COMPASS

Scoping report of data and models for impact attribution

Milestone 5

30th September 2024

Title	Scoping report of data and models for impact attribution
Lead Beneficiary	PIK
Lead Author(s)	Dominik Paprotny (PIK)
Contributors	Anaïs Couasnon (Deltares)
Milestone number	5
Work Package	WP3 Enabling impact attribution by multi-scale modeling of exposure and vulnerability
Submission date	30/09/2024

Dissemination Level

PU	Public — fully open (automatically posted online)	X
SEN	Sensitive — limited under the conditions of the Grant Agreement	
CI	EU classified —Restreint-UE/EU-Restricted, Confidentiel-UE/EU-Confidential, Secret-UE/EU-Secret under Decision 2015/444	

Version History

Date	Version	Contributors	Comments
22/07/2024	0.1	DP	First rough draft
08/08/2024	0.2	DP	Draft exposure dataset description
26/09/2024	1.0	DP, AC	First complete version

Citation

Paprotny, D., Couasnon, A. (2024): Scoping report of data and models for impact attribution. Horizon Europe project COMPASS. Milestone MS5.

Funded by the European Union

The COMPASS project has received funding from the European Union's HORIZON Research and Innovation Actions Programme under Grant Agreement No. 101135481

Disclaimer

Executive summary

This report collects information on available data, models, methods, and tools on exposure, vulnerability, and impacts that have prospective use in the project, either for use cases (Tasks 4.1 and 4.3) or the overall impact attribution framework (Task 3.4). It also contributes preliminary data and methods for other tasks in WP3, including exposure datasets at multiple scales (D3.1) and vulnerability models for multiple hazards (D3.3). It further considers the requirements of a prospective operational attribution service utilizing a methodology developed in the project. In general, we list only datasets and models which are publicly accessible without any restrictions, unless there are no such resources available for a particular aspect of exposure or vulnerability.

Table of Contents

1. Exposure data

1.1. National-level datasets

Several datasets provide national-level data for historical and future periods. They vary in coverage, often omitting certain small countries or territories. Some datasets report nationally reported data, while others adjust and gap-fill deficient or missing national data and produce more harmonized statistics. No single source provides all variables of interest (defined in Task 3.1): land use, population, gross domestic product (GDP) and fixed assets, and has full coverage of all countries and territories. The table below includes only datasets with global or near-global coverage. It excludes research publications, which usually are not maintained and mostly fill the gaps in knowledge on historical data, especially before 1950. In this and further tables in the document, the bolded title of each dataset/model links to the repository from which it is available, while the author indicated below links to the primary document describing of the dataset/model.

Dataset	Description
World Population Prospects United Nations	**Coverage:** all countries in the world divided into 237 geographies, with some smallest dependent territories aggregated. **Timeframe:** 1950-2023 and projections for 2024-2100 **Variables:** population by sex and age, fertility, mortality, migration balance and indicators **Usefulness:** this dataset harmonizes and consistently recomputes historical data and provides not only deterministic projections, but also probabilistic Bayesian projections. With full global coverage, it is the most useful population dataset available. **Update schedule:** biannually (current version: '2024 revision', July 2024)
Demographic Yearbook System United Nations	**Coverage:** all countries in the world, but availability varies per variable **Timeframe:** 1948-2023 **Variables:** large number of variables describing detailed population characteristics **Usefulness:** this data set provides raw nationally produced demographic data from censuses, registers and estimates, with coverage varying enormously between countries. It could be used to gap-fill historical information for certain countries or provide socioeconomic characteristics. **Update schedule:** twice-yearly (current version: February 2024)
SSP Scenario Explorer IIASA	**Coverage:** population for 200 geographies, GDP for 186-192 geographies depending on projection source (OECD or IIASA) **Timeframe:** SSP projections for 2025-2100, with additional 'historical reference' population data for 1950-2020 and GDP data for 1980-2020 **Variables:** population by sex, age and education, GDP, urbanization level (not yet downloadable) **Usefulness:** Shared Socioeconomic Pathways are fundamental scenarios for climate change projections. Still, gaps in coverage and 5-yearly timestep requires some harmonization with historical data and gap-filling. The two GDP projections included in SSPs show something very different already in 2025; OECD projections provided are with historical data, so that they can be harmonized, but IIASA projections have no historical data for reference. **Update schedule:** irregular (current version: 3.1, July 2024)
National Accounts Main Aggregates Database United Nations	**Coverage:** 220 geographies, including all 193 independent countries and 8 predecessor historical countries **Timeframe:** 1970-2022 **Variables:** GDP by industry, GDP expenditure decomposition, population, deflators, exchange rates **Usefulness:** this dataset harmonizes and consistently recomputes national data to provide providing the best coverage of GDP data worldwide **Update schedule:** annual (current version: December 2023)

Main Aggregates and Detailed Tables United Nations	**Coverage:** almost all countries, but availability varies per variable **Timeframe:** 1946-2021 **Variables:** GDP by industry and institutional sector, GDP expenditure and income decomposition, fixed asset stock, employment **Usefulness:** this data set provides raw nationally produced national accounts data, with coverage varying enormously between countries. Could be used to gap-fill historical information for certain countries **Update schedule:** annual (current version: '2022', actually released 2024)
World Economic Outlook International Monetary Fund	**Coverage:** 196 countries and territories **Timeframe:** 1980-2029 **Variables:** GDP, population and several other economic variables **Usefulness:** the dataset provides regularly updated most recent national economic data and produces short-term projections up to 5 years ahead, making it best dataset on current global economic conditions **Update schedule:** twice-yearly (current version: April 2024)
World Development Indicators World Bank	**Coverage:** 217 geographies, but availability varies per variable **Timeframe:** 1960-2023 **Variables:** population, GDP, agricultural land area, and hundreds of other variables **Usefulness:** comprehensive, multi-source and regularly updated collection of a vast array of socioeconomic statistics, though often drawing from other datasets in this table. Provides data for some small territories that are not covered by most other datasets. **Update schedule:** quarterly (current version: July 2024)
Penn World Table Groningen Growth and Development Centre	**Coverage:** 183 countries and territories **Timeframe:** 1950-2019 **Variables:** GDP, capital stock, and many other economic indicators **Usefulness:** the dataset harmonizes national GDP data and produces estimates of fixed asset stock and several other productivity-related indicators. Though timeseries availability varies, it is a useful resource on pre-1970 data and the only global source of timeseries of fixed asset stock **Update schedule:** irregular (current version: 10.01, January 2023)
Maddison Project Database Groningen Growth and Development Centre	**Coverage:** 169 countries and territories **Timeframe:** 1-2022 **Variables:** GDP, population **Usefulness:** the dataset collects, harmonizes and estimates long-run changes in GDP for most countries of the world. It is the primary global collection and source of pre-1950 GDP data **Update schedule:** irregular (current version: 2023)
FAOSTAT FAO	**Coverage:** 245 geographies, but availability varies per variable **Timeframe:** 1961-2023 **Variables:** land cover/use and hundreds of variables related to agriculture **Usefulness:** provides comprehensive data on land use and agricultural production for all countries **Update schedule:** continuous

1.2. Subnational-level datasets

Subnational data, i.e. data referring to administrative or statistical divisions of countries, are not available for all countries. Most comprehensive data is available for more developed countries, with limited and typically highly uncertain data for less developed countries. Outside countries covered by Eurostat and OECD databases, there is also a lack of a common classification scheme aimed at harmonizing subnational unit size

for comparability. Global subnational database ISO 3166-2 lacks such comparability and has no official map connected to it. Frequent changes to subnational divisions make it difficult for most countries to collect consistent long time series of subnational data.

Eurostat European Commission	**Coverage:** EU, EFTA and candidate countries **Timeframe:** mostly 2000-2022, which some longer series sometimes available for selected variables **Variables:** GDP, population, land use, and further detailed demographic and economic statistics, and many other topics **Usefulness:** Primary source of harmonized subnational data for most European countries, using a common classification scheme NUTS at three levels, with GDP and population available at the most detailed level 3. **Update schedule:** continuous, but individual variables typically updated once a year
OECD Data Explorer OECD	**Coverage:** OECD members and selected other countries **Timeframe:** mostly 2000-2022, which some longer series sometimes available for selected variables, availability varies by country **Variables:** GDP, population, and further detailed demographic and economic statistics, and many other topics **Usefulness:** Useful resource for countries not covered by Eurostat database **Update schedule:** continuous, but individual variables typically updated once a year
DOSE Wenz et al. (2023)	**Coverage:** 83 countries (1661 regions) **Timeframe:** 1953-2020, varies strongly by country **Variables:** GDP, total and by sector **Usefulness:** Compiles available subnational data including longer time series **Update schedule:** irregular (latest version v2, 2023)
SHDI Smits et al. (2019)	**Coverage:** 161 countries (1625 regions) **Timeframe:** 1990-2017 **Variables:** Human Development Index and its components (education, health, income) **Usefulness:** Could be used for qualitative vulnerability analyses **Update schedule:** unknown

1.3. Gridded datasets

Gridded datasets of population are mostly created by disaggregation of subnational data, except of purely census-based products. All gridded GDP or fixed asset datasets are disaggregations with considerable level of uncertainty. Very high-resolution products typically have short time span, and long-run reconstructions (spanning multiple decades or centuries) all have large uncertainty. Below we exclude pure land cover datasets already considered in Deliverable 1.1. We focus on the most useful dynamic datasets for attribution.

HANZE Paprotny et al. (2023a)	**Coverage:** 42 European countries and territories **Timeframe:** 1870-2020 **Variables:** GDP, fixed asset value, population, land use, soil sealing **Resolution:** 100 m **Usefulness:** Highest-resolution reconstruction of historical exposure, constructed by COMPASS project partner (PIK). Also contains a harmonized database of subnational statistics based on NUTS2010 level 3 classification **Update schedule:** irregular (latest version v2.0, published 2022)
HYDE Klein Goldewijk et al. (2017)	**Coverage:** global **Timeframe:** 10,000 BCE - 2023 **Variables:** population (total and urban/rural), land use **Resolution:** 1/12° = 5 arc minutes (~10 km)

	Usefulness: The only very long-run gridded reconstruction of past population and land use. The dataset suffers from many inaccuracies and is poorly documented in terms of data sources in particular but remains highly population in global climate research. **Update schedule:** irregular (current version v3.3, October 2023)
GHSL European Commission	**Coverage:** global **Timeframe:** 1975-2020 and projections for 2025-2030 **Variables:** population, build-up area **Resolution:** multiple available, down to 100 m (population) and 10 m (build-up area) **Usefulness:** The highest-resolution dataset of population covering both longer time series and short-term projections. **Update schedule:** annual (current version: 2023)
GEOSTAT European Commission	**Coverage:** EU, EFTA and some candidate countries **Timeframe:** 2006, 2011, 2018, 2021 **Variables:** population **Resolution:** 1 km **Usefulness:** Dataset with gridded population directly from geolocated census returns for 2011 and 2021, with additional estimates for 2006 and 2018. **Update schedule:** irregular (current version 2021, published June 2024)
LUH2 Hurtt et al. (2020)	**Coverage:** global **Timeframe:** 850-2015 and projections for 2015-2100 **Variables:** land use **Resolution:** 0.25° (~30 km) **Usefulness:** Primary land use dataset used in global climate research, heavily based on HYDE for the historical period **Update schedule:** irregular (latest version v2e, published 2019)
LitPop Eberenz et al. (2020)	**Coverage:** near-global **Timeframe:** 2014 **Variables:** fixed asset value **Resolution:** 30 arc seconds (~1 km) **Usefulness:** The only disaggregated fixed asset value dataset, based on nighttime lights **Update schedule:** unknown (latest version 2020)
WorldPop University of Southampton	**Coverage:** global **Timeframe:** 2000-2020 **Variables:** population and multiple demographic indicators **Resolution:** 100 m **Usefulness:** Alternative disaggregation of population to GHSL, with a smaller temporal coverage, but with additional demographic indicators **Update schedule:** irregular (latest version September 2020)
WFS-Evolution Marconcini et al. (2021)	**Coverage:** global **Timeframe:** 1985-2015 **Variables:** build-up area **Resolution:** 1 arc second (~30 m) **Usefulness:** satellite-derived build-up footprints in annual resolution, provides more temporal detail for settlement areas compared to GHSL. A 10-metre dataset for 2019 is also available. **Update schedule:** unknown

1.4. Object-level datasets

Object-level datasets are typically country-specific datasets, created by national or local authorities, with various levels of accessibility. Below, apart from building datasets with global coverage, we also include datasets and methods for disaggregating data designed to estimate missing information in other datasets.

Open Street Map OpenStreetMap Foundation	**Coverage:** global **Timeframe:** real-time **Variables:** building and land use polygons, building occupancy and other variables (where available) **Usefulness:** global crowd-sourced map with strongly varying completeness of building data and their attributes by local. Typically, the best resource of microscale exposure data, often integrates official datasets **Update schedule:** continuous
Open Building Map GFZ	**Coverage:** global **Timeframe:** near real-time **Variables:** building polygons, occupancy and other variables **Usefulness:** Derivate of OpenStreetMap, which gapfills buildings using satellite data, and then gapfills attributes. Estimated asset value is also available. The dataset is not downloadable, but accessible through personal communication with Danijel Schorlemmer **Update schedule:** continuous
Residential exposure Paprotny et al. (2020a)	**Coverage**: most European countries **Timeframe:** 2000-2020 **Variables:** residential building structure and contents replacement value **Usefulness:** dataset of average exposure of the residential sector per m^2 useful flood area, combined with a method of estimating height of buildings, where the information is missing. It was intended to be used with OSM data **Update schedule:** irregular (latest version: 2022)
Commercial exposure Paprotny et al. (2020b)	**Coverage:** most EU/EFTA countries **Timeframe:** n/a **Variables:** commercial building and equipment replacement value **Usefulness:** a method, rather than dataset, to disaggregate national fixed asset value statistics to individual buildings. It was intended to be used with OBM data using the occupancy classification scheme used there. **Update schedule:** n/a
Global Google-Microsoft Open Buildings Dataset	**Coverage:** near-global **Timeframe:** 2023 **Variables:** building polygons **Usefulness:** a dataset of buildings that could be helpful to fill gaps in OSM/OBM coverage but doesn't contain any information about occupancy of those. **Update schedule:** irregular (current version: v3, September 2023)

1.5. Multiscale approach to exposure

For COMPASS deliverable D3.1, multiple datasets will be combined with a flexible approach to downscale and upscale data according to requirements of the use cases or the future prospective operational service. The concept is to collect the best datasets available at national level for the historical period and harmonize them with projections under Shared Socioeconomic Pathways (SSPs). Then, the national data will be downscaled to target resolution using the best gridded datasets, down to 30" (~1 km) resolution. The intended coverage of the dataset is 1850-2100 to enable long-run attribution analyses in COMPASS and other research activities such as ISIMIP. The concept of combining the different datasets is shown in Fig. 1.

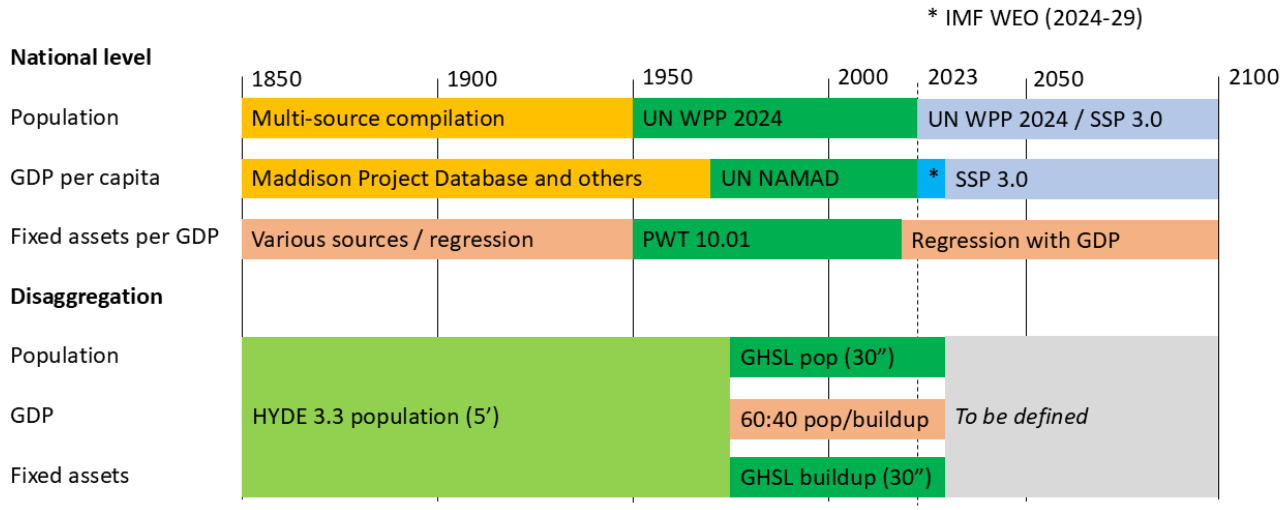

Figure 1. COMPASS multiscale model concept.

Population projections will utilize, apart from SSP datasets, the United Nations probabilistic projections to provide information on the uncertainty of future demographic developments. GDP projections will also go beyond SSPs and utilize the short-term (five years ahead) projections in the World Economic Outlook (WEO). IMF WEO will act as a bridge between historical data and long-term SSP projections. Data on fixed asset-to-GDP ratio is limited, therefore gap-filling of both historical and future trends will be done with statistical analysis. Economic data (GDP and fixed assets) will be by default measured as in SSPs (US dollars in 2017 prices and purchasing power parity), but a possibility of generating economic data in nominal value in given country is also envisioned.

GHSL dataset will be used to disaggregate population and fixed assets (according to build-up area) for the years 1975-2030. GDP will be disaggregated first using subnational GDP statistics, where available (primarily from HANZE, DOSE and OECD), and then using population and build-up area in 60:40 ratio, to approximate the relative contribution of labour and capital to generation of GDP. Extension back to 1975 will be done using coarser HYDE population data, assuming the distribution of all variables within each 5' cell be the same as in 1975. Extension to the future after 2030 is not yet defined. Several downscallings of future SSP pathways are available, but they are also rather coarse, and some were also found to show some unrealistic patterns for Europe (Steinhausen et al. 2022). As research in this area is active, it will be monitored for any new datasets.

2. Vulnerability models

2.1. Mesoscale models

Mesoscale models are those operating on an aggregated (typically gridded) level, where individual elements at risk are not present, but normally represented by patches of common land use type (e.g. urban, industrial, agricultural). Due to the aggregation, they need to be used with consistent exposure data to avoid biases resulting in assigning inaccurate economic value per area of given land use. From the list of flooded models, we exclude simple depth-damage functions, of which an enormous variety exists as highlighted in a review by Gerl et al. (2016).

2.1.1. Flooding

JRC Global Depth-Damage Functions Huizinga et al. (2017)	**Intended area of application:** global (based on data from many countries) **Damage predictors:** water depth, economic sector, continent **Publicly available:** yes (Excel tables) **Exposure data:** model provides exposure values (for year 2010) for different countries for both mesoscale and microscale applications. Usable also with other datasets. **Usefulness:** generic damage functions helpful when more detailed information (both exposure and models) is not available.
Flood mortality model Boyd et al. (2010)	**Intended area of application:** universal (based on US hurricane flooding data) **Damage predictors:** water depth **Publicly available:** yes (as equation in open-access paper by Jonkman et al. 2008) **Exposure data:** population (would work with any data) **Usefulness:** S-shaped depth-fatality function based on US hurricane impact data. It is useful in areas where only water depth is available
HANZE v2.3 Paprotny et al. (2024)	**Intended area of application:** Europe (based on data from many countries) **Damage predictors:** multiple, depends on impact type (fatalities, persons affected, economic loss). **Publicly available:** yes (Python code) **Exposure data:** intended for use with HANZE exposure data **Usefulness:** enables calculating vulnerability level for three types of impacts at NUTS3 regional level between 1950 and 2020, based on data from almost 40 countries. It includes socioeconomic conditions as predictors. It is intended for use in tandem with depth-damage functions from Huizinga et al. (2017) and Boyd et al. (2010), described above, as the model doesn't include water depth as predictors, but rather quantifies general societal vulnerability.
Flood mortality model Jonkman et al. (2008)	**Intended area of application:** universal (based on data from several countries) **Damage predictors:** water depth, water velocity **Publicly available:** yes (equations in open-access paper) **Exposure data:** population (would work with any data) **Usefulness:** allows more precise calculation of fatalities than to the use of water velocity to define different mortality zones.
Delft-FIAT Deltares	**Intended area of application:** universal **Damage predictors:** user-defined **Publicly available:** yes (Python code) **Exposure data:** user-defined **Usefulness:** as such, the model has no default data but relies on user-defined exposure and vulnerability. However, HydroMT-FIAT plugin can be used to apply JRC or HAZUS depth-damage functions.

GlobE-SoVI Reimann et al. (2024)	**Intended area of application:** global **Damage predictors:** multiple socioeconomic indicators **Publicly available:** yes (static dataset in different formats) **Exposure data:** none **Usefulness:** could be used to differentiate flood mortality between different countries, though it is not explicit how to apply the normalized vulnerability indicator for that purpose

2.1.2. Wind

PESETA wind vulnerability model JRC	**Intended area of application:** Europe (based on European data) **Damage predictors:** total construction value per country, return period **Publicly available:** yes (equation provided) **Exposure data:** total construction value per country from Eurostat **Usefulness:** more of a macroscale model, as it is based on country-level data from NatCatService (Munich Re), showing relatively good fit, but could be not very transferable to high-resolution grid due to problem with scaling the damage predictors
Wind-damage function review Gliksman et al. (2023)	**Intended area of application:** Europe **Damage predictors:** wind speed (various indicators) **Publicly available:** yes (equations provided) **Exposure data:** should be applicable to any economic exposure dataset **Usefulness:** this review provides wind-damage equations from four different studies based on different sets of data from European windstorms.
CLIMADA Schwierz et al. (2010)	**Intended area of application:** Europe **Damage predictors:** daily maximum wind speed **Publicly available:** yes (values in Excel file) **Exposure data:** should be applicable to any economic exposure dataset **Usefulness:** default CLIMADA functions based on UK insurance claims but validated from European storms. CLIMADA also provides in the same resource damage functions for other hazards.

2.1.3. Heat

Labour productivity loss curves Dasgupta et al. (2021)	**Intended area of application:** global **Damage predictors:** wet-bulb globe temperature (WBGT) **Publicly available:** yes (damage curves shown on figure) **Exposure data:** applicable to any labour supply (persons working) data **Usefulness:** the paper shows five curves from previous literature on labour productivity loss and proposes a new one, but focuses only on physically demanding work
ILO labour productivity loss curves ILO	**Intended area of application:** global **Damage predictors:** wet-bulb globe temperature (WBGT) **Publicly available:** yes (damage curves shown on figure) **Exposure data:** applicable to any labour supply data by sector **Usefulness:** report shows two sets of curves based on previous research, for three levels of labour intensity, including one set from international standard ISO 7243:1989
Heat mortality functions Ballester et al.	**Intended area of application:** Europe **Damage predictors:** mean weakly temperature, age, sex, region **Publicly available:** yes (R code) **Exposure data:** population by age and sex per region (from Eurostat)

(2023)	**Usefulness:** relatively easily reproducible approach to heat-mortality calculation, where weekly mortality data are fitted to location-specific temperature–lag–mortality functions, based on Eurostat population and mortality data.
Heat mortality functions Gallo et al. (2024)	**Intended area of application:** Europe **Damage predictors:** mean weakly temperature, age, sex, region **Publicly available:** yes (R code) **Exposure data:** population by age and sex per region (from Eurostat) **Usefulness:** modified version of Ballester et al. (2023), in which two calibration periods are used to create an adaptation counterfactual.
Heat impact on crops Agnolucci et al. (2020)	**Intended area of application:** global **Damage predictors:** mean temperature **Publicly available:** yes (R code) **Exposure data:** crop yields **Usefulness:** simple temperature-yield curves based on model predictions for 18 crops, enables quick estimation of impacts of warming on yields.

2.1.4. Fire

CLIMADA Lüthi et al. (2021)	**Intended area of application:** global **Damage predictors:** brightness of satellite image in Kelvins **Publicly available:** yes (equation and generic function in Excel file provided) **Exposure data:** calibrated using LitPop gridded asset values **Usefulness:** impact function directly converting satellite-detected fires into damages at different resolutions (1, 4, 10 km)

2.1.5. Drought

PESETA drought impact model Naumann et al. (2021)	**Intended area of application:** Europe **Damage predictors:** return period of drought (low daily river flow) **Publicly available:** yes (equation shown) **Exposure data:** GVA by selected sectors and total building stock (from Eurostat) **Usefulness:** country-level damage function for droughts, aggregating multiple sectors, fitted to NatCatService drought loss data for 1990–2016. Useful for directly computing economic loss from droughts, though rather for macroscale applications.
Wheat loss to drought and heat Zampieri et al. (2017)	**Intended area of application:** global **Damage predictors:** Heat Wave Magnitude Index daily, Standardized Precipitation Evaporation Index **Publicly available:** no (methods described) **Exposure data:** wheat yields **Usefulness:** shows that relatively simple indicators can be used to predict losses from combined drought and heat wave events
AquaCrop Vanuytrecht et al. (2014)	**Intended area of application:** global **Damage predictors:** temperature, rainfall, soil and management practices **Publicly available:** yes (stand-alone application) **Exposure data:** crop characteristics, soil properties, management practices **Usefulness:** a relatively simple crop model developed by FAO for easy application at local scales (computes one field at a time). Does not consider combined impacts of droughts and heat waves. A similar application is **DSSAT.**

2.2. Microscale models

Microscale models represent individual exposed objects, typically buildings or pieces of infrastructure. Those models are less applicable to certain hazards, particularly heat and drought, where this level of detail has limited added value.

2.2.1. Flooding

SaferPLACES residential loss model Paprotny et al. (2021)	**Intended area of application:** Europe (based on data from European countries) **Damage predictors:** water depth, velocity, return period, floor space, regional household income **Publicly available:** yes (Matlab code) **Exposure data:** intended to be used with data from Paprotny et al. (2020a) **Usefulness:** Well-calibrated model shown to provide better results than alternative models to validation case studies
SaferPLACES commercial loss model Paprotny et al. (2020b)	**Intended area of application:** Europe (based on data from European countries) **Damage predictors:** water depth, duration, use of precautionary measures, regional GVA per capita, regional investment per employee **Publicly available:** yes (Matlab code) **Exposure data:** intended to be used with data produced with the methodology described in the underlying paper **Usefulness:** Well-calibrated model shown to provide better results than alternative models to validation case studies
HAZUS FEMA (2022)	**Intended area of application:** USA (based on US data) **Damage predictors:** water depth, occupancy type, number of stories, basement **Publicly available:** yes (standalone software, but requires also ArcGIS installed) **Exposure data:** provides generic estimates for the USA, but user-defined external data can be used as well **Usefulness:** used officially in the US for FEMA's Natural Hazards Risk Assessment Program, and based on large amounts of reported losses, it is the main tool for US-based flood analyses. It also has separate modules for hurricane losses.

2.2.2. Wind

Wind speed and damage scale Feuerstein et al. (2011)	**Intended area of application:** Central Europe **Damage predictors:** building strength, wind speed **Publicly available:** yes (equations provided) **Exposure data:** requires building dataset indicating structural characteristics **Usefulness:** provides additional details of data if building construction type can be obtained.

3. Impact data

This section covers global, European and national datasets which include information on flood impacts, ranging from textual to quantitative data, covering footprints, fatalities, economic losses and the like. For national datasets, only those covering case study areas are included.

3.1. Global datasets

ReliefWeb OCHA	**Hazards:** all types **Timeframe:** c. 1980s - present **Usefulness:** provides situation reports for major disasters, it is often the most comprehensive source of data for developing countries (generally doesn't include events in high-income countries) **Update schedule:** continuous
Emergency situation reports WHO	**Hazards:** all types **Timeframe:** recent **Usefulness:** generally focused on epidemics, but sometimes also provides situation reports for major natural disasters, mainly in least developed countries in Africa and Asia **Update schedule:** continuous
Post-Disaster Needs Assessments UN/World Bank	**Hazards:** all types **Timeframe:** 2007 - present **Usefulness:** comprehensive reports on impacts from particularly large disasters in developing countries **Update schedule:** continuous
GLIDEnumber Asian Disaster Reduction Centre	**Hazards:** all types **Timeframe:** 1998 – present **Usefulness:** short event descriptions add context or sometimes more data for events that are generally available from other datasets, especially EM-DAT **Update schedule:** continuous
EM-DAT CRED	**Hazards:** all types **Timeframe:** 1900 – present (highly incomplete before 1980s) **Usefulness:** primary global database harmonizing disaster information mainly from press sources. For developing countries, often draws information from ReliefWeb (not always correctly) **Update schedule:** continuous
Dartmouth Flood Observatory University of Colorado	**Hazards:** floods **Timeframe:** 1985 – present **Usefulness:** flood footprints from remote sensing and short event descriptions based on news media, though in recent years only uses FloodList for impact data **Update schedule:** continuous
Global Flood Database Tellman et al. (2021)	**Hazards:** floods **Timeframe:** 2000–2018 **Usefulness:** flood footprints from remote sensing and impact data **Update schedule:** unknown
FloodList ECMWF, Copernicus	**Hazards:** floods **Timeframe:** 2013 – present **Usefulness:** descriptions of flood events with impact data, photographs and links to news and social media on the event. Website is privately run, but financed by ECMWF and Copernicus **Update schedule:** continuous

DesInventar UNDRR	**Hazards:** all types **Timeframe:** varies by country **Usefulness:** the database is a collection of national databases, each with different coverage and level of detail, but is usually the most comprehensive source for many countries, almost all of which are low- and medium-income countries. **Update schedule:** continuous
Copernicus EMS European Commission	**Hazards:** all types **Timeframe:** 2012 - present **Usefulness:** short event descriptions, often of still ongoing events, and occasionally post-disaster assessments. Often provide more precise spatial information on location of event than other sources. **Update schedule:** continuous

3.2. European datasets

Weekly mortality data Eurostat	**Hazards:** none specific **Timeframe:** 2000–2024 **Usefulness:** weekly mortality can be used to detect excess mortality due to cold and heat waves. **Update schedule:** continuous
HANZE Paprotny et al. (2023b)	**Hazards:** floods **Timeframe:** 1870–2020 **Usefulness:** main resource on European flood impacts, standardizing footprints (NUTS3 regions), impact indicators and descriptions. **Update schedule:** irregular (current version: v2.1, August 2023)
Floods Directive	**Hazards:** floods **Timeframe:** various **Usefulness:** compilation of national preliminary flood risk assessments, which have to include information on historical floods, prepared under the Floods Directive. A combined database prepared by the European Environment Agency is no longer accessible online. **Update schedule:** irregular (7-yearly cycle of the Floods Directive, first version 2012, updated 2019)

3.3. Local datasets

Xynthia insured loss data FFSA / GEMA	**Hazards:** coastal flood and windstorm **Timeframe:** 2010 Xynthia storm **Usefulness:** provides data on insured loss across France, the most detailed data available for the event: at regional (department, NUTS3) level (by cause and type of asset), for several most affected regions main cause also split by assets, for two most affected regions also by municipality and cause. Also shows total insured loss from biggest disasters in France since 1988. **Update schedule:** one-off
SurgeWatch University of Southampton	**Hazards:** coastal floods **Timeframe:** 1014-2018 **Usefulness:** collection of information on causes and impacts of major storm surges affecting the United Kingdom, sourced from literature and news media. Most comprehensive source of coastal flooding data for the UK. **Update schedule:** unknown
Recorded Flood Outlines	**Hazards:** floods **Timeframe:** 1946-present

Environment Agency	**Usefulness:** collection of flood footprints obtained for England through various means (from visual to remote sensing), collected continuously since 1946. **Update schedule:** several times a year

4. References

Agnolucci, P., Rapti, C., Alexander, P. et al. (2020) Impacts of rising temperatures and farm management practices on global yields of 18 crops. Nature Food, 1, 562–571, https://doi.org/10.1038/s43016-020-00148-x

Ballester, J., Quijal-Zamorano, M., Méndez Turrubiates, R.F. et al. (2023) Heat-related mortality in Europe during the summer of 2022. Nature Medicine, 29, 1857–1866, https://doi.org/10.1038/s41591-023-02419-z

Boyd, E., Levitan, M., and van Heerden, I. (2010) Improvements in Flood Fatality Estimation Techniques Based on Flood Depths, in: Wind Storm and Storm Surge Mitigation, edited by: Uddin, N., American Society of Civil Engineers, Reston, Virginia, USA, 126–139, https://doi.org/10.1061/9780784410813.ch11

Dasgupta, S. et al. (2021) Effects of climate change on combined labour productivity and supply: an empirical, multi-model study. Lancet Planetary Health, 5, e455–e465, https://doi.org/10.1016/S2542-5196(21)00170-4

Eberenz, S., Stocker, D., Röösli, T., Bresch, D. N. (2020) Asset exposure data for global physical risk assessment. Earth System Science Data, 12, 817–833, https://doi.org/10.5194/essd-12-817-2020

FEMA (2022) Hazus Flood Technical Manual, https://www.fema.gov/sites/default/files/documents/fema_hazus-flood-model-technical-manual-5-1.pdf

Gallo, E., Quijal-Zamorano, M., Méndez Turrubiates, R.F. et al. (2024) Heat-related mortality in Europe during 2023 and the role of adaptation in protecting health. Nature Medicine, https://doi.org/10.1038/s41591-024-03186-1

Gerl, T., Kreibich, H., Franco, G., Marechal, D., & Schröter, K. (2016). A review of flood loss models as basis for harmonizationand benchmarking. PLoS One, 11, 1–22. https://doi.org/10.1371/journal.pone.0159791

Gliksman, D., Averbeck, P., Becker, N., Gardiner, B., Goldberg, V., Grieger, J., Handorf, D., Haustein, K., Karwat, A., Knutzen, F., Lentink, H. S., Lorenz, R., Niermann, D., Pinto, J. G., Queck, R., Ziemann, A., Franzke, C. L. E. (2023) Review article: A European perspective on wind and storm damage – from the meteorological background to index-based approaches to assess impacts. Natural Hazards and Earth System Sciences, 23, 2171–2201, https://doi.org/10.5194/nhess-23-2171-2023

Huizinga, J., de Moel, H., Szewczyk, W. (2017) Global flood depth-damage functions. Methodology and the database with guidelines. Publications Office of the European Union, Luxembourg, https://doi.org/10.2760/16510

Hurtt, G. C., Chini, L., Sahajpal, R., Frolking, S., Bodirsky, B. L., Calvin, K., Doelman, J. C., Fisk, J., Fujimori, S., Klein Goldewijk, K., Hasegawa, T., Havlik, P., Heinimann, A., Humpenöder, F., Jungclaus, J., Kaplan, J. O., Kennedy, J., Krisztin, T., Lawrence, D., Lawrence, P., Ma, L., Mertz, O., Pongratz, J., Popp, A., Poulter, B., Riahi, K., Shevliakova, E., Stehfest, E., Thornton, P., Tubiello, F. N., van Vuuren, D. P., Zhang, X. (2020) Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. Geoscientific Model Development, 13, 5425–5464, https://doi.org/10.5194/gmd-13-5425-2020

Jonkman, S. N., Vrijling, J. K., and Vrouwenvelder, A. C. W. M. (2008) Methods for the estimation of loss of life due to floods: a literature review and a proposal for a new method. Natural Hazards, 46, 353–389, https://doi.org/10.1007/s11069-008-9227-5

Klein Goldewijk, C.G.M., Beusen, A., Doelman, J., Stehfest, E. (2017) Anthropogenic land use estimates for the Holocene – HYDE 3.2. Earth System Science Data, 9, 927–953, https://doi.org/10.5194/essd-9-927-2017

Lüthi, S., Aznar-Siguan, G., Fairless, C., Bresch, D. N. (2021) Globally consistent assessment of economic impacts of wildfires in CLIMADA v2.2. Geoscientific Model Development, 14, 7175–7187, https://doi.org/10.5194/gmd-14-7175-2021

Marconcini, M., Metz-Marconcini, A., Esch, T., Gorelick, N. (2021) Understanding Current Trends in Global Urbanisation - The World Settlement Footprint Suite. 12th International Symposium on Digital Earth, https://doi.org/10.1553/giscience2021_01_s33

Naumann, G., Cammalleri, C., Mentaschi, L. et al. (2021) Increased economic drought impacts in Europe with anthropogenic warming. Nature Climate Change, 11, 485–491, https://doi.org/10.1038/s41558-021-01044-3

Paprotny D., Kreibich H., Morales Nápoles O., Castellarin A., Carisi F., Schröter K. (2020a) Exposure and vulnerability estimation for modelling flood losses to commercial assets in Europe. Science of the Total Environment, 737, 140011, https://doi.org/10.1016/j.scitotenv.2020.140011

Paprotny D., Kreibich H., Morales Nápoles O., Terefenko P., Schröter K. (2020b) Estimating exposure of residential assets to natural hazards in Europe using open data. Natural Hazards and Earth System Sciences, 20, 323–343, https://doi.org/10.5194/nhess-20-323-2020

Paprotny D., Kreibich H., Morales Nápoles O., Wagenaar D., Castellarin A., Carisi F., Bertin X., Merz B., Schröter K. (2021) A probabilistic approach to estimating residential losses from different flood types. Natural Hazards, 105(3), 2569–2601, https://doi.org/10.1007/s11069-020-04413-x

Paprotny D., Mengel M. (2023a) Population, land use and economic exposure estimates for Europe at 100 m resolution from 1870 to 2020. Scientific Data, 10, 372, https://doi.org/10.1038/s41597-023-02282-0

Paprotny D., 't Hart C. M. P., Morales-Napoles O. (2024) Evolution of flood protection levels and flood vulnerability in Europe since 1950 estimated with vine-copula models. Natural Hazards, in revision, https://doi.org/10.21203/rs.3.rs-4213746/v1

Paprotny D., Terefenko P., Śledziowski J. 2023b. An improved database of flood impacts in Europe, 1870–2020: HANZE v2.1. Earth System Science Data, accepted in print, https://doi.org/10.5194/essd-2023-321

Reimann, L., Koks, E., de Moel, H., Ton, M. J., & Aerts, J. C. J. H. (2024) An Empirical Social Vulnerability map for flood risk assessment at global scale ("GlobE-SoVI"). Earth's Future, 12, e2023EF003895, https://doi.org/10.1029/2023EF003895

Schwierz, C., Köllner-Heck, P., Zenklusen Mutter, E. et al. (2010) Modelling European winter wind storm losses in current and future climate. Climatic Change, 101, 485–514, https://doi.org/10.1007/s10584-009-9712-1

Smits, J., Permanyer, I. (2019) The Subnational Human Development Database. Scientific Data, 6, 190038 https://doi.org/10.1038/sdata.2019.38

Steinhausen M., Paprotny D., Dottori F., Sairam N., Mentaschi L., Alfieri L., Lüdtke S., Kreibich H., Schröter K. (2022) Drivers of future fluvial flood risk change for residential buildings in Europe. Global Environmental Change, 76, 102559, https://doi.org/10.1016/j.gloenvcha.2022.102559

Tellman, B., Sullivan, J.A., Kuhn, C. et al. (2021) Satellite imaging reveals increased proportion of population exposed to floods. Nature 596, 80–86, https://doi.org/10.1038/s41586-021-03695-w

Vanuytrecht, E., Raes, D., Steduto, P., Hsiao, T. C., Fereres, E., Heng, L. K., Vila, M. G., and Moreno, P. M. (2014) AquaCrop: FAO's crop water productivity and yield response model. Environmental Modelling & Software, 62, 351–360, https://doi.org/10.1016/j.envsoft.2014.08.005

Wenz, L., Carr, R.D., Kögel, N. et al. (2023) DOSE – Global data set of reported sub-national economic output. Scientici Data, 10, 425 https://doi.org/10.1038/s41597-023-02323-8

Zampieri, M., Ceglar, A., Dentener, F., Toreti, A. (2017) Wheat yield loss attributable to heat waves, drought and water excess at the global, national and subnational scales. Environmental Research Letters, 12, 6, https://doi.org/10.1088/1748-9326/aa723b