

Spoke 0 - WP2

Armonisation and Validation of Data and Indicators in Amelia

Alessandro Fassò
Bologna, 19/20 febbraio 2024



**UNIVERSITÀ
DEGLI STUDI
DI BERGAMO**

Dipartimento
di Scienze Economiche

WP0.2 main activity

D0.2.1 Harmonised dataset/indicators

- ▶ Geostatistical **methodology** for Amelia
- ▶ Showcase on environmental hazards

D0.2.2 Datasets obtained by merging households data sources

D0.2.3 S.A.F.E. assessment, **validation** and comparison of the proposed models and indicators.

TALK AIMS

My presentation aims at

Part 1 Sharing with you an approach to quality assessment of Amelia datasets

Part 2 Showing you the datasets we are handling and how

PART 1

Quality assessment

Quality assessment of Amelia Data Products

BOLLINATURA

Aim: to "classify" the quality of each **data product** in the Amelia digital ecosystem.

This could be a part of an overall GRINS quality certification.

Data Product

Typical data products are

- ▶ Source data sets
- ▶ Harmonised data sets
- ▶ Statistical models/ML (georeferenced) outputs
- ▶ GRINS Indicators
- ▶ GRINS surveys

Quality dimensions

BOLLINATURA should be based on

- ▶ Data source quality
- ▶ Data transformation quality
- ▶ (Graphical interface and Output layout)
- ▶ (Privacy, Intellectual property, etc.)

Data Source Quality

A Official sources ("certified")

1. Validated data
 2. Interim data
- Mandatory Metadata

B Publishable sources

1. GRINS Surveys
 2. Peer review publication
- Mandatory Metadata

C Other sources (internal use only)

1. Working data
 2. Others
- Recommended Metadata

Data Transformation Quality

Harmonisation/Downscaling/Aggregation/Indicators/Statistical models/ML

- ▶ Code

- ▶ Standardised documentation
- ▶ Fully traceable to other qualified data sources
- ▶ Code availability
- ▶ Repeatability

- ▶ Statistical/ML models

- ▶ Peer review publication
- ▶ Model validation and its transparency
- ▶ Uncertainty assessment and propagation

Data Transformation Quality Levels

A Fully satisfactory \Rightarrow Public Certified

B Partially satisfactory \Rightarrow Public Not
Certified

C Internal use only

Data Quality Matrix

| | | Data source | | |
|-------------|---|-------------|---------|---------|
| | | A | B | C |
| Data transf | A | aa1/aa2 | ab1/ab2 | ac1/ac2 |
| | B | ba1/ba2 | bb1/bb2 | bc1/bc2 |
| | A | ca1/ca2 | cb1/cb2 | cc1/cc2 |

Validation transparency - The Copernicus example

Evaluation and Quality Assurance (EQA) of CAMS-AQ NRT



PART 2

Datasets covered by WP0.2/D0.2.1 and Examples of Harmonisation

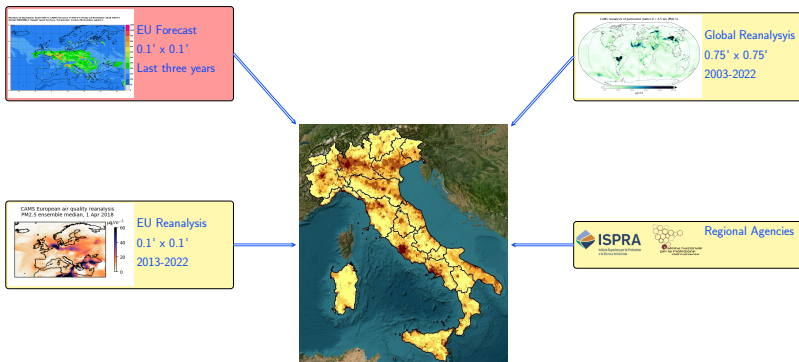
Datasets covered by WP02.2

- ▶ Atmosphere
 - ▶ Meteo-climatic data (ERA5)
 - ▶ AQ concentrations (CAM5+ISPRA+ARPAs)
 - ▶ AQ emissions (CAM5)
- ▶ Fires (CEMS)
- ▶ Hydraulic hazard (ISPRA)

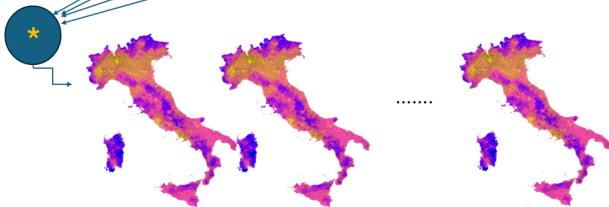
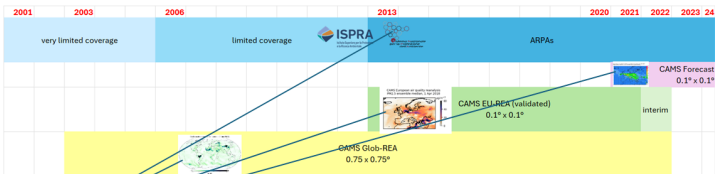
Additional output:

Methodological report/paper to be shared in GRINS

Example: CAMS AQ harmonisation



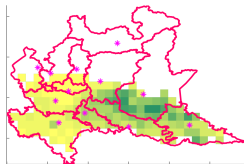
Example: CAMS AQ harmonisation



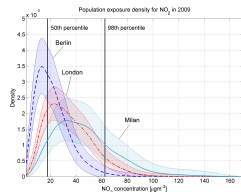
Example: CAMS AQ harmonisation

Statistical methods:

- ▶ Data fusion
- ▶ Downscaling
- ▶ Change of support
- ▶ Uncertainty assessment
- ▶ High resolution + Municipal level
 - ▶ AQ indicators (KPI)
 - ▶ Spatial distribution
 - ▶ Exposure distribution
 - ▶ Long time series



Fassò et al. (2023) Scientific data



Fassò et al. (2016) JABES

GRAZIE!