Quality Control and Creation of Grids of Meteorological Variables for the Copernicus Emergency Management Services

Markus Ziese[1], Zora L. Schirmeister[1], Carina-Denise Lemke[1], Jakub P. Walawender[1], Christoph Schweim[2], Damien Pichon[3], Stefania Grimaldi[4], Gonçalo Gomes[4], Peter Salamon[4]

1 Deutscher Wetterdienst

Emergency Management

Background CEMS-MDCC

- Copernicus Emergency Management Services (CEMS)– Meteorological Data Collection Centre (MDCC):
  - Collects near-real time meteorological observations for Europe and neighboring areas
  - Stores them in a data bank
  - Controls the quality of the data
  - Post-processes data (aggregation, minimum/maximum/mean calculation, etc.)
  - Creates tailored analyses for CEMS components EFAS, EDO and EFFIS
- Established in 2003 at JRC Ispra, operated since 2016 by a consortium of KISTERS AG and German Meteorological Service

Real-time Data Provider

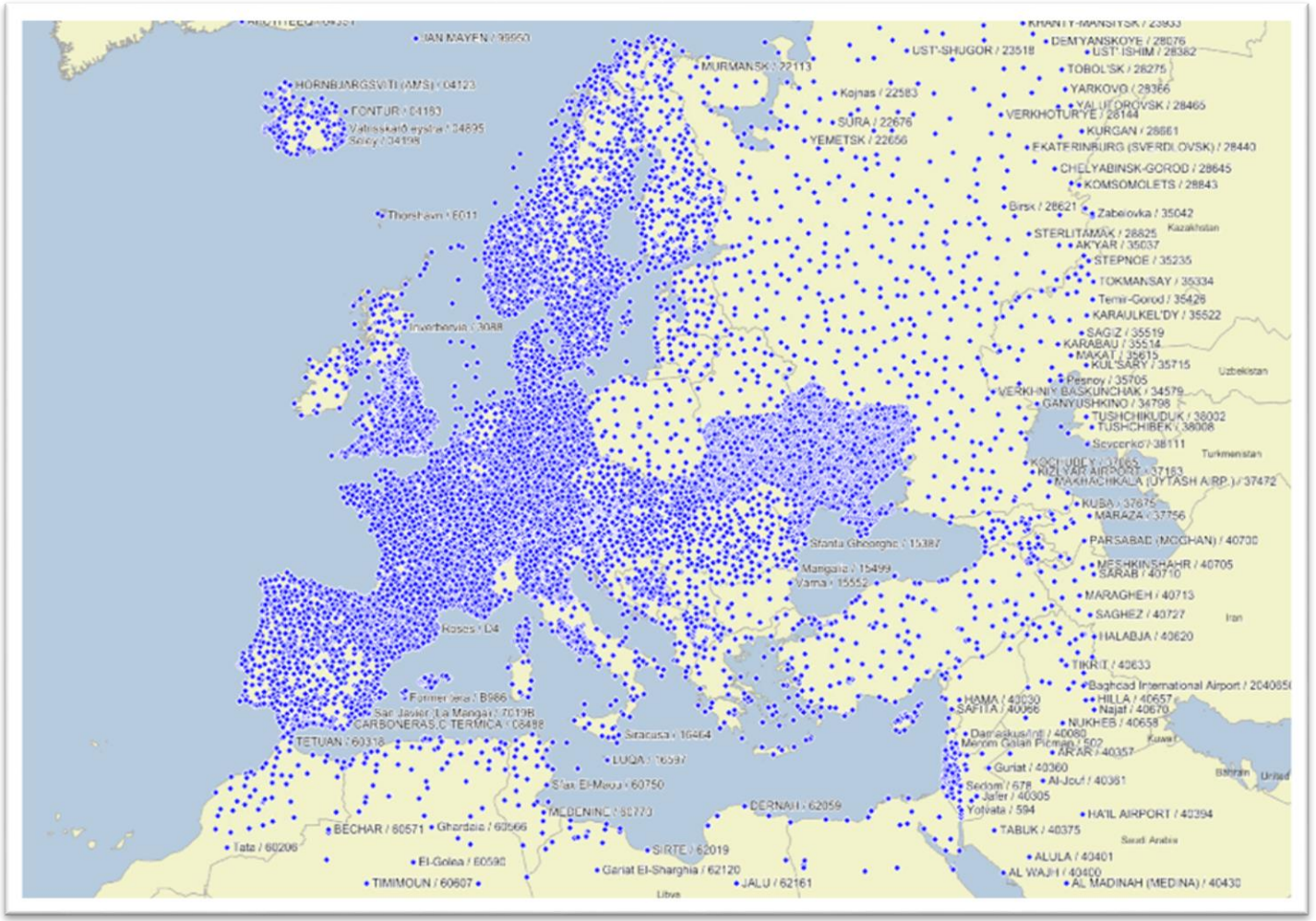
Emergency Management

32 currently active data providers

www.efas.eu

Station Distribution

Active stations in CEMS-MDCC database (EFAS domain)



Number of stations:

- Total: >54.500
- Active: ~22.000
  - thereof 4695 virtual stations
- Inactive: >31.500
  - thereof 9310 virtual stations

Emergency Management

www.efas.eu

Copernicus Europe's eyes on Earth

European Commission

Some numbers

Emergency Management

- > 70,246 stations in database
- > 9,000,000 data records to process each day (to be QC'ed)
- > 1300 GB disc space of database

Parameter description	Active stations in EFAS domain
Precipitation	~ 17300
Relative Air humidity	~ 8600
Solar radiation	~ 5800
Air temperature	~ 16800
Water vapor pressure	~ 4170
Wind speed	~ 13300

Quality Control at the MDCC

Emergency Management

- Data are QC‘ed:
  - When they are imported to the data bank system WISKI (Water Information System KISTERS; agent or synchronous validation)
  - External data validation framework outside WISKI (asynchronous data validation)
  - Defined QC-flags: good, estimated, suspect, rejected, missing
- Synchronous validation:
  - \+ Triggered by the import into WISKI
  - \+ Post-processed data always QC‘ed
  - \- Can only compare with existing data (‚order of import‘ matters)
  - \- No spatial validation possible
  - \- Rely on implemented validation agents
- Asynchronous Validation:
  - \+ Run of QC scheduled by user
  - \+ Applies pre-defined and user-defined rules
  - \+ Applicable for within time series and spatial validation
  - \+ Uses delayed data for comparison
  - \+ Data are extracted from data base and QC-flags written back

Synchronous Data Validation

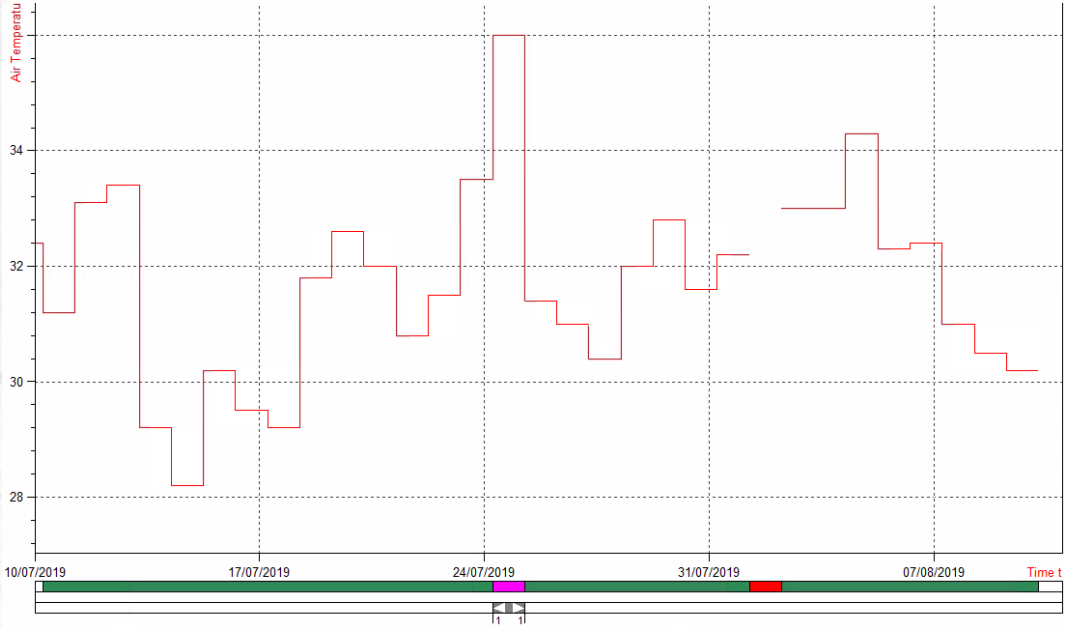
Emergency Management

- Comparison against thresholds, other time series (parameters), temporal consistency and rate of change
  - Temporal consistency: checks reporting frequency and adds missings in data gaps -> needed to get correct temporal coverage for post-processing
  - Rate of change: detects abrupt changes in time series, e.g. wrong sign in air temperature
  - Cross-validation: Consistency between parameters, e.g. air temperature and dew point temperature
  - Thresholds: checks exceedance of minimum and maximum limits
- Current setting: one threshold for whole EFAS-domain (Northern Africa to Arctic) and no annual cycle (except air temperature)

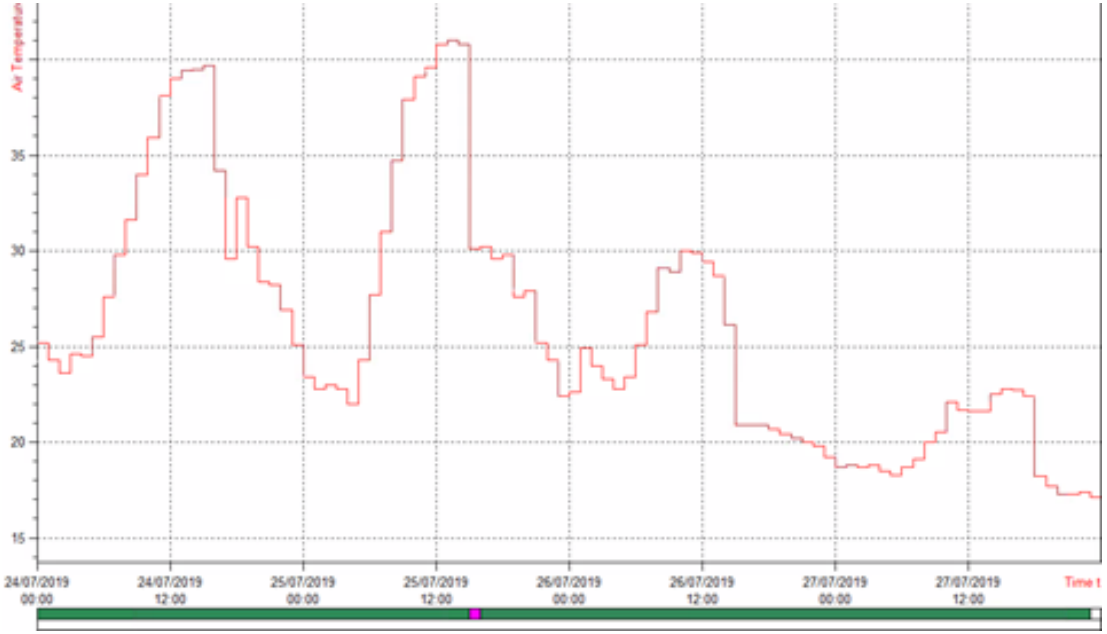
Synchronous Data Validation

Emergency Management

- Real (extreme) values can be outside the thresholds:



Minimum temperature above threshold during heat wave



Value ok, difference to previous due to cold front

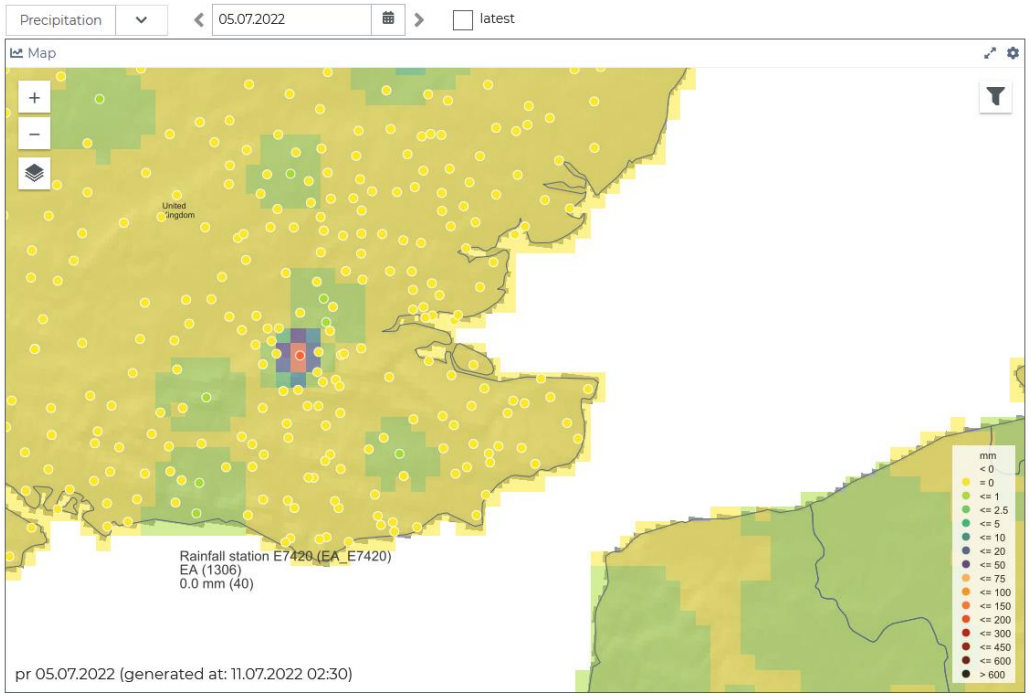
Copernicus Europe's eyes on Earth

European Commission

Synchronous Data Validation

Emergency Management

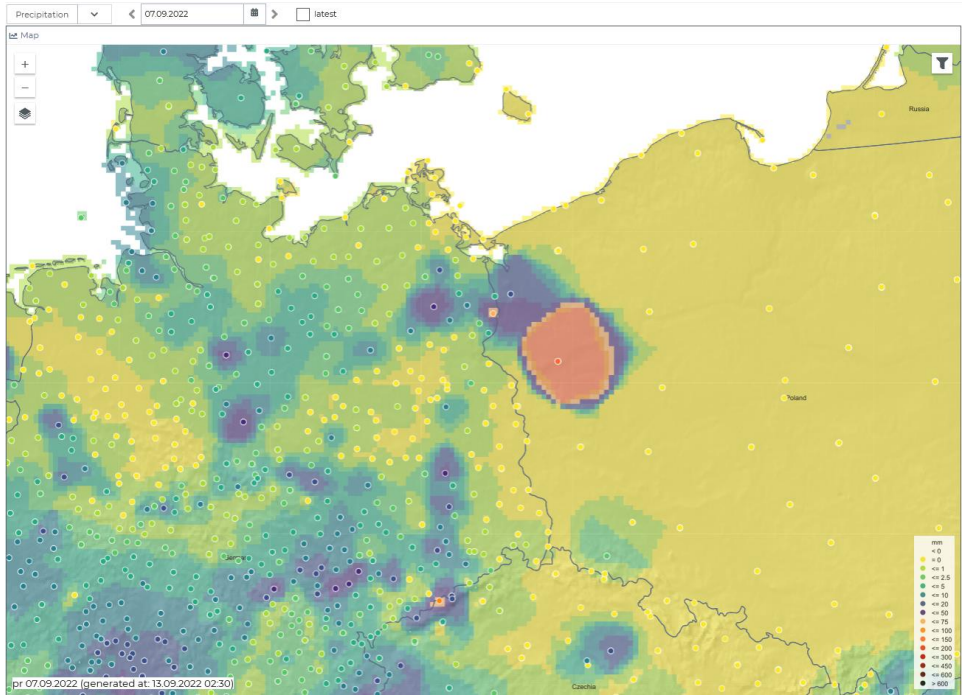
- Erroneous values can be inside the thresholds:
  - not all errors were detected (left), but real extremes have to pass (right)



144 mm for 6 July 2022 in England

Yellow: 0.0 mm

Dark blue: > 50 mm



130 mm for 9 September 2022 in Poland

Orange: > 130 mm

Confirmed by Radar data

Synchronous Data Validation

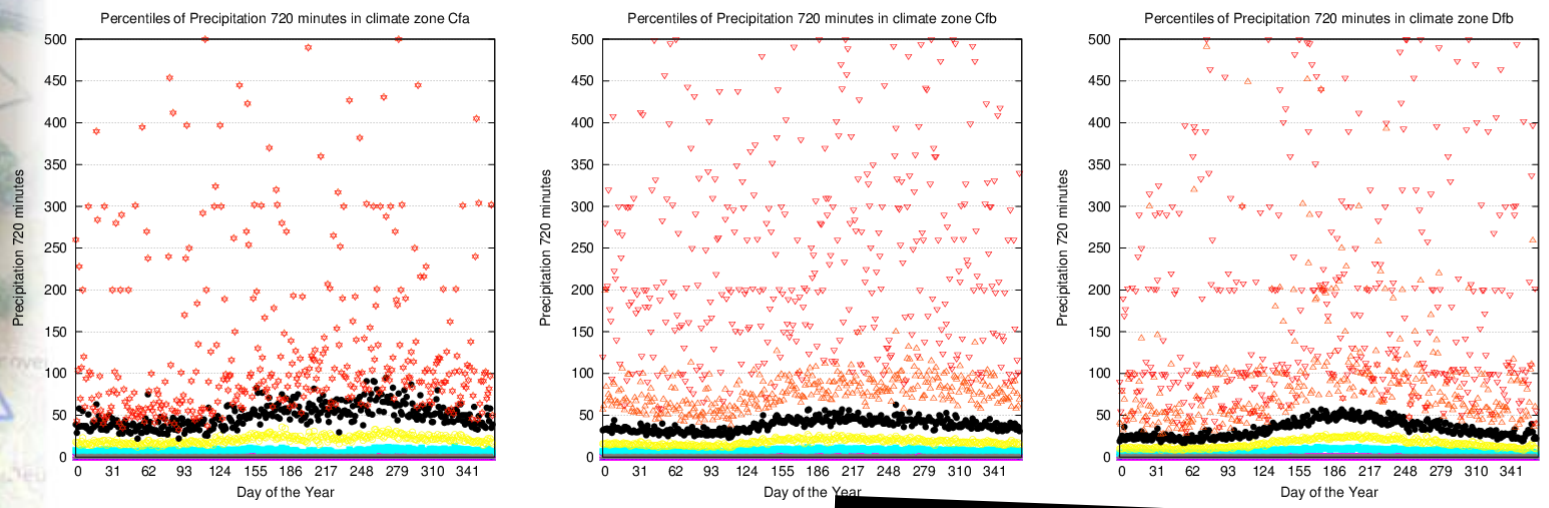
Emergency Management

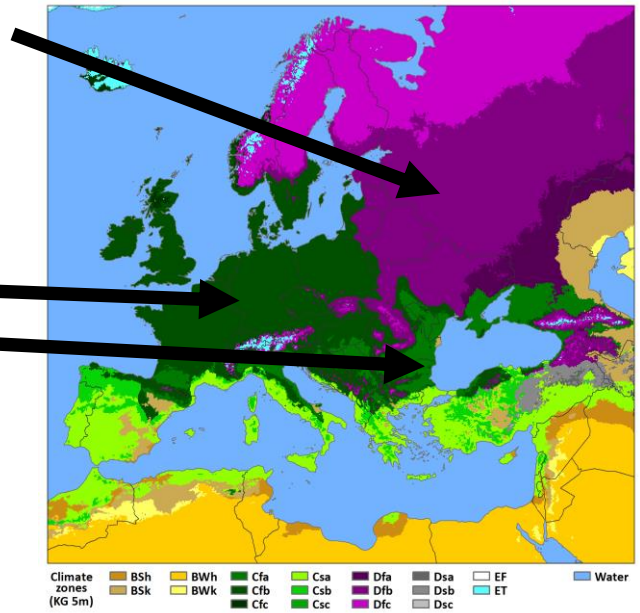
- Improvement: define seasonal and geographical thresholds
  - Use existing data in data base to calculate thresholds
  - Group stations by climate zone according to Köppen-Geiger climate classification
  - Calculate various percentiles
  - Visual inspection to define which percentile in which climate zone and offset is applied
- Turned out that this approach is not applicable to all parameters
  - Wind speed: low seasonal and geographical dependency, no improvement with new approach, but adjusted existing threshold (45 m/s -> 50 m/s)
  - Precipitation: too noisy in real extremes for calculation of reliable seasonal thresholds
  - Wind direction and cloud cover have neither a geographical nor a seasonal dependency

Synchronous Data Validation

Emergency Management

- 12-hourly precipitation data





www.efas.eu

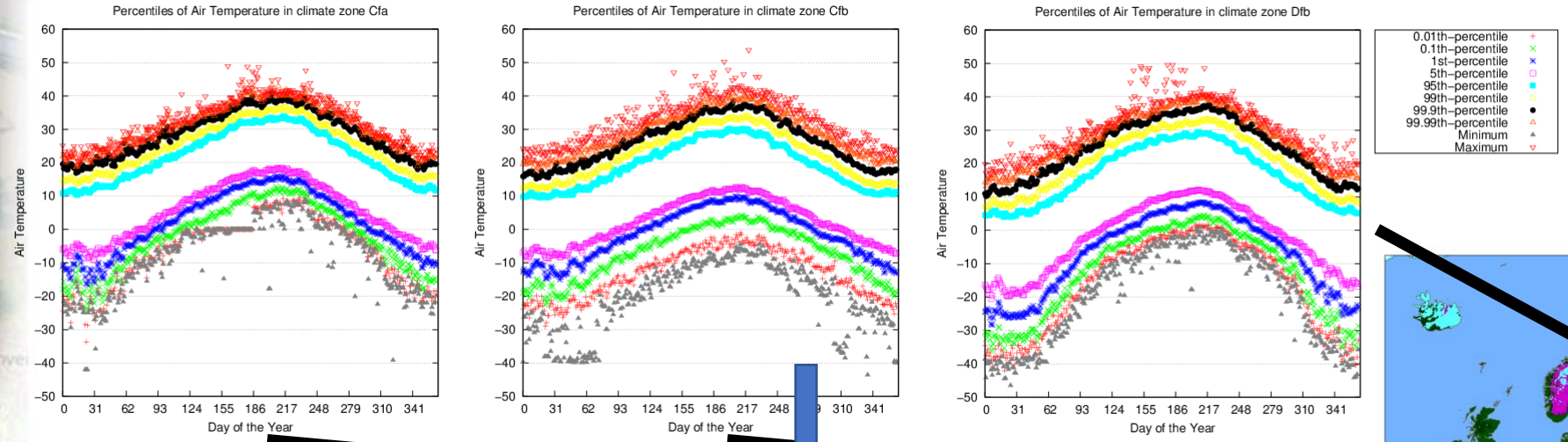
Copernicus Europe's eyes on Earth

European Commission

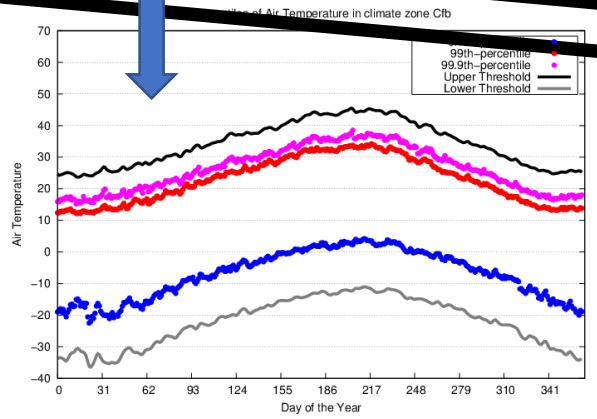
Synchronous Data Validation

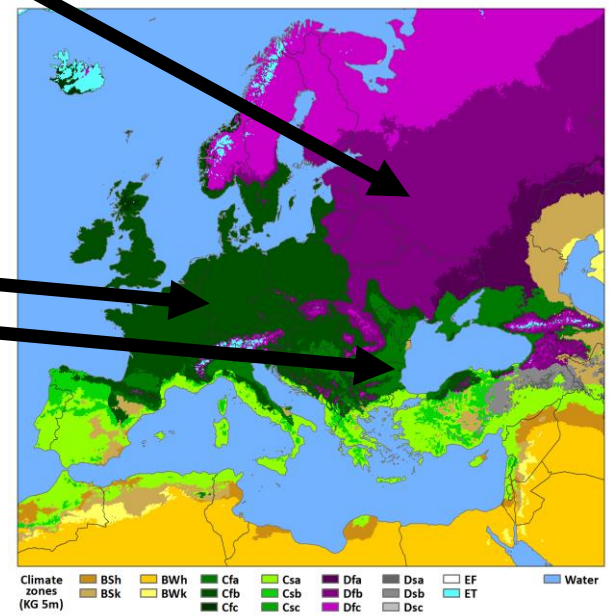
Emergency Management

- Air temperature



- Added 7-days low-pass filter
- Upper limit: 99.9% + 5K offset
- Lower limit: 0.1% - 10K offset





www.efas.eu

Copernicus Europe's eyes on Earth

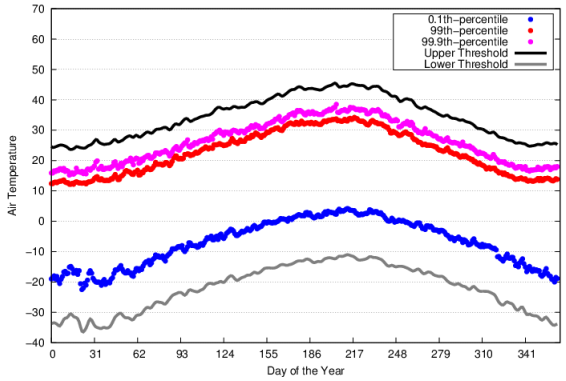
European Commission

Synchronous Data Validation

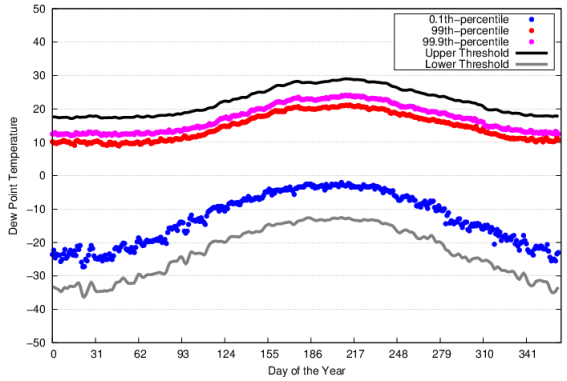
Emergency Management

- Expect elimination of more outliers by the seasonal and geographical thresholds (examples for Cfb)

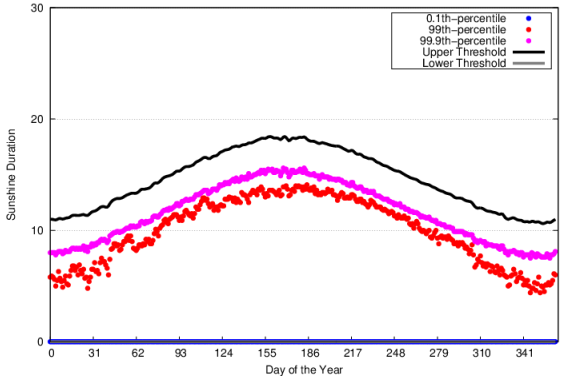
Air temperature



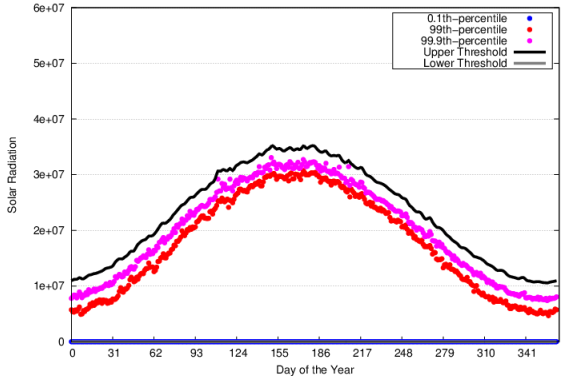
Dew point temperature



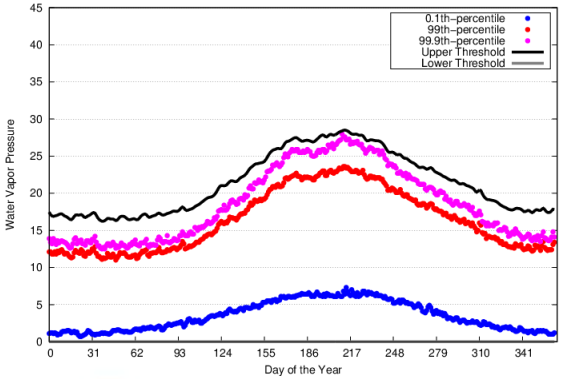
Sunshine duration



Solar radiation



Water vapor pressure



www.efas.eu

Copernicus Europe's eyes on Earth

European Commission

Asynchronous Data Validation

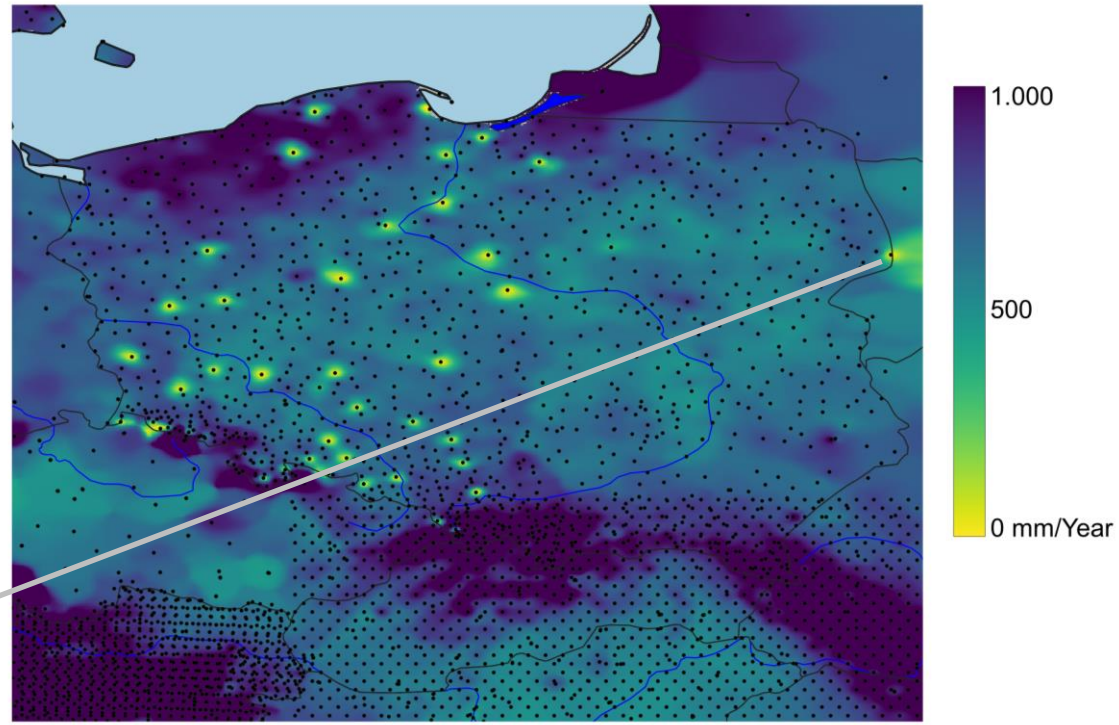
Emergency Management

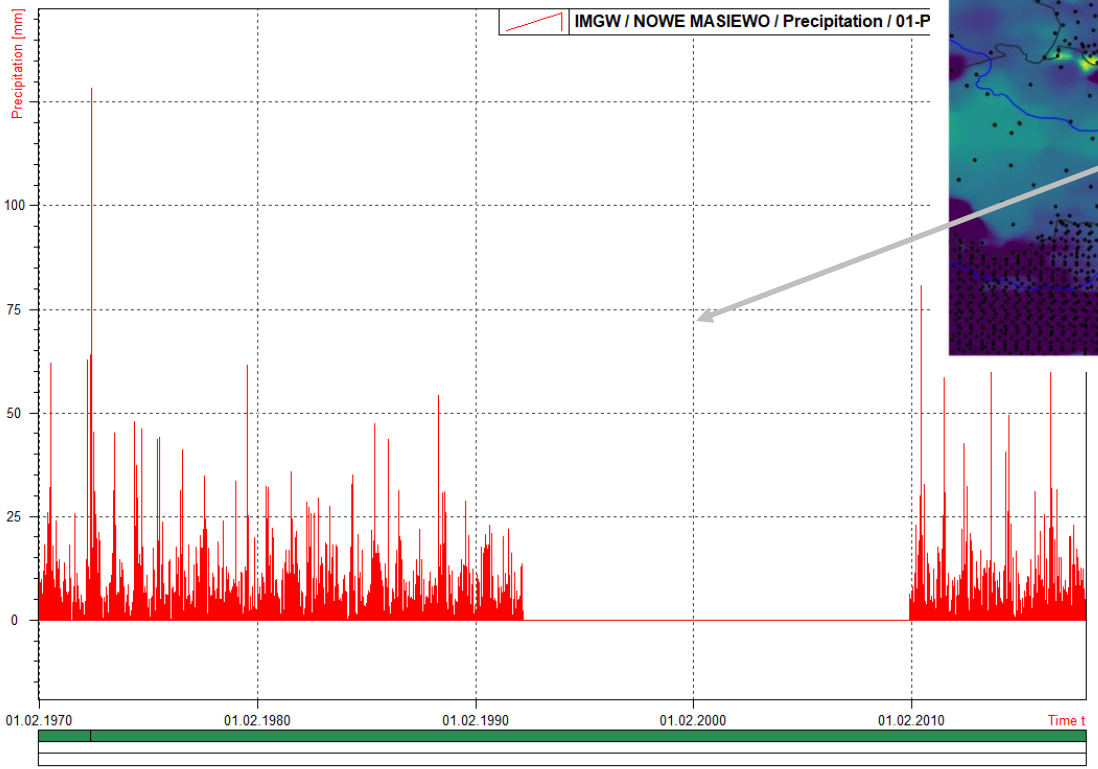
- Many predefined rules available:
  - Flatliner detection (value sequence)
  - Completeness
  - Range
  - Rate of change
  - Spatial comparison / spatial zero comparison
  - As well as user developed rules (Python interface)
- Currently best settings for rules under investigation, e.g. zero precipitation flatliner (missing data were filled with zeros)

Asynchronous Data Validation

Emergency Management

- Use case: flatliner detection





Note: Zeros may have been introduced by mishandling gaps during data preparation - they are unlikely to be in your copy of the record!

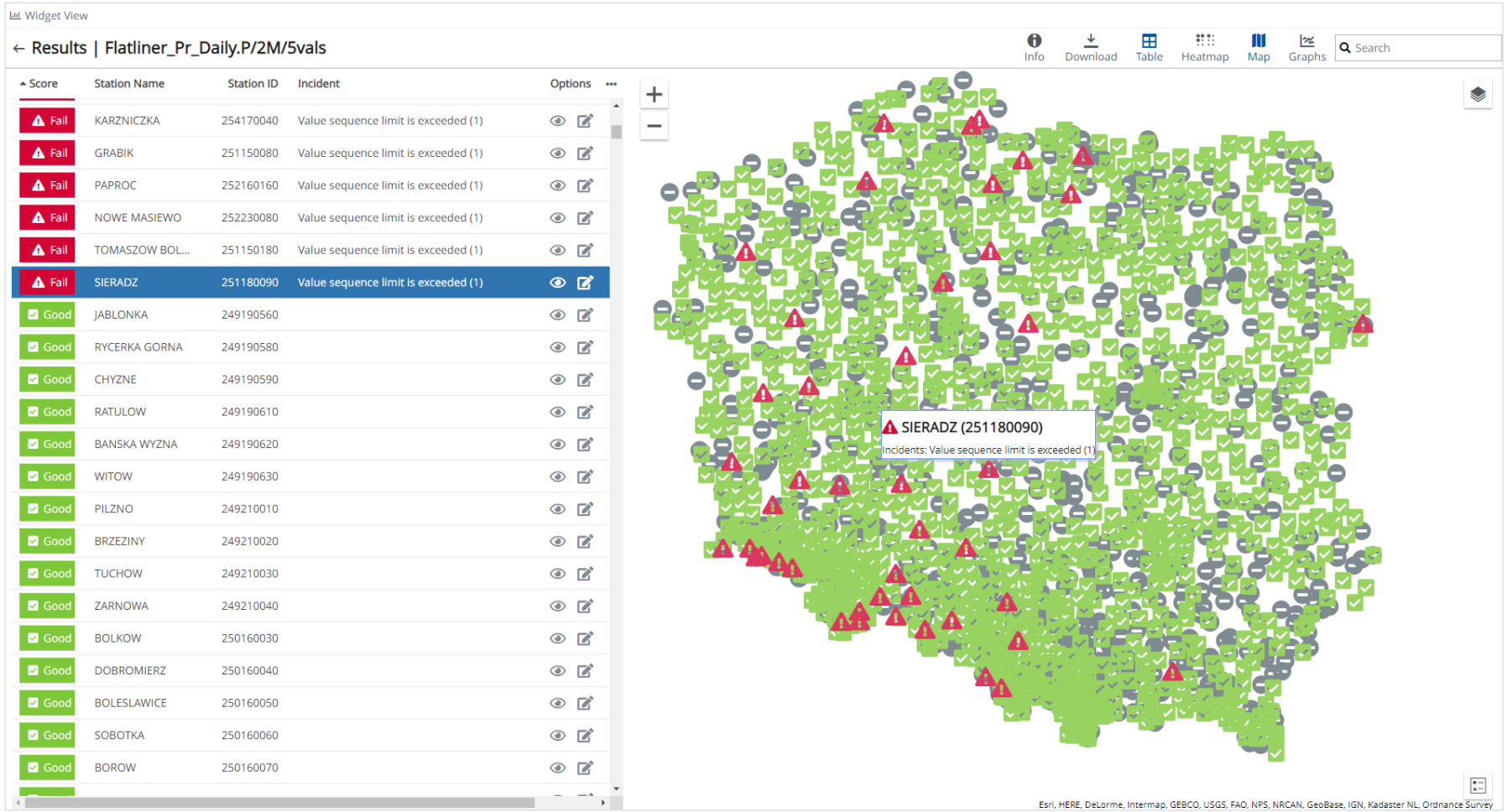
Copernicus Europe's eyes on Earth

European Commission

Asynchronous Data Validation

Emergency Management

- Use case: flatliner detection



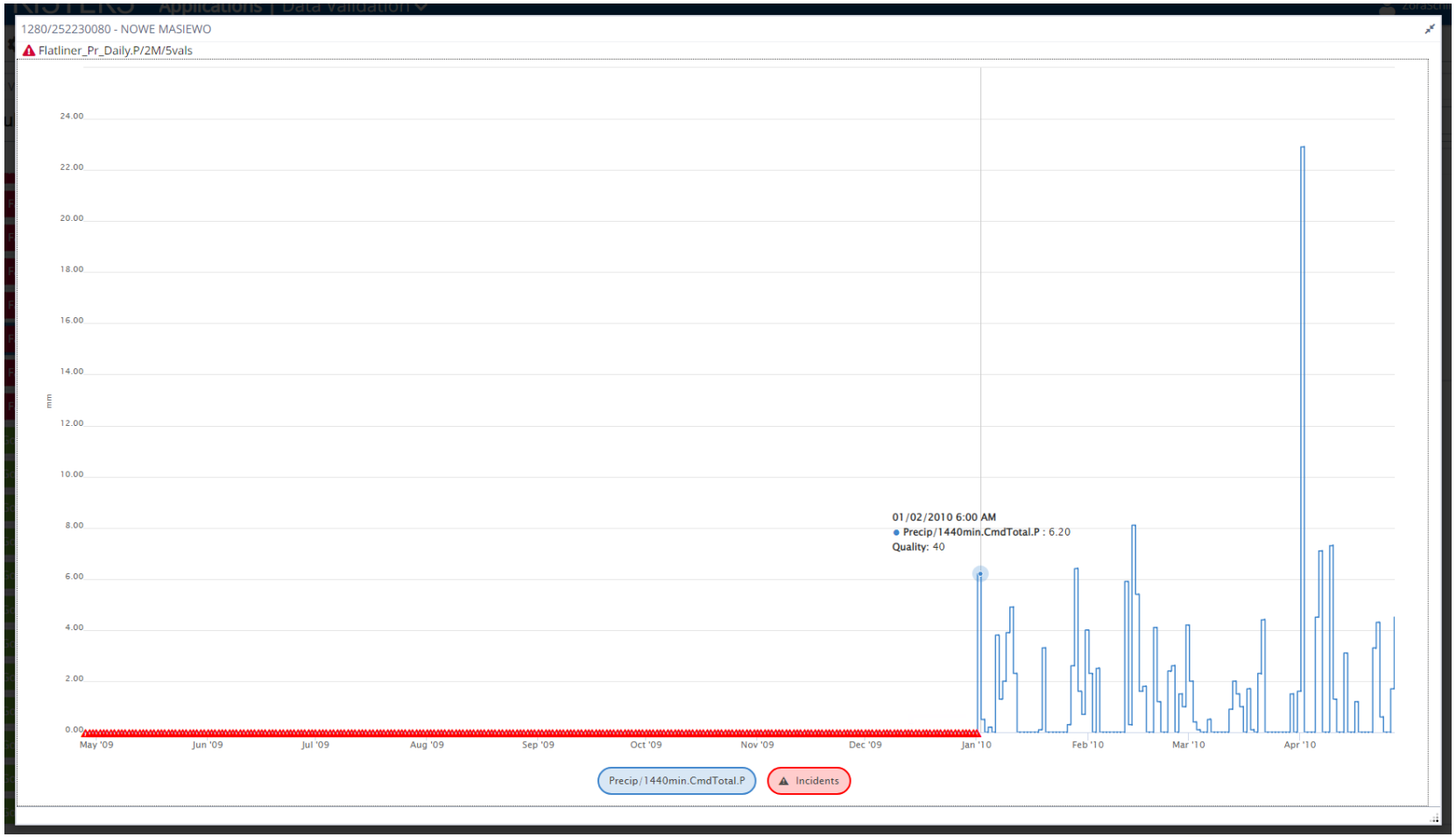
Copernicus Europe's eyes on Earth

European Commission

Asynchronous Data Validation

Emergency Management

- Use case: flatliner detection



Copernicus Europe's eyes on Earth

European Commission

MDCC Grid Creation

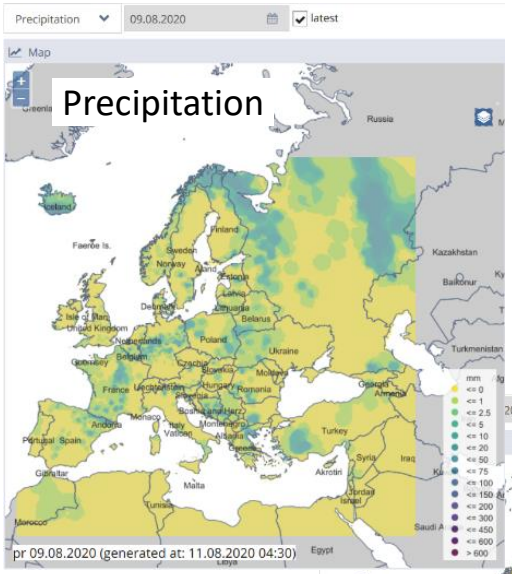
Emergency Management

- Only data with quality good, estimated or suspect are used for gridding
- Gridding by means of a modified SPHEREMAP scheme (4 – 10 station per grid point)
- Calculate an uncertainty estimation for each grid, depending on data availability and variance between input data
- Spatial resolution: 5 x 5 km² (950.000 cells), soon 1 x 1 arcmin² (13.454.100 cells)
- Gridded parameters:
  - Precipitation (6-hourly and daily)
  - Air temperature (6-hourly mean, daily minimum and maximum)
  - Wind speed
  - Solar radiation
  - Water vapour pressure

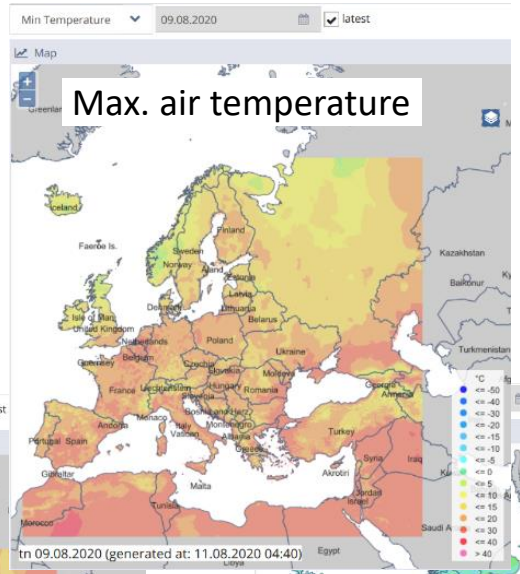
Emergency Management

Interpolated grids for LISFLOOD model

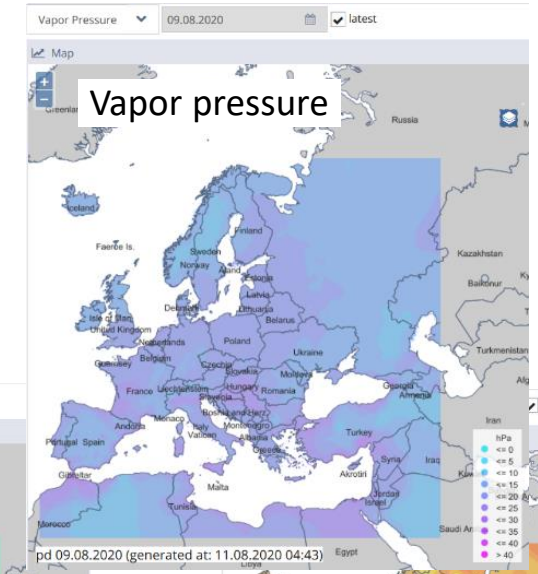
- Daily grids produced by CEMS-MDCC



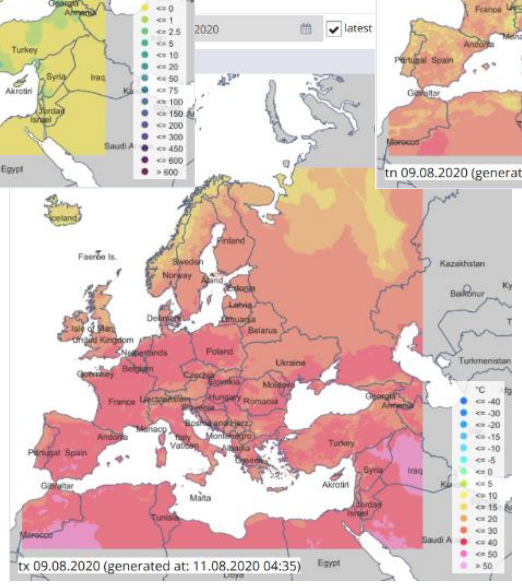
Precipitation



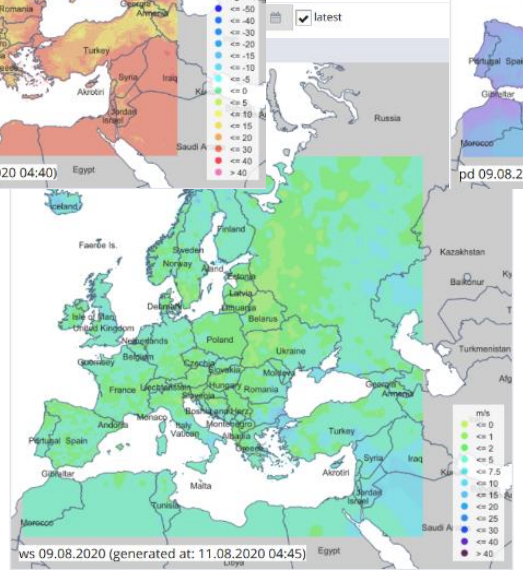
Max. air temperature



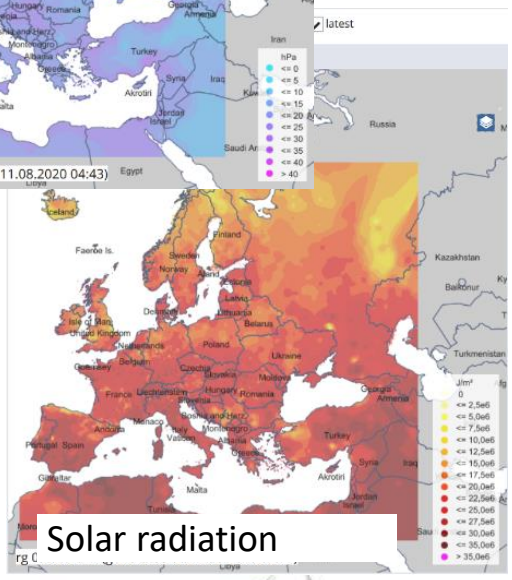
Vapor pressure



Min. air temperature



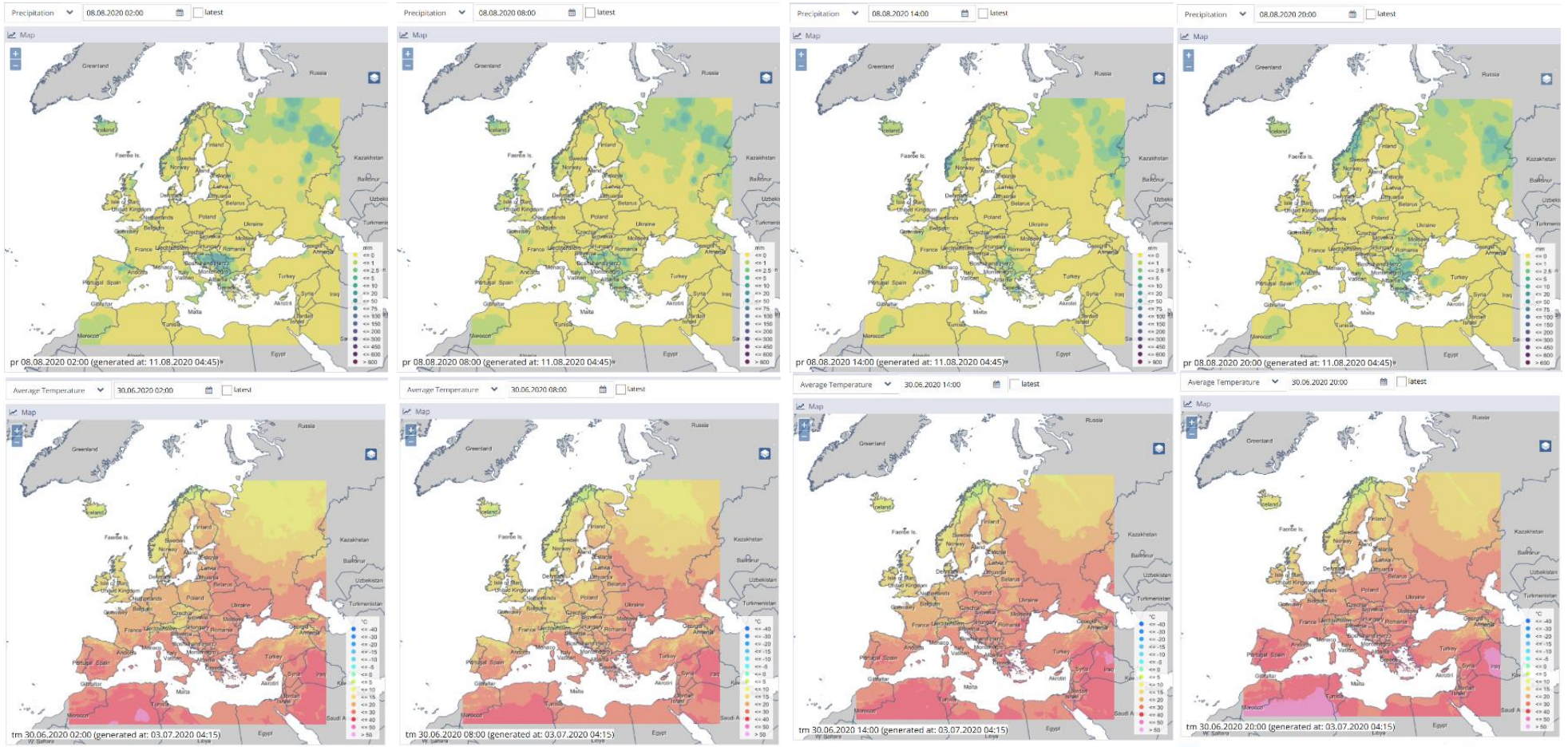
Wind speed



Emergency Management

Interpolated grids for LISFLOOD model

- 6-hourly precipitation and average temperature grids produced by CEMS-MDCC



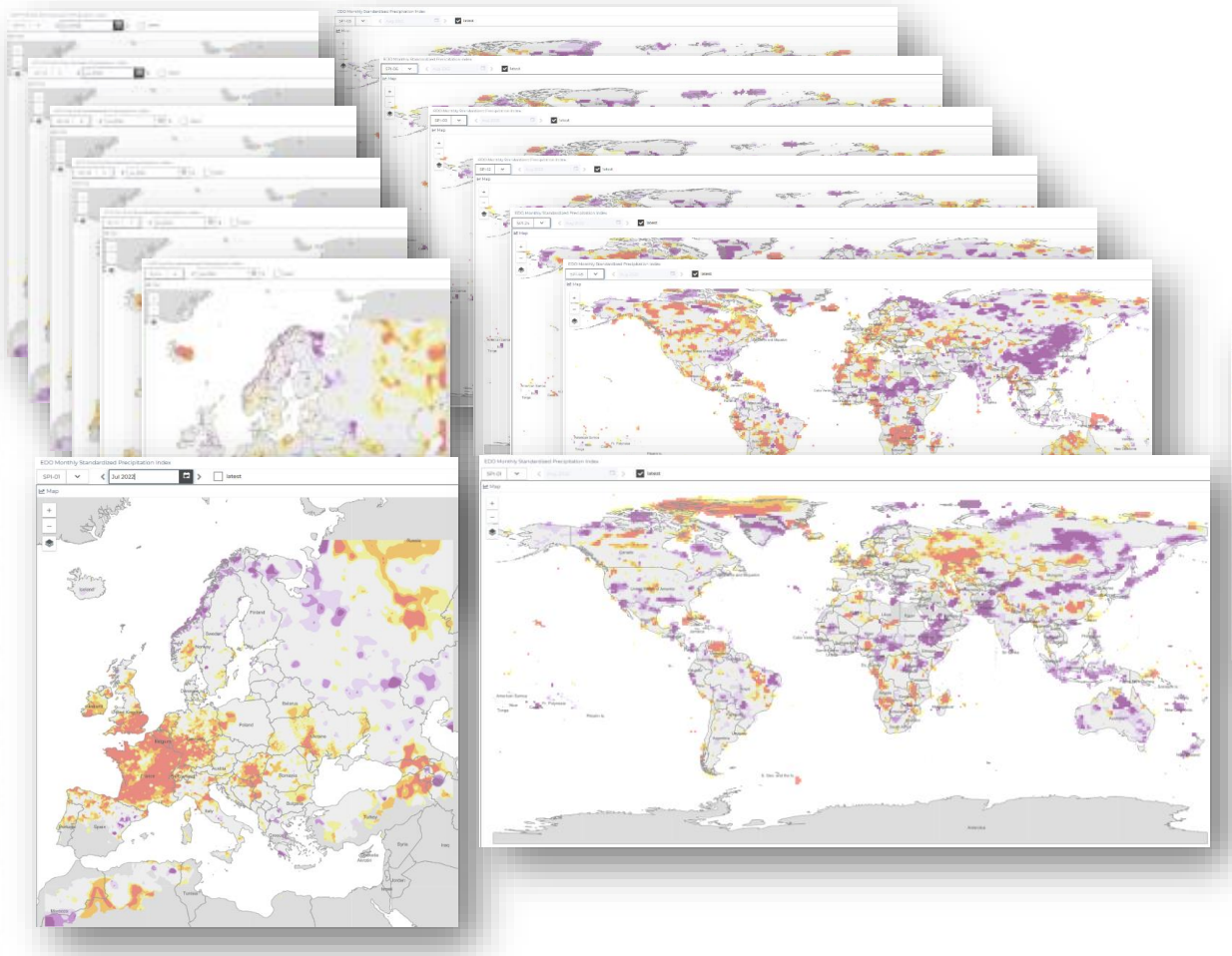
Copernicus Europe's eyes on Earth

European Commission

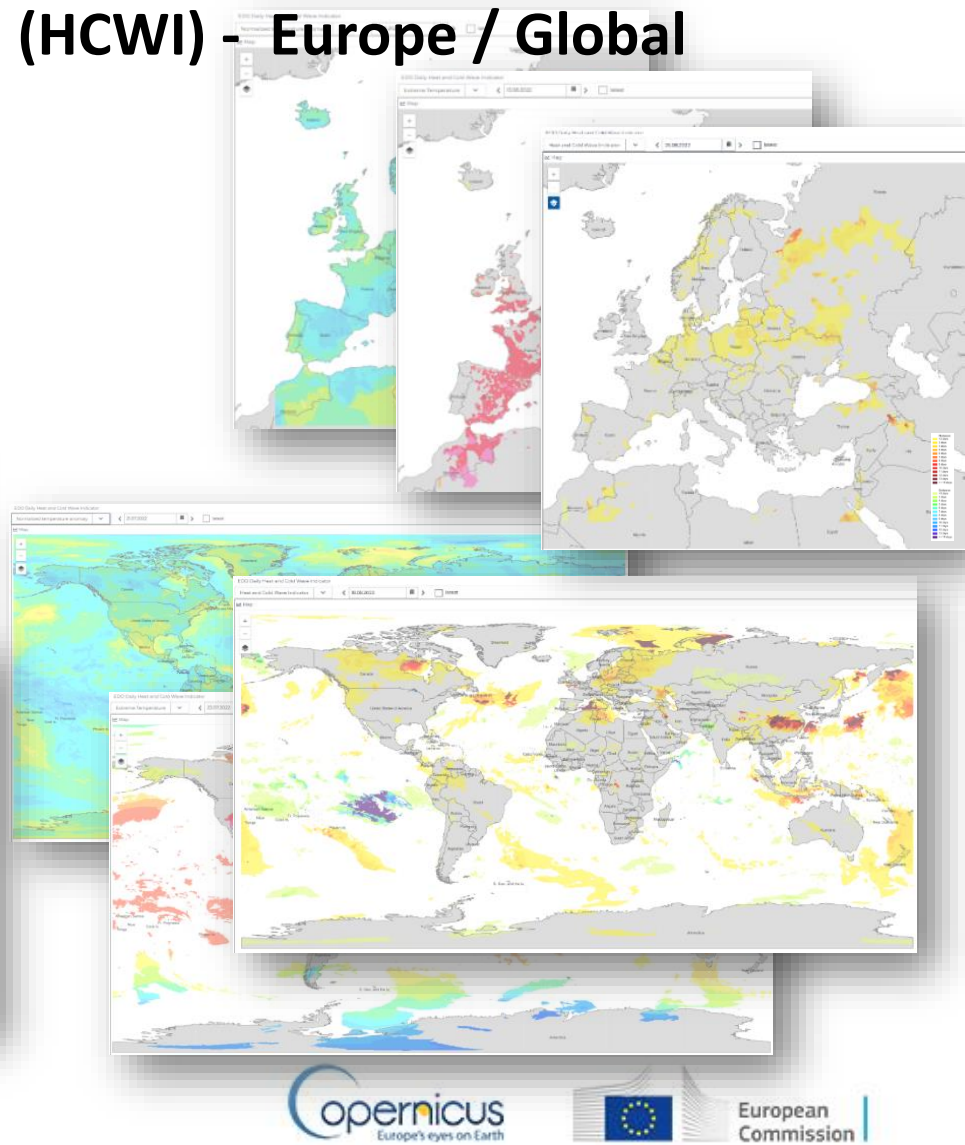
Additional MDCC Products

Emergency Management

Standardized Precipitation Index (SPI) - Europe / Global



Heat and Cold Wave Indicator (HCWI) - Europe / Global



www.efas.eu

Copernicus Europe's eyes on Earth

European Commission

How to contribute to CEMS-MDCC?

Emergency Management

Meteorological station data

- As many parameters as possible:
  - **Precipitation**
  - **Relative humidity**
  - **Solar radiation**
  - **Temperature (including min/max)**
  - **Water vapour pressure**
  - **Wind speed**
  - Cloud cover
  - Dew-point temperature
  - Evaporation
  - Sunshine duration
  - Wind direction
- Time-period: near real-time but also historic data from 1 January 1970 to present (if available)
- Time-resolution: at least 6-hourly, if possible hourly or better
- **Station metadata**: number/ID, name, coordinates, elevation, instrument type (optional)
- Usage in accordance to EUMETNET-Copernicus data license

Emergency Management

Contact details

Email: efas.mdcc@dwd.de

Markus Ziese (CEMS-MDCC/DWD), phone: +49 69 8062 2973

For questions in relation to the data license:

Peter Salamon (JRC Ispra), phone: +39 0332786013,

email: peter.salamon@ec.europa.eu

We are looking forward to work with you and your data!

Further Information

- LISFLOOD (hydrological model):
  - https://github.com/ec-jrc/lisflood-code

- Gridded data:
  - https://data.jrc.ec.europa.eu/dataset/0bd84be4-cec8-4180-97a6-8b3adaac4d26#dataaccess


Preparedness
Response
Recovery
Rapid Mapping
Risk & Recovery Mapping
Floods
Fires
Droughts
Population
Built-up areas