J.-L. Widlowski (2007). Retrieving surface parameters for climate models from Moderate Resolution Imaging Spectroradiometer (MODIS) - Multiangle Imaging Spectroradiometer (MISR) albedo products, J. Geophys. Res., 112, D10116. https://doi.org/10.1029/2006JD008105			
Technical fitness for purpose	Both models can be driven by land surface albedo and in inverse mode derive different vegetation biophysical parameters required within HARMONIA. - PROSAIL has also been used to develop new methods for retrieval of vegetation biophysical properties. Since it links the spectral variation of canopy reflectance, which is mainly related to leaf biochemical contents, with its directional variation, which is primarily related to canopy architecture and soil/vegetation contrast, this link is key to simultaneous estimation of canopy biophysical/structural variables which is key within the HARMONIA applications. - TIP implements the adjoint and Hessian codes, generated using automatic differentiation techniques, of a cost function balancing: 1) the deviation from the a priori knowledge on the model parameter values, and 2) the misfit between the observed remote-sensing fluxes and the two-stream model simulations. The individual weights of these contributions are specified notably via covariance matrices of the uncertainties in the a priori knowledge on the model parameters and the measurements. This uncertainty characterisation is relevant within HARMONIA to add traceability to any applications developed.			G
Development Level	PROSAIL code available on GitHub (https://github.com/jgomezdans/prosail). TIP code available at the Joint Research Centre (JRC) FAPAR website (https://fapar.jrc.ec.europa.eu/_www/models.php#sft_model_two)			G
Model Validation	Both models have been extensively validated for different EO datasets in several biomes			G
Software QA	Source code for both models is publicly available but not formally maintained. Relatively limited user base.		A	
Legal	IPR belongs to University College London, but code is open source and can be used and modified. Fully available for the HARMONIA applications			G

Dry deposition model (TBD)

	TBD	R	A	G
Summary, Inputs, outputs	Dry deposition is the transport and removal of pollutants from the atmosphere to a surface (e.g., building materials, vegetation, soil) in the absence of precipitation. Dry deposition is the most important physical removal mechanism for air pollutants during the dry season. Although there are several dry deposition models in the literature, an assessment of the most adequate one should be carried out and decide how to adapt it to the HARMONIA sites local conditions.			
Key Referenc e(s)				
Technical fitness for purpose				
Develop ment Level				
Model Validation				
Softwa re QA				
Legal				

Traffic emissions factors (TBD)

	TBD	R	A	G
Summary, Inputs, outputs	Traffic emission factors allow users to calculate road vehicle pollutant emission rates for oxides of nitrogen (NOx) and particulate matter (PM - PM10 and PM2.5), based on road type, vehicle speed and vehicle fleet composition. Carbon dioxide (CO_2) exhaust emission rates can also be calculated. These outputs would be used to validate and interpret EO based air quality estimates.			
Key Referenc e(s)				
Technical fitness for purpose	Various models are available, but they depend upon local information on road types, speeds and vehicle fleet composition (age, fuel types etc..)			
Develop ment Level				
Model Validation				
Softwa re QA				
Legal				

7.2. Urban heat fluxes and urban heat islands

Urban heat fluxes and heat islands Data

Sentinel-2 Data Specification

	Sentinel-2	R	A	G
Summary	The Copernicus Sentinel-2 mission comprises a constellation of two platforms in the same sun-synchronous orbit with a swath width of 290km with a revisit time of ~5 days with both satellites. S2 carries a sensor delivering visible / near infra-(VNIR) and shortwave infrared (SWIR) red imagery.			
Technical characteristics	- 4 x 10m bands: blue (~493nm), green (560nm), and red (~665nm), near infra-red (~833nm); - 6 x 20m bands: 4 narrow Bands in the VNIR vegetation red edge spectral domain (~704nm, ~740nm, ~783nm and ~865nm) and 2 wider SWIR bands (~1610nm and ~2190nm); - 3 x 60m bands mainly focused towards cloud screening and atmospheric correction (~443nm for aerosols and ~945nm for water vapour) and cirrus detection (~1374nm). - Radiometric resolution is 12bit The characteristics of the data are well suited for identification of the surface cover fractions, necessary for the urban heat fluxes models.			
Availab-ility	Sentinel-2A: data available from late 2015 Sentinel-2B: data available from mid 2017 Good support for long term sustainability but historical record is short			
Technical Access	Sentinel-2 products available for users (either generated by the ground segment or by the Sentinel-2 Toolbox) are Level-1C and Level-2A. The Copernicus Open Access Hub provides complete, free and open access to Sentinel-2 data. Additionally, Sentinel-2 data products are available in the Copernicus Data and Information Access Service (DIAS) cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services. Extensive M2M discovery and download options.			
QA	The radiometric accuracy is less than 5% (goal 3%). Radiometric resolution is also dependent upon the Signal to Noise Ratio (SNR) of the detector. The system is under constant quality monitoring by ESA			
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

Sentinel-3 Data Specification

	Sentinel-3	R	A	G
Summary	The Copernicus Sentinel-3 mission comprises a constellation of two platforms in the same sun-synchronous orbit with a swath width of 1 420 km (nadir) and 750 km (backwards) for SLSTR instrument with a revisit time of less than one day for SLSTR at the equator with both satellites. S3 carries a Sea and Land Surface Temperature Radiometer (SLSTR) for sea- and land-surface temperature monitoring.			
Technical characteristics	-6 x 500m: 3 narrow Bands in the VIS (~0.554µm, ~0.659µm and ~0.868µm) , and 3 narrow Bands in the SWIR (~1.374µm, ~1.613µm and ~2.25µm) -1 km: 3 narrow Bands in the MWIR/TIR (~3.742µm,~10.85µm and 12.02µm) 2 narrow Bands in the Fire (~3.742µm and ~10.85µm) 1km for thermal measurement is not sufficient for cities. Neighborhood scale measurements (~100 m) are valuable for assessing heat fluxes.			
Availab-ility	-Sentinel-3A: data available from early 2016 -Sentinel-3B: data available from mid 2018 Good support for long term sustainability but historical record is short. MODIS data could complement since 2000.			
Technical Access	Sentinel-3 products available for users (either generated by the ground segment or by the Sentinel-3 Toolbox) are Level-1C and Level-2A. The Copernicus Open Access Hub provides complete, free and open access to Sentinel-3 data. Additionally, Sentinel-3 data products are available in the Copernicus Data and Information Access Service (DIAS) cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services. Extensive M2M discovery and download options.			
QA	The radiometric accuracy is less than 5%. Radiometric resolution is also dependent upon the Signal to Noise Ratio (SNR) of the detector. The system is under constant quality monitoring by ESA			
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

ERA5 Data Specification

	ERA-5	R	A	G
Summary	ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables. The data cover the Earth on a 30km grid and resolve the atmosphere using 137 levels from the surface up to a height of 80km. ERA5 includes information about uncertainties for all variables at reduced spatial and temporal resolutions			
Technical characteristics	-2m temperature : spatial resolution 9km x 9km and temporal resolution 1 hour. Temporal resolution is sufficient, but spatial resolution is not sufficient for cities.			
Availab-ility	-2m temperature: data available from early 1950			
Technical Access	ERA5 provides complete, free and open access data. ERA5 data can be downloaded through the CDS either via the CDS web interface or programmatically using the CDS API service.			
QA				
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

CAMS Data Specification

	CAMS	R	A	G
Summary	The Copernicus Atmosphere Monitoring Service (CAMS) provides consistent and quality-controlled information related to air pollution and health, solar energy, greenhouse gases and climate forcing, everywhere in the world, among which is the Clear-sky surface solar irradiation product			
Technical characteristics	CAMS Clear-sky surface solar irradiation product data are provided with 3-5km spatial resolution and 1m, 15m, 30m, 1h, 1 day and 1 month temporal resolution			
Availab-ility	-CAMS: data available from early 2004			
Technical Access	CAMS provides complete, free and open access data. CAMS data can be downloaded through the web interface.			
QA				
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

Urban heat fluxes and heat islands Models

Local Scale Land Surface Temperature Model (LSLST)

	Local Scale Land Surface Temperature Model (LSLST)	R	A	G
Summary, Inputs, outputs	LSLST methodology is employing physically-based downscaling approaches for achieving high spatio-temporal retrieval of Land Surface Temperature (LST). Thermal infrared (TIR) imagery of 1 km spatial resolution is downscaled to 100 m using higher spatial resolution surface cover information. Ancillary information for spectral libraries combined with the dynamic land cover is used for the emissivity estimation. Inputs: Sentinel-3 SLSTR brightness temperature, land cover fractions (Sentinel-2), representative emissivity (spectral libraries), atmospheric water vapour (ERA5).			
Key Referenc e(s)	Mitraka, Z. et al. (2015) 'Urban Surface Temperature Time Series Estimation at the Local Scale by Spatial-Spectral Unmixing of Satellite Observations', Remote Sensing. Multidisciplinary Digital Publishing Institute, 7(4), pp. 4139–4156. doi: 10.3390/rs70404139.			
Technical fitness for purpose	LSLST was successfully applied in URBANFLUXES project (Chrysoulakis et al., 2018) for the calculation of citywide heat fluxes in a spatial resolution of 100 m x 100 m. The model is suitable for HARMONIA.			
Develop ment Level	The model is fully developed, but needs to be adapted for the cities/areas in question. Parameterization and adjustment is necessary for a specific city/area since several local conditions, alter the behaviour of the model.			
Model Validation	The model results were evaluated with in situ outgoing thermal radiation measurements at urban flux towers in the URBANFLUXES project and good agreement was observed (Chrysoulakis et al., 2018).			
Softwa re QA	There is no software currently available. The method is implemented by FORTH in Matlab. No need for QA.			
Legal	IPR belongs to FORTH.			

Aerodynamic Resistance Method (ARM)

	Aerodynamic Resistance Method (ARM)	R	A	G
Summary, Inputs, outputs	ARM uses the Monin-Obukhov similarity theory (MOST) as the theoretical basis to estimate momentum and scalar fluxes in the atmospheric surface layer. MOST is commonly used in meteorological numerical modelling systems. In the MOST framework, roughness lengths for momentum and heat (z_{0m} and z_{0h}, respectively) are the key parameters identifying the aerodynamic features of underlying surfaces. Inputs: Digital Object Model (i.e. 3D representation of buildings and trees), Land surface temperature, air temperature, humidity, wind velocity, wind direction, radiation			
Key Reference(s)	Chrysoulakis, N., Grimmond, S., Feigenwinter, C., Lindberg, F., Gastellu-Etchegorry, J.-P., Marconcini, M., Mitraka, Z., Stagakis, S., Crawford, B., Olofson, F., Landier, L., Morrison, W., Parlow, E., 2018. Urban energy exchanges monitoring from space. *Sci. Rep.* 8, 11498. https://doi.org/10.1038/s41598-018-29873-x Feigenwinter, C., Vogt, R., Christen, A., 2012. Eddy Covariance Measurements Over Urban Areas, in: Eddy Covariance. Springer Netherlands, Dordrecht, pp. 377–397. https://doi.org/10.1007/978-94-007-2351-1_16			
Technical fitness for purpose	ARM was successfully applied in URBANFLUXES project (Chrysoulakis et al., 2018) for the calculation of citywide sensible and latent heat flux in a spatial resolution of 100 m x 100 m. The model is suitable for HARMONIA assessing sensible heat flux in appropriate scale for cities.			G
Development Level	The model is fully developed, but needs to be adapted for the cities/areas in question. Parameterization and adjustment is necessary for a specific city/area since several local conditions, alter the behaviour of the model.		A	
Model Validation	The model results were evaluated with in situ turbulent sensible heat fluxes measured by the Eddy Covariance (EC) method at urban flux towers in the URBANFLUXES project and good agreement was observed (Feigenwinter et al., 2012).			G
Software QA	There is no software currently available. The method is implemented by FORTH in Python. No need for QA.			G
Legal	IPR belongs to FORTH.			G

Objective Hysteresis Model (OHM)

	Objective Hysteresis Model (OHM)	R	A	G
Summary, Inputs, outputs	The Objective Hysteresis Model (OHM) is a well-established and robust method using net radiation Q* and a set of Land Use/Land Cover dependent coefficients for the calculation of heat storage flux (ΔQs). Inputs: Landcover fractions, land cover coefficients, Net all-wave radiation			
Key Reference(s)	S. Grimmond, T.R. Oke 1991. Heat Storage in Urban Areas: Local-Scale Observations and Evaluation of a Simple Model. Joyrnal of Applied Meteorology, vol 38, pp 922-940. doi: https://doi.org/10.1175/1520-0450(1999)038<0922:HSIUAL>2.0.CO;2 H.C. Ward, S. Kotthaus, L. Järvi, C.S.B. Grimmond. Surface Urban Energy and Water Balance Scheme (SUEWS): development and evaluation at two UK sites Urban Climate, 18 (2016), pp. 1-32, 10.1016/j.uclim.2016.05.001			
Technical fitness for purpose	The model is suitable for HARMONIA since it estimated heat storage flux at neighbourhood scale.			
Development Level	The model is fully developed, but needs to be adapted for the cities/areas in question. Parameterization and adjustment is necessary for a specific city/area since several local conditions, alter the behaviour of the model.			
Model Validation	The model has been validated in several different research projects. The model parameterized according to H.C.Ward et al. 2016.			
Software QA	There is no software currently available. The method is implemented by FORTH in Python. No need for QA.			
Legal	IPR belongs to FORTH.			

7.3. Humidity and Temperature Anomalies & complex scenarios

Humidity and Temperature Anomalies & complex scenarios Data

MODIS Burned Area Specification

	MODIS Burned Area	R	A	G
Summary	The Moderate Resolution Imaging Spectroradiometer (MODIS) MCD64A1 Version 6.1 contains burning and quality information on a per-pixel basis. Produced from both the Terra and Aqua MODIS-derived daily surface reflectance inputs, the algorithm analyses the daily surface reflectance dynamics to locate rapid changes, and uses that information to detect the approximate date of burning, mapping the spatial extent of recent fires only.			
Technical characteristics	Burned areas are characterized by deposits of charcoal and ash, removal of vegetation, and alteration of the vegetation structure (Pereira et al.,1997, Roy et al., 1999). The MODIS burned area mapping algorithm takes advantage of these spectral, temporal, and structural changes. It detects the approximate date of burning at a spatial resolution of 500 m by locating the occurrence of rapid changes in daily surface reflectance time series data. - Burned Date 500m – burn day of year - Burned Date Uncertainty 500m – estimated uncertainty in burn day - QA 500m – Quality Assurance indicators MODIS burned area products are suitable to map areas affected by fire even within urban environments. The latest collection (V061) has significantly better detection of small burns, a modest reduction in burn-date temporal uncertainty, and a large reduction in the extent of unmapped areas compared to the previous collection.			
Availab-ility	Combined products, Terra + Aqua, are available from July 2001 making them ideal for long term analysis. The methodology to derive burned areas has been adapted to create a similar product but using VIIRS and OLCI data. This will ensure continuity once the Terra and Aqua platforms are deactivated.			
Technical Access	MODIS products are freely available from a wide range of data providers. The Land Processes Distributed Active Archive Center (LP DAAC) is the most common data portal for MODIS data and provides access using different interfaces from automated APIs to data servers. Additionally, MODIS data can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	MODIS data has been validated and the BRDF product sits at stage 3 on the CEOS LPV validation hierarchy. Part of the validation involves comparison with *in situ* albedo observations such as AmeriFlux, NEON and SURFRAD. Liu et al. (2017) used this technique to validate the MODIS BRDF C006 snow-free product (using only full inversion retrievals) and found an overall 0.0195 RMSE with a -0.0038 bias.			
Legal	Open access. NASA/USGS give freedom to use, modify and distribute basic and adapted products.			

MODIS Active Fires Specification

	MODIS Active Fires	R	A	G
Summary	The Moderate Resolution Imaging Spectroradiometer (MODIS) MOD14 Version 6.1 active fire product detects fires in 1-km pixels that are burning at the time of overpass under relatively cloud-free conditions using a contextual algorithm.			
Technical characteristics	The most basic fire product in which active fires and other thermal anomalies, such as volcanoes, are identified. The Level 2 product is defined in the MODIS orbit geometry covering an area of approximately 2340 × 2030 km in the along-scan and along-track directions, respectively. It is used to generate all of the higher-level fire products, and contains the following components: - An active fire mask that flags fires and other relevant pixels - A pixel-level quality assurance (QA) image that includes 19 bits of QA information about each pixel - A fire-pixel table which provides 27 separate pieces of radiometric and internal-algorithm information about each fire pixel detected within a granule MODIS active fire products are suitable to identify fires even within urban environments. However, urban areas are particularly a source of thermal anomalies (factories, large urban developments, landfills, tec) , therefore not all identifications will be wildfires, additional masking might be required to avoid false alarms.			
Availab-ility	Combined products, Terra + Aqua, are available from July 2001 making them ideal for long term analysis. The methodology to derive burned areas has been adapted to create a similar product but using VIIRS and OLCI data. This will ensure continuity once the Terra and Aqua platforms are deactivated.			
Technical Access	MODIS products are freely available from a wide range of data providers. The Land Processes Distributed Active Archive Center (LP DAAC) is the most common data portal for MODIS data and provides access using different interfaces from automated APIs to data servers. Additionally, MODIS data can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	MODIS data has been validated and the BRDF product sits at stage 3 on the CEOS LPV validation hierarchy. Part of the validation involves comparison with *in situ* albedo observations such as AmeriFlux, NEON and SURFRAD. Liu et al. (2017) used this technique to validate the MODIS BRDF C006 snow-free product (using only full inversion retrievals) and found an overall 0.0195 RMSE with a -0.0038 bias.			
Legal	Open access. NASA/USGS give freedom to use, modify and distribute basic and adapted products.			

MODIS BRDF Data Specification

	MODIS BRDF	R	A	G
Summary	The Moderate Resolution Imaging Spectroradiometer (MODIS) MCD43A1 Version 6.1 Bidirectional Reflectance Distribution Function and Albedo (BRDF/Albedo) Model Parameters dataset is produced daily using 16 days of Terra and Aqua MODIS data at 500 metre resolution. The BRDF model parameters are used to derive surface reflectance under any feasible acquisition and illuminations conditions for all MODIOS spectral bands in the visible, near-infrared and shortwave infrared.			
Technical characteristics	Daily MODIS surface reflectance data are used to invert a BRDF model. The daily observations are temporally weighted to the ninth day of the retrieval period. The product coded as MCD43A1, it provides the three model weighting parameters (isotropic, volumetric, and geometric) used to derive the Albedo (MCD43A3) and Nadir BRDF-Adjusted Reflectance (NBAR) (MCD43A4) products. Relevant broad bands: - 3 x 500m broad bands: visible (0.3-0.7µm), infrared (0.7-5.0µm), and the full shortwave (0.3-5.0µm). - Radiometric resolution is 12bit MODIS BRDF model parameters are suited for characterising the land surface on a daily basis and help creating the surface reflectance expectation required by the atmospheric correction scheme used within the Urban Air Quality application.			
Availab-ility	Combined products, Terra + Aqua, are available from July 2001 making them ideal for long term analysis. However, in this application only data from 2017 will be used to match the Sentinel-2 era. The methodology to derive the BRDF model parameters has been adapted to create a similar product but using VIIRS data. This will ensure continuity once the Terra and Aqua platforms are deactivated.			
Technical Access	MODIS products are freely available from a wide range of data providers. The Land Processes Distributed Active Archive Center (LP DAAC) is the most common data portal for MODIS data and provides access using different interfaces from automated APIs to data servers. Additionally, MODIS data can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	MODIS data has been validated and the BRDF product sits at stage 3 on the CEOS LPV validation hierarchy. Part of the validation involves comparison with *in situ* albedo observations such as AmeriFlux, NEON and SURFRAD. Liu et al. (2017) used this technique to validate the MODIS BRDF C006 snow-free product (using only full inversion retrievals) and found an overall 0.0195 RMSE with a -0.0038 bias.			
Legal	Open access. NASA/USGS give freedom to use, modify and distribute basic and adapted products.			

C3S ERA-5

	C3S ERA5	R	A	G
Summary	ERA5 provides hourly estimates of a large number of atmospheric, land and oceanic climate variables. The data cover the Earth on a 30km grid and resolve the atmosphere using 137 levels from the surface up to a height of 80km. ERA5 includes information about uncertainties for all variables at reduced spatial and temporal resolutions.			
Technical characteristics	Quality-assured monthly updates of ERA5 (1979 to present) are published within 3 months of real time. The main variables that will be used within this application are: - 2m temperature: Temperature of air at 2m above the surface of land, sea or in-land waters. Temperature measured in kelvin can be converted to degrees Celsius (°C) by subtracting 273.15. - 2m dewpoint temperature: Temperature to which the air, at 2 metres above the surface of the Earth, would have to be cooled for saturation to occur. It is a measure of the humidity of the air. Combined with temperature and pressure, it can be used to calculate the relative humidity. - 10m u- and v-components of wind: eastward and northward components of the 10m wind. ERA5 variables are very well characterised, provide a long term data record. Nevertheless, given the ~30km grid cell size, they are not necessarily suitable for local scale analysis.			
Availab-ility	ERA5 products are available from 1979 to present. Preliminary daily updates of the dataset are available to users within 5 days of real time. Any near real-time applications will need access to the ECMWF operational archive to obtain data with a ~1 day lag.			
Technical Access	All ERA5 data products are freely available at the Climate Data Store (CDS). Additionally, some ERA5 data variables can be acquired in cloud environments and commercial platforms such as Google Earth Engine and Amazon Web Services.			
QA	ERA5 has been assessed in multiple peer-reviewed publications. (Hersbach et al., 2020) performed a first-guess departure to assess against observations just prior to their assimilation finding a good agreement for different conditions.			
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

Humidity and Temperature Anomalies & complex scenarios Models

Fire Weather Index (FWI)

	Fire Weather Index (FWI)	R	A	G
Summary, Inputs, outputs	The FWI system is able to model fuel moisture response to atmospheric forcing. Tree fuel beds are considered at increasing depths. The deeper and more compact the fuel is, the slower its response to the atmospheric forcings. These moisture codes are then combined to derive fire behaviour indices in terms of rate of spread and intensity. The FWI system is used to estimate fire danger.			
Key Reference(s)	- Vitolo, Claudia; Di Giuseppe, Francesca; Barnard, Christopher; Coughlan, Ruth; San-Miguel-Ayanz, Jesus ; Libertá, Giorgio ; Krzeminski, B. ERA5-based global meteorological wildfire danger maps. Nat. Sci. Data 1–11 (2020). https://doi.org/10.1038/s41597-020-0554-z - Van Wagner, C. E. et al (1974). Structure of the Canadian forest fire weather index, vol. 1333. Environment Canada, Forestry Service.			
Technical fitness for purpose	As FWI quantifies a potential danger, it is not an observable quantity but it can be validated against other existing reanalysis products. Compare against EO derived burned area and active fires from the Global Fire Emission Database (GFED) v4, which identify different fire regimes, the average bias goes from −5 to +5, while the anomaly correlation ranges between 0.52 to 0.87.			G
Development Level	The source code to reproduce the results obtained by Vitolo (2020) is openly available on a public repository on GitHub: https://github.com/cvitolo/paper_gef_era5)			G
Model Validation	Good technical validation results have been demonstrated in Vitolo (2020). Products for previous years are available for assessment.		A	
Software QA	Software is publicly available but not formally maintained. Limited user base.	R		
Legal	Model and source code are publicly available and there are no legal limitations to its use and reproduction.			G

7.4. Urban Flash Flooding

Urban flash flooding Data

	EO-Based Flood Inundation Obs. (Validation)	R	A	G
Summary	Usage of Earth-Observation data could prove to be a valuable source of validation information for urban flood modelling. Each municipality requesting urban flood modelling services should provide, at a minimum, dates of significant past flooding events in-order for FMI to check which satellite data products might be available for producing flood inundation area/depth maps.			
Technical characteristics	At this point in time, it is impossible to determine what EO-data products will be used, if any.			
Availab-ility	TBD			
Technical Access	TBD			
QA	TBD			
Legal	TBD			

	Drainage Flow Observations (Validation)	R	A	G
Summary	Each municipality requesting an urban flood modelling service should try to provide or point to a place where FMI can inquire about city wide drainage system flow measurements for Itzï - distributed hydrologic and hydraulic modelling system validation purposes. Without validation it is next to impossible to assess the predictive capability of the system and/or calibrate model.			
Technical characteristics	Distributed hydrologic and hydraulic modelling system validation purposes. Without validation it is next to impossible to assess the predictive capability of the system and/or calibrate model			
Availab-ility	TBD			
Technical Access	TBD			
QA	TBD			
Legal	TBD			

	Meteorological Observations (Calibration / Validation)	R	A	G
Summary	Each municipality requesting an urban flood modelling service should try to provide or point to a place where FMI can inquire about meteorological in situ. Without validation it is next to impossible to assess the quality of the proposed meteorological bias corrections and synthetic weather radar image products.			
Technical characteristics	In situ observations for ecPoint-calibrate synthetic radar image production validation purposes			
Availab-ility	TBD			
Technical Access	TBD			
QA	TBD			
Legal	TBD			

	EU DEM	R	A	G
Summary	A Digital Elevation Model (DEM) is a numerical representation of the Earth's surface that contains actual height points representing the topography, as well as the method to calculate elevations between the height points. Typically, a DEM is stored in a data system as a regular grid or a triangulated irregular network (TIN). EU-DEM v1.0 and v1.1 is a digital surface model (DSM) of EEA39 countries representing the first surface as illuminated by the sensors. It is a hybrid product based on SRTM and ASTER GDEM data fused by a weighted averaging approach. A high-resolution DEM in this context is used to calculate surface runoff flow directions and potential flow velocity fields for the urban flooding hydrologic model.			
Technical characteristics	Digital Surface Model (DSM) is a concept that has become common due to widespread use of airborne laser scanning. It means a model that represents the highest elevation of the terrain. Thus, a DSM represents the Earth's surface only in open areas, while in other regions the model follows the forest canopy and the roofs of buildings. A spatial resolution of 25 m may not however be sufficient for deriving detailed urban flow direction and potential flow velocity estimates.		(orange)	
Availab -ility	Static (in principle) with only very slow changes in time - reference year is 2017. Product will be maintained but specific future updates are not certain.			(green)
Technical Access	Open download from Copernicus Land Service			(green)
QA	The statistical validation of EU-DEM v1.0 shows a relatively unbiased (-0.56 meters) overall vertical accuracy of 2.9 meters RMSE, which is fully within the contractual specification of 7m RMSE (European Commission 2009). - https://land.copernicus.eu/user-corner/technical-library/eu-dem-2013-report-on-the-results-of-the-statistical-validation The following corrections and improvements have been implemented in EU-DEM v1.1: - Systematic correction of geo-positioning issues (found and corrected for Malta and Lampedusa islands); - Bias adjustment with ICESat; - Screening and removal of artefacts, including the presence of blunders (i.e. negative or positive anomalies); more than 75.000 artefacts have been detected and corrected; - Consistency with the upgraded version of EU-Hydro, in order to produce a better river network topology. EU-DEM v1.1 has not been validated yet		(orange)	
Legal	Access to data is based on a principle of full, open and free access as established by the Copernicus data and information policy Regulation (EU) No 1159/2013 of 12 July 2013.			(green)

	Urban Atlas	R	A	G
Summary	The Urban Atlas provides pan-European comparable land cover and land use data for Functional Urban Areas (FUA).. Each Urban Atlas product is generated over the city and its surroundings, according to the Functional Urban Area (FUA) defined by the implementation of the approach developed by the DG Regional and Urban Policy (REGIO) of the European Commission. The Urban Atlas dataset allows for the creation of a harmonized set of stormwater runoff potential characteristics parameters for the urban flooding model for all the selected cities in the HARMONIA project.			
Technical characteristics	The dataset appears to fit the project's goals, however more detailed data may provide more accurate simulation results, or it may turn-out that more detailed data would essentially disappear into noise caused by uncertainties in other input/forcing data. Since urban environments are constantly evolving, this dataset must also be constantly updated in-order-to ensure that the urban flooding model's static parameters remain up to date.			
Availab-ility	The Urban Atlas is available for the 2006, 2012 and 2018 reference years over nearly 800 cities with more than 50,000 inhabitants distributed among EU, EFTA and West Balkan countries plus United Kingdom and Turkey. Future periodic updates should be assured but may not always reflect rapidly developing urban areas.			
Technical Access	Direct download from Copernicus Land Service			
QA	Extensive validation reports have been undertaken: • https://land.copernicus.eu/user-corner/technical-library/building-heights-2012-validation-report • https://land.copernicus.eu/user-corner/technical-library/urban-atlas-2018-validation-report			
Legal	Access to data is based on a principle of full, open and free access as established by the Copernicus data and information policy Regulation (EU) No 1159/2013 of 12 July 2013.			

	HDWS Soils	R	A	G
Summary	The Harmonized World Soil Database is a 30 arc-second raster database with over 15 000 different soil mapping units that combines existing regional and national updates of soil information worldwide (SOTER, ESD, Soil Map of China, WISE) with the information contained within the 1:5 000 000 scale FAO-UNESCO Soil Map of the World (FAO, 1971-1981). The resulting raster database consists of 21600 rows and 43200 columns, which are linked to harmonized soil property data. The use of a standardized structure allows for the linkage of the attribute data with the raster map to display or query the composition in terms of soil units and the characterization of selected soil parameters (organic Carbon, pH, water storage capacity, soil depth, cation exchange capacity of the soil and the clay fraction, total exchangeable nutrients, lime and gypsum contents, sodium exchange percentage, salinity, textural class and granulometry			
Technical characteristics	The generalized nature and low resolution of the product may cause issues when developing a hydrologic model for relatively small areas, where more detail soil information would be preferable. Harmonized soil texture data is typically however not available. If new more reliable data sources are found, the project may switch to another data source. In urban hydrology soils typically do not however play a very significant role in model setup and/or parametrization since most of the modelling domain is in an urban environment and other factors play a more important role in the hydrology regime of cities.	■ (red)		
Availab -ility	The database is a single reference product with components of varying dates.		■ (amber)	
Technical Access	Downloadable from the FAO Soils Portal https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/harmonized-world-soil-database-v12/en/			■ (green)
QA	Extensive product validation is lacking due to scarcity of reference data.	■ (red)		
Legal	All rights reserved. No part of this Harmonized World Soil Database may be reproduced, stored in a retrieval system or transmitted by any means for resale or other commercial purposes without written permission of the copyright holders. Reproduction and dissemination of material in this information product for educational or other non-commercial purposes are authorized without any prior written permission from the copyright holders provided the source is fully acknowledged. Full acknowledgement and referencing of all sources must be included in any documentation using any of the material contained in the Harmonized World Soil Database.		■ (amber)	

Urban flash flooding Models

	Itzï - distributed hydrologic and hydraulic model	R	A	G
Summary, Inputs, outputs	Itzï is a dynamic, fully distributed hydrologic and hydraulic model that simulates: • Hydrological processes like rainfall and infiltration • 2D surface flows on a raster grid using simplified shallow water equations. • 1D drainage flow via the SWMM model • Bi-directional coupling between the drainage and the surface.			
Key Reference(s)	Courty, L. G., Soriano-Monzalvo, J. C., & Pedrozo-Acuña, A. (2019). Evaluation of open-access global digital elevation models (AW3D30, SRTM and ASTER) for flood modelling purposes. Journal of Flood Risk Management, 1–14. https://doi.org/10.1111/jfr3.12550 Courty, L. G., Rico-Ramirez, M. Á., & Pedrozo-Acuña, A. (2018). The Significance of the Spatial Variability of Rainfall on the Numerical Simulation of Urban Floods. Water, 10(2), 207. https://doi.org/10.3390/W10020207 Courty, L. G., Pedrozo-Acuña, A., & Bates, P. D. (2017). Itzï (version 17.1): an open-source, distributed GIS model for dynamic flood simulation. Geoscientific Model Development, 10(4), 1835–1847. http://doi.org/10.5194/gmd-10-1835-2017			
Technical fitness for purpose	FMI has already demonstrated the technical feasibility of application of this model in similar projects (e.g., CITYZER) for the city of Helsinki, Finland. The modelling system contains all the required components to achieve the aims of task WP 5 Urban Flooding pilots in any given area if all the required input datasets are made available.			
Development Level	Code library available on Bitbucket (itzi-model / itzi — Bitbucket)			
Model Validation	Urban flooding models in general are challenging to validate often due to the lack of availability of fit-for-purpose validation datasets. There are however numerous case studies where the modelling system has seemingly been utilised effectively to simulate urban flooding scenarios.			
Software QA	Software is publicly available and maintained through contributions from the user base.			
Legal	Itzï is released under the GNU GPL license. It is free to use for any purposes.			

	ecPoint-Calibrate	R	A	G
Summary, Inputs, outputs	ecPoint-Calibrate is a conditional verification tool to compare numerical weather prediction (NWP) model outputs against point observations and, in this way, anticipate sub-grid variability and identify biases at grid scale. It's decision trees can used to downscale coarse models to better point forecasts. It provides a dynamic and user-friendly environment to post-process NWP model parameters (such as precipitation, wind, temperature, etc.) and produce probabilistic products for geographical locations (everywhere in the world, and up to seasonal range forecasts).			
Key Reference(s)	Hewson, T.D., Pillosu, F.M. A low-cost post-processing technique improves weather forecasts around the world. Commun Earth Environ 2, 132 (2021). https://doi.org/10.1038/s43247-021-00185-9			
Technical fitness for purpose	Preliminary studies into statistical post-processing for ensemble forecasts show encouraging results: https://www.nature.com/articles/s43247-021-00185-9			
Development Level	Early release versions available, currently 0.29.0. Application to seasonal forecasts is a novel step, but as the underlying model is the same ECMWF code, there are no barriers to doing it.			
Model Validation	When applying this post-processing method, forecast skill was shown to improve substantially; for extreme rainfall, useful forecasts were extended 5 days ahead, compared to less than 1 day without post-processing.			
Software QA	Software is publicly available but not formally maintained. Limited user base.			
Legal	Open source, code library available on GitHub (https://github.com/esowc/ecPoint-Calibrate)			

7.5. Geohazards

Geohazards Data

Sentinel-1 Data Specification

	Sentinel-1	R	A	G
Summary	The Copernicus Sentinel-1 mission comprises a constellation of two platforms in the same sun-synchronous orbit. S1 carries a Synthetic Aperture Radar (SAR) sensor able to acquire radar images with a variable swath and a revisit time of 6 days with both satellites (12 days with one satellite). The Sentinel-1 mission includes C-band imaging operating in four exclusive imaging modes with different resolution (down to 5 m) and coverage (up to 400 km). It provides dual polarisation capability, very short revisit times and rapid product delivery. For each observation, precise measurements of spacecraft position and attitude are available.			
Technical characteri-stics	Sentinel-1 carries a single C-band SAR instrument operating at a centre frequency of 5.405 GHz. It is equipped with a right-looking active phased array antenna providing fast scanning in elevation and azimuth, data storage capacity of 1 410 Gb and 520 Mbit/s X-band downlink capacity. The C-SAR instrument supports operation in dual polarisation (HH+HV, VV+VH) implemented through one transmit chain (switchable to H or V) and two parallel receive chains for H and V polarisation. Dual polarisation data is useful for land cover classification and sea-ice applications. Sentinel-1 operates in four exclusive acquisition modes: ∉ Stripmap (SM) ∉ Interferometric Wide swath (IW) ∉ Extra-Wide swath (EW) ∉ Wave mode (WV). We will use the Interferometric Wide swath (IW) mode, that is the default mode on land, characterised by a large swath width (250 km) with a moderate geometric resolution (5 m by 20 m). The IW mode returns three sub-swaths using Terrain Observation with Progressive Scans SAR (TOPSAR). The characteristics of such a data are well suited for identification of landslides, subsidence and man-made structure instabilities both in urban and rural areas.			
Availability	Sentinel-1A: data available from 2014 Sentinel-1B: data available from mid 2016 Good support for long term sustainability and a well-established historical database.			

Technical Access	Sentinel-1 products available for users (either generated by the ground segment or by the Sentinel-1 Toolbox) are Level-1, i.e. the Single Look Complex (SLC) data. The Copernicus Open Access Hub provides complete, free and open access to Sentinel-1 data, but only of recent acquisition (about 2 months). In order to overcome this limitation, Sentinel-1 data products are also available in the Copernicus Data and Information Access Service (DIAS) cloud environments and in the NASA platform Alaska Satellite Facility (ASF) service via a dedicated web GIS.			
QA	Precise orbit files are periodically released by ESA to improve the accuracy of the satellite position along its orbit. The NTC orbital product (AUX_POEORB) is generated with a timeliness of 20 days and covers 26 hours, from 1 hour before the start of day N-20 until 1 hour after the end of day N-20 (being N the current day) providing an accuracy of 5 cm. The overall system is under constant quality monitoring by ESA.			
Legal	Open access. Copernicus license gives freedom to use, modify and distribute basic and adapted products.			

DEM Data Specification

	Digital Elevation Model (DEM)	R	A	G
Summary	A Digital Elevation Model (DEM) is needed to remove the topography contribution during the InSAR processing chain, other than geocoding the final InSAR outputs in the desired reference system. Since the last decades several DEM have been produced from different space missions worldwide using different sensors. For our purpose, we will mainly adopt the Shuttle Radar Topography Mission (SRTM). The SRTM project produced digital topographic data for 80% of the Earth's land surface, with data points located every 1-arc-second (approximately 30 meters) on a latitude/longitude grid.			
Technical characteristics	The Shuttle Radar Topography Mission (SRTM) was flown aboard the space shuttle *Endeavour* February 11-22, 2000. This mission used single-pass interferometry, which acquired two signals at the same time by using two different radar antennas. An antenna located on board the space shuttle collected one data set and the other data set was collected by an antenna located at the end of a 60-meter mast that extended from the shuttle. Differences between the two signals allowed for the calculation of surface elevation. The SRTM DEM shows a ground pixel resolution of 30 (1 arc-second) or 90 meters (3 arc-sec). Each file or cell contains a matrix of vertical elevation values spaced at regular horizontal intervals measured in geographic latitude and longitude units. File size is approximately 25 MB for 1-arc-second data files and approximately 3 MB for 3-arc-second data files. Band interleaved by line (BIL) is a binary raster format with an accompanying header file which describes the layout and formatting of the file. File size is approximately 7 MB for 1-arc-second data files and approximately 1 MB for 3- arc-second data files. Georeferenced Tagged Image File Format (GeoTIFF) is a TIFF file with embedded geographic information. This is a standard image format for GIS applications. File size is approximately 25 MB for 1-arc-second data files and approximately 3 MB for 3-arc-second data files.			
Availability	SRTM data are freely available and distributed. Each data file covers a one-degree-of-latitude by one-degree-of-longitude block of Earth's surface. These files are in "raw" format (no headers and not compressed), 16-bit signed integers, elevation measured in meters above sea level, in a "geographic" (latitude and longitude array) projection, with data voids indicated by -32768.			
Technical Access	USGS EarthExplorer can be used to search, preview, and download Shuttle Radar Topography Mission (SRTM) 1/3 Arc-Second Global data.			
QA	The accuracy in elevation is about 3 to 6 meters. The SRTM DEM covers all land areas between 60º north and 56° south latitude.			

Legal	Open access. USGS gives freedom to use, modify and distribute basic and adapted products.			

GACOS Data Specification

	Generic Atmospheric Correction Online Service for InSAR (GACOS)	R	A	G
Summary	GACOS utilises the Iterative Tropospheric Decomposition (ITD) model to separate stratified and turbulent signals from tropospheric total delays, and generate high spatial resolution zenith total delay maps to be used for correcting InSAR measurements and other applications. GACOS has the following key features: (i) globally available; (ii) operational in a near real time mode; (iii) easy to implement; and (iv) users to be informed how the model performs and whether the correction is recommended.			
Technical characteristics	Datasets used in GACOS include: • High Resolution ECMWF weather model at 0.1-degree and 6-hour resolutions; • SRTM DEM (90m,S60-N60); • ASTER GDEM (90m,N60-N83,S60-S83); GACOS tropospheric delay maps are given in a grid binary format (4-byte float little endian, naming convention YYYYMMDD.ztd). A ReadMe file is provided to demonstrate how to use GACOS tropospheric correction maps. For any further details see the paper at the following link.			
Availability	GACOS: data available from 2019. Good support for long term sustainability. GACOS is supported by COMET, Newcastle University, through the centre for seismic, volcanic and tectonic observation and modeling (comet, Ref.: come30001) and LICs (Ref. Ne / k010794 / 1) and esa-most dragon-4 (Ref. 32244).			
Technical Access	GACOS products available for users are in a grid Geotiff format. The GACOS online website provides complete, free and open access to the GACOS data.			
QA	The performance of the InSAR tropospheric corrections is assessed by considering both the phase standard deviation and the agreement with GPS-measured displacements. Further details can be found at this link.			
Legal	Open access. Users are free to share, copy and redistribute the data in any medium or format, provided that they give appropriate credit to the author (e.g., citing the above-mentioned papers).			

Geohazards Models

	Persistent Scatterer Interferometry	R	A	G
Summary, Inputs, outputs	Persistent Scatterer Interferometry is by now a well-consolidated technique applied to large stacks of SAR (Synthetic Aperture Radar) acquisitions able to retrieve mean ground velocity and time series of deformation, at each image date, of point targets observed by the SAR sensor. The method is based on the selection of pixels within the SAR images that have a stable amplitude of the backscattered electromagnetic signal measured by SAR, i.e. the Persistent Scatterers (PS). The processing starts with the generation of N-1 interferograms, being N the number of available SAR images, and at the end of the processing chain the electromagnetic phase information, and its evolution in time, is extracted for each PS. By applying spatio-temporal filters the PS algorithm can estimate atmospheric delay and topographic residuals that are then removed from the displacement time series. The final outcomes are: 1) the map of the mean velocity of each PS, in mm/yr, and 2) for each PS the time series of deformation in mm. It is worth noting that SAR sensors can measure such movements along its own Line Of Sight (LOS) depending on the orbit of the satellite (ascending or descending). Furthermore, the combination of the two orbits allows estimating the vertical and horizontal (east-west) component of the detected deformation.			
Key Reference(s)	- Ferretti, A., Prati, C., Rocca, F., 2000. Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry. IEEE TGRS 38 (5),2202–2212. - Ferretti, A., Prati, C., Rocca, F., 2001. Permanent scatterers in SAR interferometry. IEEE TGRS 39 (1), 8–20.			
Technical fitness for purpose	In the last 3 decades, SAR interferometry (InSAR) has proven to be a particularly powerful technique for Earth geodesy. It has been applied for mapping topography and deformation at the Earth's surface, and its products are widely used in tectonics, seismology, geomorphology, volcanology, urban monitoring, with sub-centimeter accuracy.			
Development Level	The PS algorithm adopted in HARMONIA is the one developed in the SARScape interferometric module, which is part of the commercial software ENVI (Harris Geospatial Solutions).			
Model Validation	Model has been extensivelyvalidated by the huge range of literature and applications based on PS.			

Software QA	The software is maintained by SARmap SA. Periodical upgrades and releases are delivered with new functionalities and tools providing algorithm improvements, automation of the different processing steps, accuracy of the final products, and usability.			
Legal	The licensed software is supported by the INGV funding. Fully available for the HARMONIA applications.			

	Geotechnical approaches	R	A	G
Summary, Inputs, outputs	Several geotechnical approaches can be applied to investigate the observed ground displacements from the processing of SAR (Synthetic Aperture Radar) images. These approaches range from simplified analytical methods aimed at assessing the origin of urban subsidence or landslide movements, to complex numerical methods. The choice of the most suitable model to be applied to the deformation data will be evaluated with Task 4.4.			
Key Reference(s)				
Technical fitness for purpose	TBD			
Development Level	TBD			
Model Validation	TBD			
Software QA	TBD			
Legal	TBD			