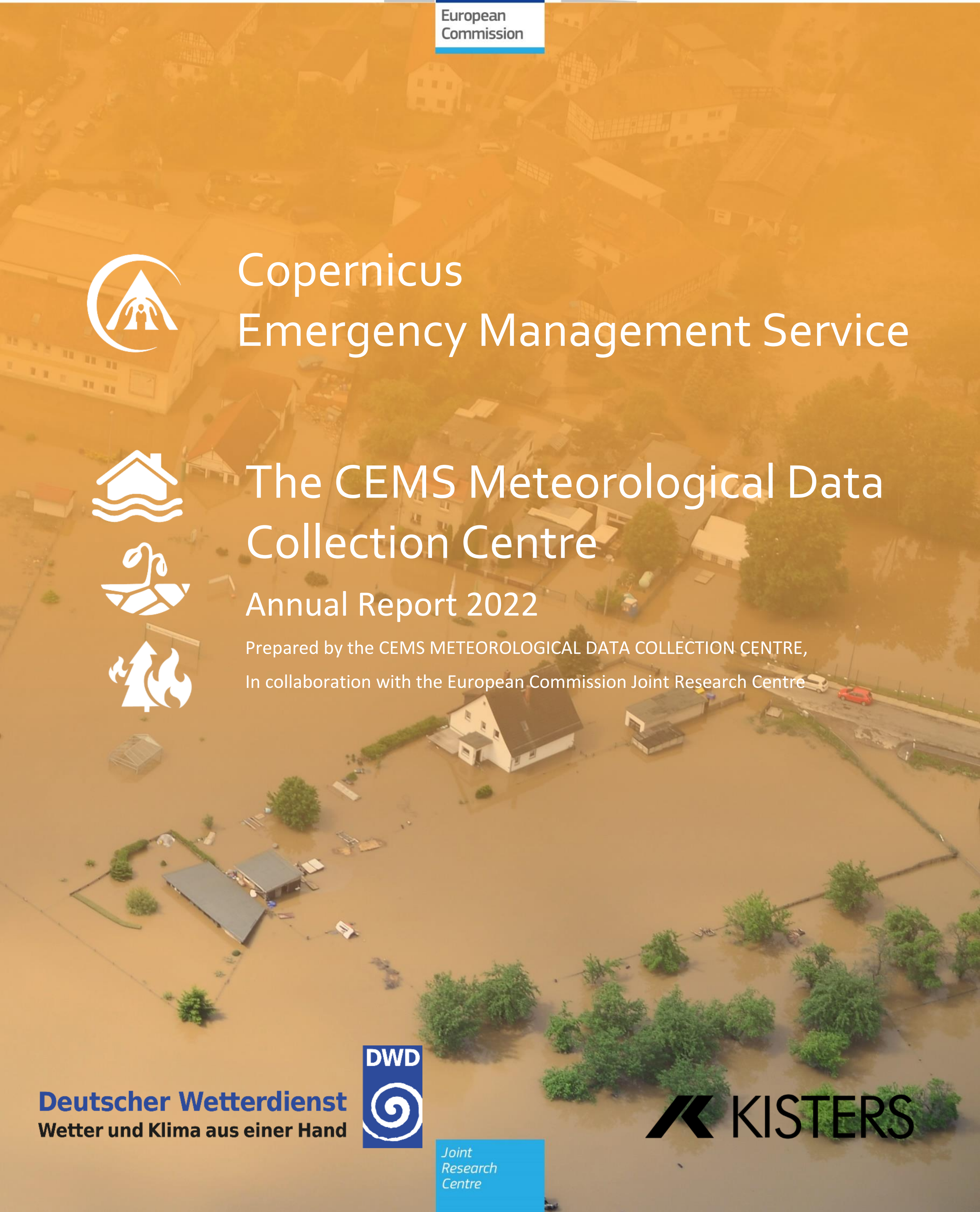
European
Commission



Copernicus
Emergency Management Service

The CEMS Meteorological Data Collection Centre

Annual Report 2022

Prepared by the CEMS METEOROLOGICAL DATA COLLECTION CENTRE,
In collaboration with the European Commission Joint Research Centre

Deutscher Wetterdienst
Wetter und Klima aus einer Hand
DWD

KISTERS

Joint
Research
Centre

This publication is a Technical report by the Copernicus Emergency Management Service - Meteorological Data Collection Centre (operated by Kisters AG and Deutscher Wetterdienst) in collaboration with the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The contents of this publication do not necessarily reflect the position or opinion of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC134489

Ispra: European Commission, 2023

© European Union, 2023

The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (https://creativecommons.org/licenses/by/4.0/). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

How to cite this report: Lemke, C., Schirmeister, Z., Walawender, J., Ziese, M., Pichon, D., Radke-Fretz, M., Schweim, C., Gomes, G., Grimaldi, S. and Salamon, P., CEMS Meteorological Data Collection Centre – Annual report 2022, European Commission, 2023, JRC134489

Contents

Abstract

The Copernicus Emergency Management Service (CEMS) Meteorological Data Collection Centre (METEO) collects, quality-controls and post-processes in situ and ground-based remote sensed meteorological data to provide input data tailored to the needs of CEMS European Flood Awareness System (EFAS), European Forest Fire Information System (EFFIS), European and Global Drought Observatory (EDO and GDO).

All data are quality-controlled. Only the data that passed this check are processed further as reliable input data for the CEMS. Data post-processing includes the calculation of minimum, maximum and mean values as well as the aggregation of totals over different accumulation periods. Data are then provided to CEMS as station lists or grids.

This report provides an overview of the data collection, quality control, and post-processing activities completed during the year 2022. In 2022, METEO collected real-time data for 11 parameters from 35 data providers and around 22,000 stations. On average, 9,000,000 records were added to the database each day. Furthermore, the database included more than 31,500 stations with historical data.

METEO constantly strives to improve the configuration of the database to enhance the product generation. For this purpose, in 2022, METEO developed a new set of quality checks (the operational implementation will start in 2023). Finally, an analysis on data availability led to a proposal for future data collection strategy.

Acknowledgements

The EFAS team, and particularly METEO, would like to thank the EFAS Partners and Data Providers that contributed to the CEMS meteorological data collection. We would like to acknowledge their dedication to the EFAS project, their commitment and the sharing of their meteorological data. We thank them for their cooperation with METEO, both in the provision of data and for their proactive role in responding to the questions and solving issues. Without their collaboration, the delivery of this report would not be possible.

1 Introduction

The Meteorological Data Collection Centre (METEO) was established to provide application tailored quality-controlled meteorological surface data to the Copernicus Emergency Management Services (CEMS). It is operated since 2016 by the KISTERS AG and Deutscher Wetterdienst (DWD). Current CEMS components served by CEMS METEO are the European Flood Awareness System (EFAS), the European Forest Fire Information System (EFFIS), the European and Global Drought Observatory (EDO and GDO). To fulfil this task, CEMS METEO collects in situ meteorological observations as well as ground-based remote-sensed data like radar observations of precipitation within the EFAS Extended domain (Arnal et al., 2019). The EFAS Extended domain covers the European continent and it includes also parts of North Africa and Middle East (to include more river basins located near the southern and eastern borders of Europe). When using the INSPIRE compliant ETRS89 Lambert Azimuthal Equal Area Coordinate Reference System (ETRS-LAEA, the boundaries of the EFAS Extended domain are top: 5500000 (including Scandinavia), left: 2500000 (including the Iberian Peninsula), right: 7500000 (including Turkey), bottom: 750000 (including coastal catchments of the southern shore of the Mediterranean Sea).

The data are received from many data providers using various channels, e.g. ftp-server, APIs or other webservices delivering different file formats. All received data are quality controlled and integrated into a database before they are then post-processed to the needs of the CEMS applications. This includes the calculation of minimum, maximum and mean values as well as the aggregation and disaggregation of totals. Depending on the component, the data are either provided as station data or gridded fields.

This annual report is intended to provide a general overview of METEO's approach and to highlight changes and improvements that happened in 2022.

2 Data Providers and Provision

An overview about the current data providers and which data they provide to CEMS METEO is given in this section.

2.1 Data Providers, Parameter and Time Resolution

By the end of 2022, 35 data providers deliver real-time data to CEMS METEO. The data providers and delivered parameters are summarised in Table 1.

Table 1. List of active real-time data providers and delivered parameters (abbreviation). Parameter abbreviations are explained in Table 2.

Name	Parameter										
	ClCo	**DT**	**Evap**	**Precip**	**ReAiHu**	**SunRad**	**SunD**	**AT**	**VP**	**WDir**	**WSpeed**
Agencia Estatal de Meteorología (Spain)	-	x	-	x	x	x	x	x	-	x	X
Agenzia Regionale Di Protezione Civile (Lazio, Italy)	-	-	-	x	x	x	-	x	-	x	X
Agenzia Regionale per la Prevenzione e l'Ambiente dell'Emilia-Romagna (Italy)	-	-	-	x	-	-	-	x	-	-	-
Automatic System of Hydrological Information (SAIH) for the Ebro river basin (Spain)	-	-	-	x	x	x	-	x	-	x	X
Czech Hydro-Meteorological Institute	-	-	-	x	-	-	-	x	-	-	-
Danish Meteorological Institute (Denmark & Greenland)	x	x	-	x	x	x	x	x	-	x	X
Deutscher Wetterdienst (Germany)	x	x	-	x	-	-	x	x	-	x	X
Deutscher Wetterdienst (Global)	x	x	-	x	x	x	-	x	-	x	X
Environment Agency (England)	-	-	-	x	-	-	-	-	-	-	-

Name	Parameter										
	ClCo	DT	Evap	Precip	ReAiHu	SunRad	SunD	AT	VP	WDir	WSpeed
Federal Hydrometeorological Institute (Bosnia-Herzegovina)	-	-	-	x	-	-	-	x	-	-	-
Finnish Meteorological Institute	x	x	-	x	x	-	-	x	-	x	X
Flanders Environment Agency (Belgium)	-	x	x	x	x	-	-	x	-	x	X
Hungarian Meteorological Service	-	-	-	x	x	x	-	x	-	x	X
Hydrological Information Centre (HIC) - Flanders Hydraulics Research (Belgium)	-	-	-	x	-	-	-	-	-	-	-
ICON Numeric Weather Prediction model (Ukraine) (see section 2.2)	-	x	-	x	x	x	-	x	-	x	X
Institute for Ocean and Atmosphere (Portugal)	-	x	-	x	x	x	-	x	-	x	X
Institute of Meteorology and Water Management (Poland)	x	x	-	x	-	x	x	x	-	x	X
Israel Meteorological Service	x	x	x	x	x	x	x	x	-	x	X
Kosovo Hydrometeorological Institute	-	-	-	x	x	x	-	x	-	x	X
Met Éireann (Ireland)	-	-	-	x	-	-	-	x	-	x	X
MeteoFrance	-	x	-	x	x	x	x	x	-	x	X

Name	**Parameter**										
	ClCo	**DT**	**Evap**	**Precip**	**ReAiHu**	**SunRad**	**SunD**	**AT**	**VP**	**WDir**	**WSpeed**
MeteoLux (Luxembourg)	x	x	-	x	x	-	x	x	-	x	X
MeteoSchweiz (Switzerland)	x	x	x	x	x	x	x	x	-	x	X
MeteoSchweiz CombiPrecip (Switzerland)	-	-	-	x	-	-	-	-	-	-	-
Monitoring Agricultural Resource system from JRC, European Commission (MARS; Global)	x	-	-	x	-	x	-	x	x	-	X
National Environmental Agency (Georgia)	x	x	-	x	-	-	-	x	-	x	X
Norwegian Meteorological Institute	x	x	-	x	x	-	-	x	-	x	X
Republic Hydrometeorological Service of the Republic of Srpska (Bosnia and Herzegovina)	-	-	-	x	-	-	-	x	-	-	-
Royal Netherlands Meteorological Institute	x	x	-	x	x	x	x	x	-	x	X
Servei Meteorològic de Catalunya (Spain)	-	-	-	x	x	x	-	x	-	x	X
Service public de Wallonie (Belgium)	-	-	-	x	-	-	-	-	-	-	-
Slovak Hydro-Meteorological Institute	-	-	x	x	-	-	-	x	-	-	-
Slovenian Environment Agency	-	-	-	x	x	x	-	x	-	x	X

Name	Parameter										
	ClCo	**DT**	**Evap**	**Precip**	**ReAiHu**	**SunRad**	**SunD**	**AT**	**VP**	**WDir**	**WSpeed**
Swedish Meteorological and Hydrological Institute	x	-	-	x	x	x	x	x	-	x	X
Zentralanstalt für Meteorologie und Geodynamik (Austria)	-	-	-	x	-	-	-	x	-	-	-

Source: METEO

Table 2. Parameter abbreviation and description.

Parameter abbreviation	Parameter description
ClCo	Cloud cover
DT	Dew point temperature at 2 m above ground (WMO)
Evap	Evaporation
Precip	Precipitation
ReAiHu	Relative Air humidity
SunRad	Solar radiation
SunD	Sunshine duration
AT	Air temperature at 2 m above ground (WMO)
VP	Water vapour pressure
WDir	Wind direction
WSpeed	Wind speed at 10 m above ground (WMO)

Source: METEO

In order to provide reliable maps for historical periods, used for example as input datasets to calibrate the hydrological model of EFAS, data providers were asked to deliver also historical data back to 1970 if possible. Some data providers provided only historical data. Historical data were also retrieved from research projects and gridded data sets. In case a country or region provided only historical data, these are covered with real-time data by the two global data deliveries from MARS[1] and DWDSynop[2] (global). Table 3 shows all sources from which only historical data are available.

Table 3. List of data providers with only historical data and delivered parameters (abbreviation). Parameter abbreviations are given in Table 2.

Name	**Parameter**										
	ClCo	**DT**	**Evap**	**Precip**	**ReAiHu**	**SunRad**	**SunD**	**AT**	**VP**	**WDir**	**WSpeed**
CarpatClim	-	-	-	x	-	-	-	-	-	-	-
Danube*	-	-	-	x	-	-	x	x	-	-	-
DWD Climatic*	x	-	-	x	x	-	x	x	-	-	-
ECA*	-	-	-	x	-	-	-	-	-	-	-
ERA-Interim-land	-	-	-	x	-	-	-	-	-	-	-
Euro Synop*	x	x	-	x	-	-	-	x	-	x	x
EURO4M-APGD by MeteoSchweiz (Switzerland)	-	-	-	x	-	-	-	-	-	-	-
Hellenic National Meteorological Service (Greece)	x	-	-	x	x	-	x	x	x	-	x
MeteoConsult*	-	-	-	x	-	-	-	-	-	-	-

* historic data received from JRC

A brief explanation of the main sources (research projects and gridded datasets) of historical data is provided below.

- Source: METEOCarpatClim is a historical daily precipitation product, including 2946 virtual stations covering the area between 44 and 50◦ E as well as 17 to 27◦ N (Hungary, Serbia, Romania, Ukraine, Slovakia, Poland, Czech Republic, Croatia) at a horizontal resolution of 0.1◦ (Spinoni et al., 2015), from 1970 to 2010.
- Danube is a dataset of historical observations covering the years 1990-2003 and collected by the JRC from a number of national and regional hydro-meteorological institutes of the Danube basin (e.g. Romania National Institute for Hydrology and Water Management, Czech Hydro-Meteorological Institute , Bulgaria National Institute of Meteorology and Hydrology, Serbia Institute for Development of Water Resources Institute, Meteorological and Hydrological Service for the Republic of Croatia).
- DWD Climatic Historical is a dataset made available from the Deutscher Wetterdienst (DWD) Climate Data Centre (Wetter und Klima - Deutscher Wetterdienst - Climate monitoring (dwd.de)).

---

[1] https://marswiki.jrc.ec.europa.eu/agri4castwiki/index.php/Meteorological_data_from_ground_stations#Calculation_of_advanced_parameters

[2] Copy of global data exchange between Meteorological Services, distributed via WMO Information System (WIS)

- ECA – European Climate Assessment dataset contains series of daily observations at meteorological stations throughout Europe and the Mediterranean area. More information is available from https://www.ecad.eu/dailydata/
- ERA-Interim-Land is a global multi-variable reanalysis dataset for all land surface areas from 1979 to October 2016 at a 0.75◦ × 0.75◦ spatial resolution (Dee et al., 2011)
- Euro Synop is a collection of data from manned and unmanned meteorological stations, as well as from mobile stations. SYNOP (surface synoptic observations) is a numerical code used for reporting weather observations.
- EURO4M-APGD is a daily precipitation dataset from 1971 to 2008 covering the Alps and adjacent flatlands (area between 4.8 and 17.5◦ E as well as 43 to 49◦ N) (Isotta et al., 2014).
- Meteo Consult is a 6 hourly precipitation dataset from 2007 to 2015 produced by Meteo Consult (part of MeteoGroup).

Data are provided with various temporal resolutions and aggregation intervals, depending on the parameter and data provider. The highest received temporal resolution and accumulation period is one minute (air temperature and precipitation from one data provider). Instantaneous parameters, like air temperature or wind speed, are mainly provided with temporal resolution of one, three and six hours, but also with higher and lower temporal resolutions. The majority of provided precipitation totals are accumulated over six and twelve hours, but also daily and one hourly total are often provided. Minimum and maximum air temperature are mostly provided with daily aggregation.

2.2 New Data Providers in 2022

A continuous task for CEMS METEO is the acquisition and integration of additional data providers for real-time and historical data. This is done for two reasons: (1) to get more real-time information into the products and to increase reliability of the grids and (2) to enlarge the database with additional historical data needed for the calibration and validation of the hydrological model.

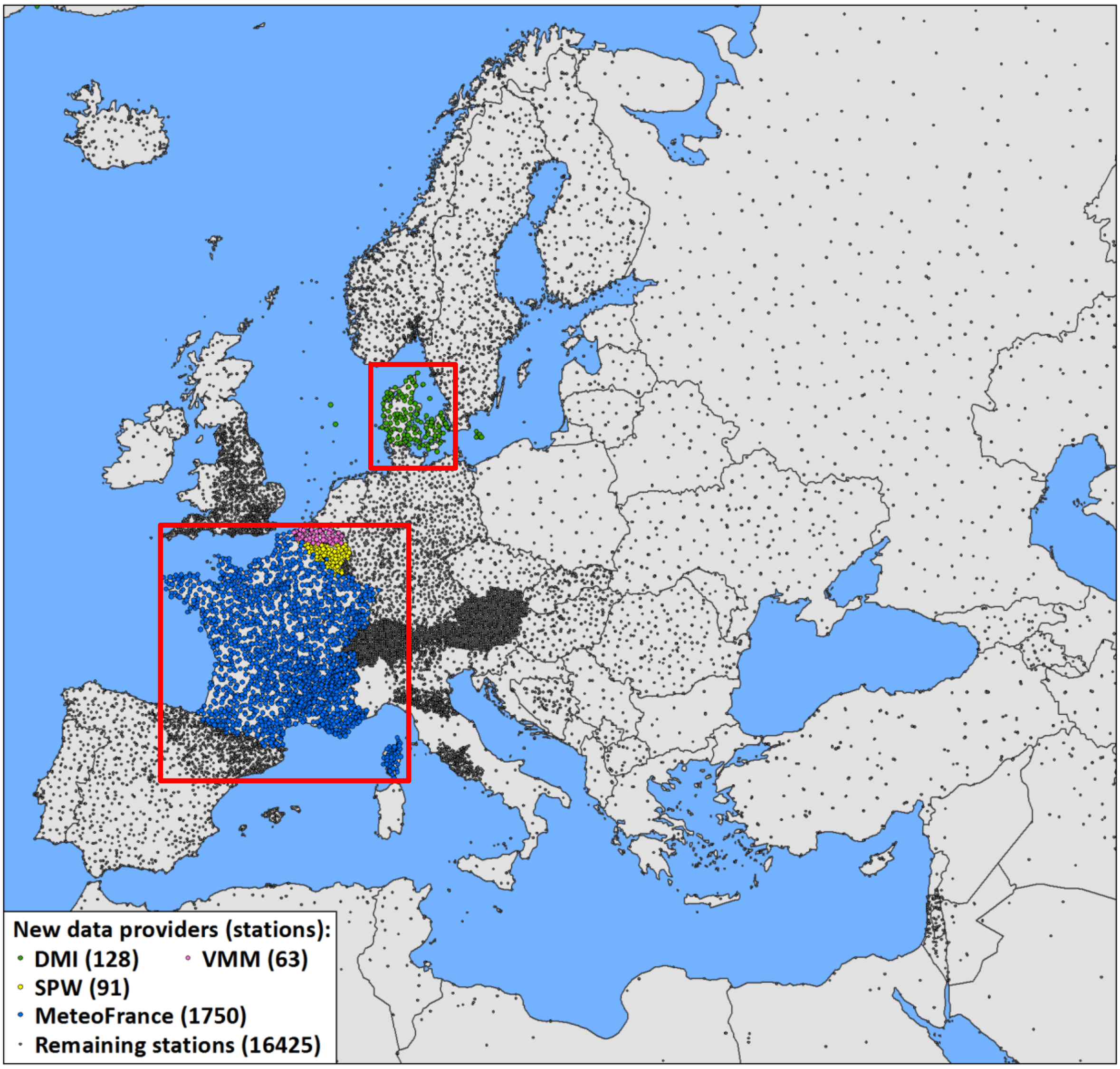
To enlarge the CEMS METEO database, the following data providers were added within 2022:

- Météo-France
- Service public de Wallonie (SPW; Belgium)
- Flanders Environment Agency (VMM; Belgium)
- Danish Meteorological Institute (DMI; Denmark & Greenland)

The war in Ukraine caused instabilities in the delivery of in situ data. Missing meteorological data for an area as large as the Ukraine for a long time period could lead to degradation of the EFAS forecasts. The European Commission Joint Research Centre, which is the entrusted entity responsible for CEMS EFAS, the CEMS Hydrological Forecast Computational Centre, which is responsible for the operational computations of EFAS, and METEO agreed on a temporary solution to prevent a consistent degradation of the EFAS forecasts. Specifically, 3371 gridded stations of DWD’s numerical weather prediction model ICON (Icosahedral Nonhydrostatic) were included in the operational processing routines as a temporary complement to the in situ stations (Figure 3).

The station density is shown in Figure 1. The inclusion of Météo-France, SPW, VMM, and DMI contributed to increase the density of in situ stations compared to the previous year. It is here noted that two of the new data providers, namely SPW and VMM, also deliver historical data.

Figure 1. Additional stations due to the integration of new data providers in comparison to the existing station network. Red rectangles mark the regions with the new data providers. The number of stations for each new data provider is given in brackets.



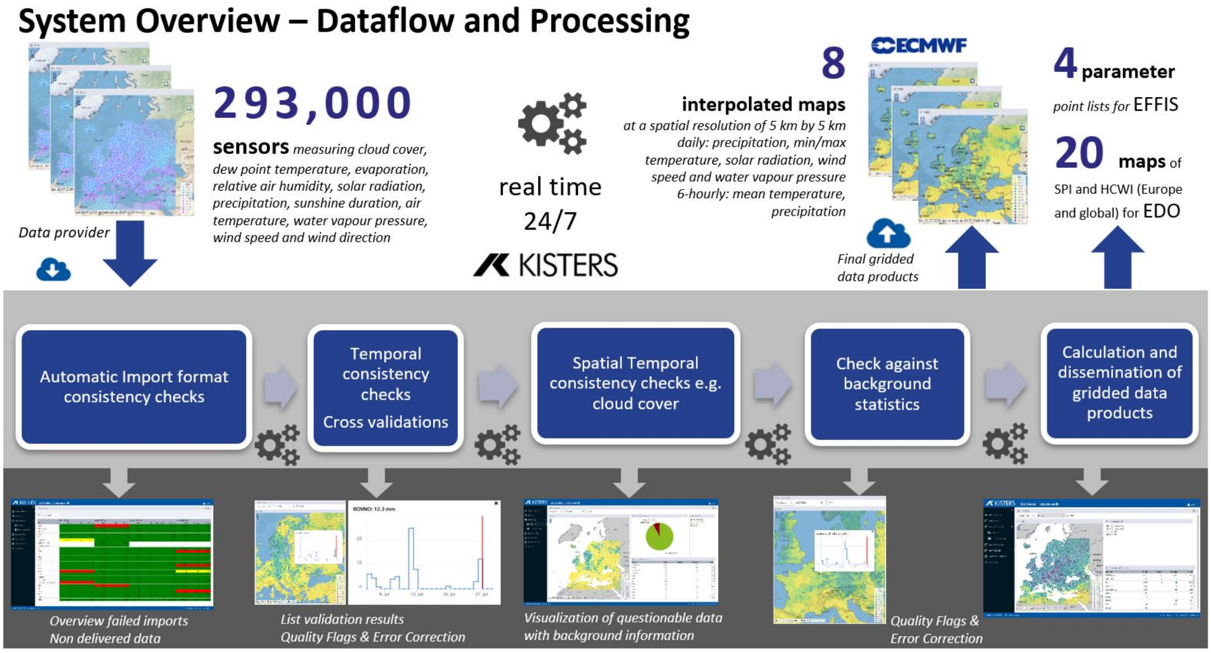
Source: METEO

3 Database

3.1 Data Flow

The data flow within CEMS METEO is illustrated in Figure 2. Real-time data are delivered by the data providers via (s)ftp-servers (pulled by METEO/pushed to METEO), web services/APIs and email attachments in provider specific file formats. All files are converted into a uniform and optimised file format for the integration into the METEO database.

Figure 2. Schematic illustration of the data flow within METEO. The amount of input and output data are shown as well as the processing and quality control steps within the data bank system.



System Overview – Dataflow and Processing

293,000 sensors *measuring cloud cover, dew point temperature, evaporation, relative air humidity, solar radiation, precipitation, sunshine duration, air temperature, water vapour pressure, wind speed and wind direction*

8 interpolated maps *at a spatial resolution of 5 km by 5 km daily: precipitation, min/max temperature, solar radiation, wind speed and water vapour pressure 6-hourly: mean temperature, precipitation*

4 parameter *point lists for* EFFIS

20 maps of SPI and HCWI (Europe and global) for EDO

Source: METEO

The provided data are stored in a data bank super-system, named WISKI. Within WISKI, the quality control (see section 3.2) of the data and necessary aggregations (see section 4.1) are done. Finally, the data are extracted from the database as input for the interpolated maps and station lists provided to the CEMS components. The interpolation procedure is described in detail in section 4.2.

3.2 Quality Control

The following section covers the quality control as it is currently carried out in the operational EFAS routines and the newly implemented routines of the new validation framework along with the new seasonal and geographical thresholds.

3.2.1 Operational Quality Control in EFAS

Although the data is usually quality controlled by the data providers, an independent quality control procedure was established by CEMS METEO based on the experience that real-time data contain erroneous data points from time to time. This applies to historical data too. Data providers are regularly informed about detected errors, this feedback is an added value correlated to the data provision to METEO. The quality control procedure is triggered every time by the import of real-time, historical and redelivered data. The latest refers to data which are re-sent by data providers in order to either fill in missing data points or replace already existing data in the database. This guarantees the availability of quality checked data in the database used as input for the interpolated maps and station lists. If needed, a quality control is repeated on the aggregated data (see section 4.1). This is necessary, as for example the twelve hourly precipitation threshold is not twice the six hourly thresholds (Table 5).

Quality flags are added to each data record in the quality control procedure. The following flags are in use: "good" if the value passes all quality checks, "suspect", if it is inconsistent with other parameters (e.g. dew point temperature higher than air temperature) and "rejected", if it did not pass at least one of the threshold checks. Additionally, a quality flag was defined for missing values in the time series. Moreover, it is here noted that the quality flag "suspect" is also added to data points shifted in time. For example, some stations provide six-, twelve-hourly and daily precipitation totals outside the needed time steps at 00, 06, 12 and 18 UTC. Such totals are shifted to the nearest needed date, as by doing so the uncertainty is lower compared to splitting such data into hourly totals and then aggregate them.

The quality control is mainly based on fixed thresholds as shown in Table 4. At the current stage, the same threshold values (minimum and maximum) per parameter is used for the whole EFAS-Domain. Climate zone-specific thresholds were implemented in 2022 and will enter the operational process in 2023. Other improvements in the data validation were also introduced last year, which will be described in section 6. Additionally, cross-validation procedure against data from other parameters at the same station is carried out.

Table 4. Parameters with fixed thresholds. For precipitation see Table 5.

Parameter	Min. threshold	Max threshold
Cloud cover (ClCo)	0	9 octas
Evaporation (Evap)	0	2 mm/15 min, 15 mm/day or 3 mm/hour
Relative air humidity (ReAiHu)	5	100 %
Solar radiation (SunRad)	0	1360 cos(lat) W/m^2
Sunshine duration (SunD)	0	Astronomic max
Water vapour pressure (VP)	0	35 hPa
Wind direction (WDir)	0	360 deg
Wind speed (WSpeed)	0	45 m/s

Source: METEO

The thresholds for precipitation depend on the aggregation period (Table 5).

Temperature data are checked against time-dependent thresholds taking the annual cycle into account. In winter, only data between -50°C and 25°C are accepted while the thresholds in summer are -10°C and 55°C.

Dew point temperature is checked against the air temperature. If the dew point temperature is 30°C below the air temperature or 0.2°C above the air temperature, then the dew point temperature is flagged as "suspect".

In an effort to constantly improve the routines and deliverables for the CEMS, METEO implemented two new data validation procedures in 2022, of which the operational use will start in 2023. More Information will be given in section 6.

Table 5. Thresholds for precipitation depending on the aggregation interval.

Aggregation interval [min]	Max precipitation threshold [mm]
15	125
30	200
60	250
180	350
360	425
540	475
720	500
900	525
1080	550
1440	600

Source: METEO

3.2.2 Quality Control Statistics

Data are controlled and flagged as described in section 3.2.1. Table 6 shows the summary for the parameters used for gridding (EFAS) and station lists (EFFIS).

Table 6. Quality control statistics of the quality codes “good”, “suspect” and “rejected” for each parameter, given as relative and absolute values in 2022

		Quality					
		good		**suspect**		**rejected**	
	Parameter	**relative [%]**	**absolute**	**relative [%]**	**absolute**	**relative [%]**	**absolute**
EFAS Gridding	Temperature (6hourly mean) (AT)	97,64%	30198155	0,00%	214	2,36%	730004
	Temperature (daily maximum) (AT)	97,04%	7576068	0,00%	87	2,96%	231348
	Temperature (daily minimum) (AT)	98,06%	7535394	0,00%	93	1,94%	149087
	Precipitation (6hourly totals) (Precip)	91,62%	26158995	8,30%	2368703	0,08%	22627
	Precipitation (daily totals; 6 to 6 UTC) (Precip)	95,69%	7236568	4,15%	313540	0,17%	12732
	Solar radiation (SunRad)	99,15%	3145391	0,77%	24538	0,07%	2315
	Vapour pressure (VP)	100,00%	1629588	0,00%	0	0,00%	0
	Wind speed (WSpeed)	99,96%	6153946	0,00%	272	0,04%	2328
EFFIS Station lists	Precipitation (daily totals; 12 to 12 UTC) (Precip)	88,51%	6051276	11,35%	776275	0,13%	9171
	Relative air humidity (at Noon) (ReAiHu)	99,87%	3359261	0,00%	4	0,13%	4345
	Temperature (at Noon) (AT)	98,39%	7853156	0,00%	5	1,61%	128886
	Wind speed (at Noon) (WSpeed)	99,99%	6612509	0,00%	26	0,01%	501

Source: METEO

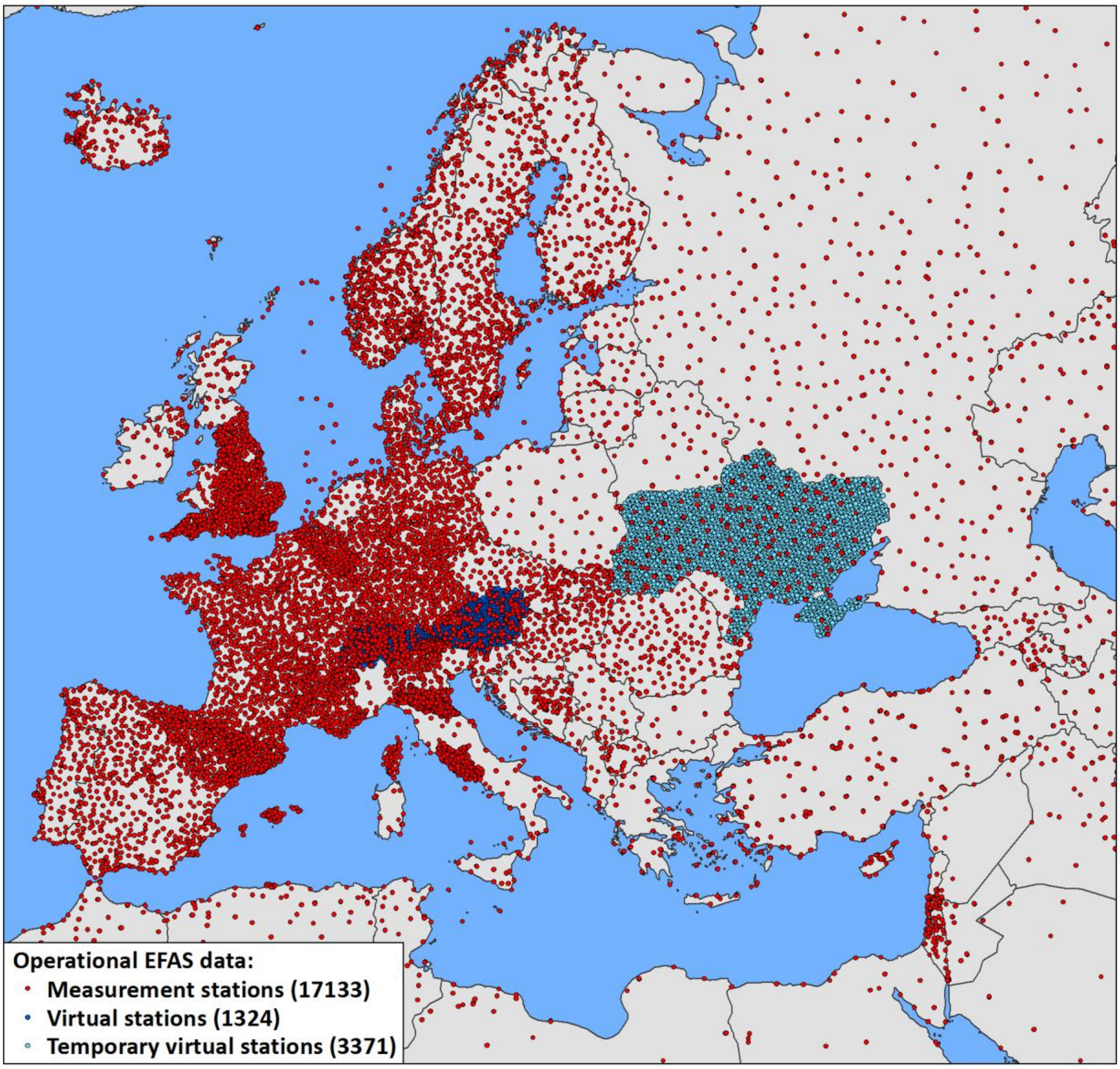
It should be noted that the quality “suspect” in precipitation is mainly due to a shift of the reporting times in the post processing to meet the CEMS component needs.

3.3 Database Statistics

By the end of 2022, the CEMS METEO database contained more than 54,500 stations, of which around 22,000 stations are delivering real-time data. The other stations provided real-time data in previous times or only historical data. 4,695 of the stations with real-time data are so-called 'virtual stations', which are extracted from high-resolution gridded data sets (e.g. station adjusted radar quantitative precipitation estimations). It is here noted that 4,695 is the sum of the 1324 virtual stations and of the 3371 temporary virtual stations. 1324 virtual stations are extracted from CombiPrecip (488 virtual stations with hourly precipitation, https://www.meteoswiss.admin.ch/weather/warning-and-forecasting-systems/nowcasting.html) and ZAMG-INCA (836 virtual stations with 6-hourly precipitation and temperature, https://www.zamg.ac.at/cms/de/forschung/wetter/inca). 3371 virtual stations from DWD's numerical weather prediction model ICON (Icosahedral Nonhydrostatic) were included in the operational processing routines in 2022 as a temporary complement to the in situ stations (see Table 1, section 2.2, and Figure 3).

CEMS METEO receives approximately 62,000 data files per day. All these files are processed, leading to on average 9,000,000 data records added to the database – per day.

Figure 3. Spatial distribution of stations delivering real-time data of at least one parameter stored in CEMS METEO database. Maps per parameter are shown in Figure 14. The number of stations is given in brackets, with the 3371 ICON stations being only an interim solution.

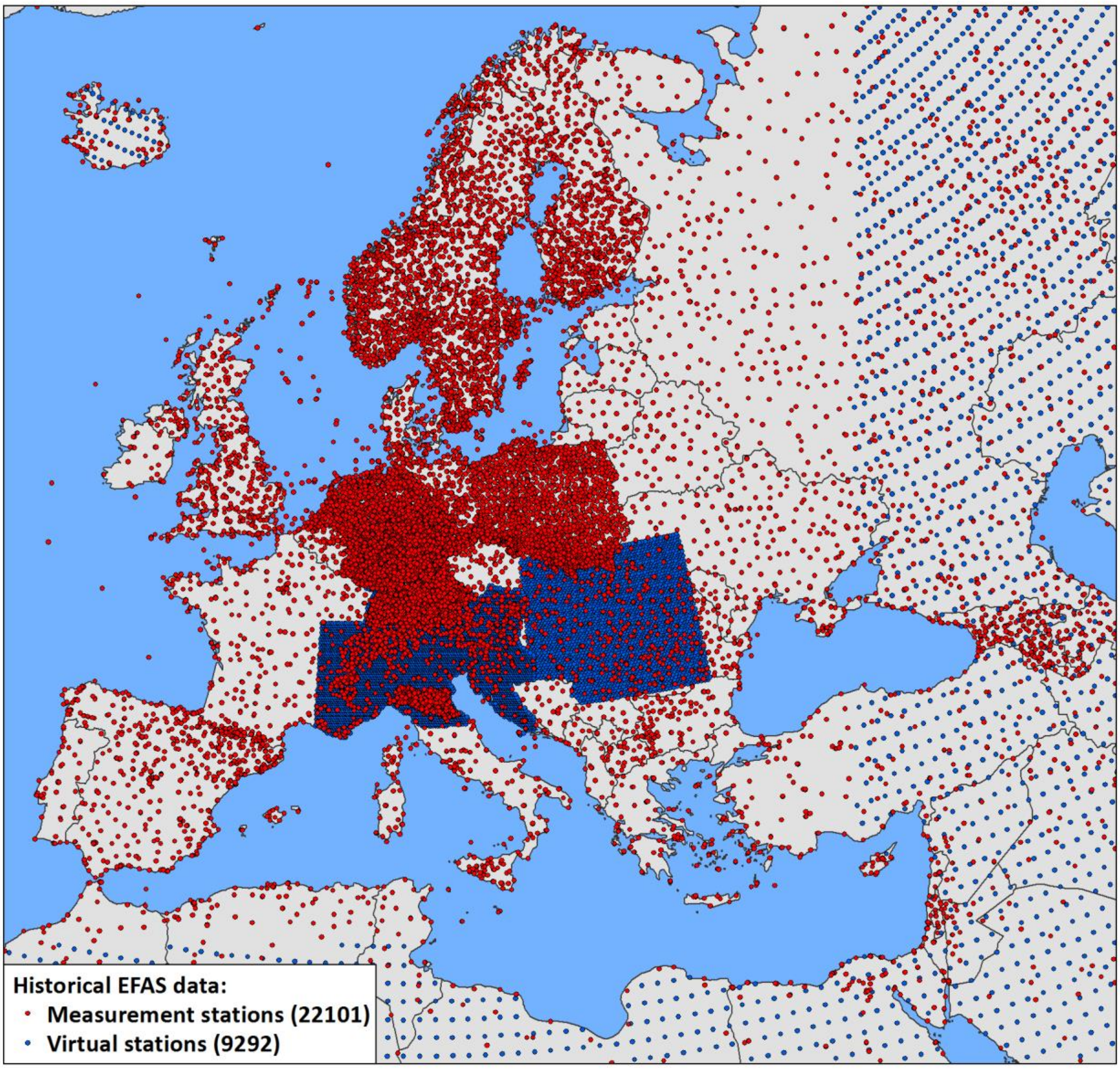


Source: METEO

The spatial distribution of active (i.e. with real-time data) stations within the EFAS-Domain is depicted in Figure 3. It is obvious that the data coverage varies between countries and even within the countries. Please note that not all active stations deliver all needed parameters, e.g. many stations only deliver precipitation and temperature. As mentioned in section 2.2, 3371 gridded stations of DWD's numerical weather prediction model ICON were included in the operational processing routines as a temporary complement to the in situ stations.

All stations currently not delivering real-time data are classified as 'inactive'. Even if these stations do not contribute to the real-time grids, they are highly valuable for historical grids: from time to time all data from the database are extracted to compute grids of historical periods (e.g. EMO-5, a high-resolution multi- variable gridded meteorological data set for Europe, Thiemig et al., 2022). Figure 4 shows the spatial distribution of the inactive stations within the EFAS-Domain. To increase the data coverage over highly complex terrain and in data sparse areas, data sets from research projects (e.g. CarpatClim and EURO4M-APGD), but also operational data sets like ERA-Interim land were integrated. Those gridded data sets are integrated as so-called 'virtual stations' on a regular grid.

Figure 4. Spatial distribution of inactive stations. These stations deliver currently no real-time data, but data of former periods. The number of stations is given in brackets.



Source: METEO

4 Post-Processing

The received data are post-processed to fulfil the needs of the various CEMS components. The post-processing routines applied are partially different depending on the application case of each CEMS. Only the quality-controlled data are used in the post-processing. It is done for the following subset of collected parameters:

- Precipitation;
- Air temperature;
- Wind speed;
- Solar radiation;
- Water vapour pressure;
- Relative air humidity.

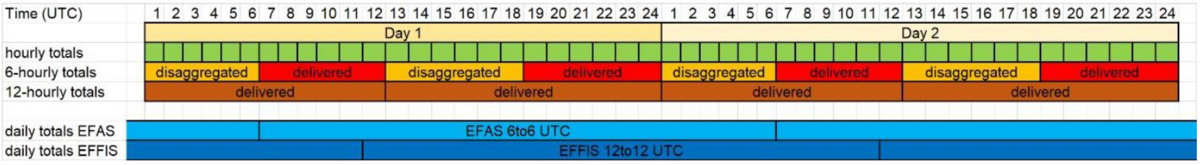
Post-processing for all parameters includes:

- Calculations of 6-hourly means/totals, except for solar radiation and water vapour pressure;
- Calculations of daily minimum, maximum and mean values, except for water vapour pressure;
- Calculation of daily totals for precipitation and solar radiation;
- Different definitions of the start and end time at a day, depending on the CEMS component;
- Aggregation and disaggregation of precipitation totals;
- Extraction of data from database;
- Spatial interpolation of station data to generate grids;
- Generation of station lists.

4.1 Aggregation and Disaggregation, Calculation of Minimum, Maximum and Mean

Precipitation totals are delivered with various accumulation periods, mainly 6 hours, 12 hours and 24 hours but also 10 minutes, 15 minutes, 30 minutes, 1 hour or 3 hours. Additionally, the reporting behaviour of the stations differs between the data providers and even within some data providers.

Figure 5. Aggregation and disaggregation scheme of precipitation totals. Disaggregated 6-hourly totals (orange) are the difference between the 12-hourly totals (brown) and the enclosed 6-hourly totals (red). Hourly totals are shown in green and can be aggregation periods.



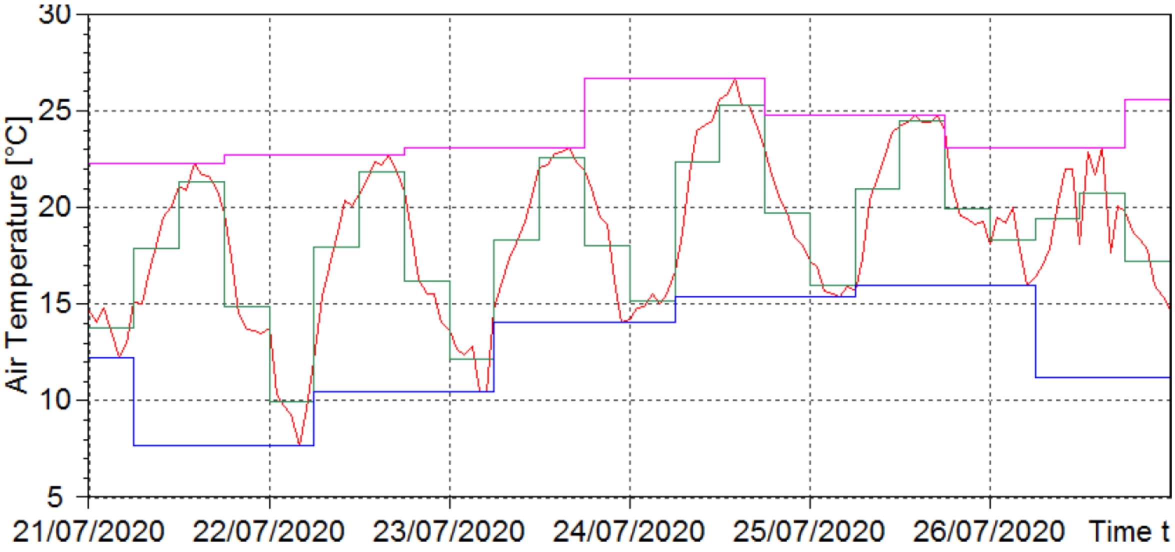
Source: METEO

To achieve a high temporal and spatial coverage of 6-hourly precipitation totals, the 12-hourly and 6-hourly totals have to be disaggregated, if the 6-hourly totals are within the 12-hourly total (Figure 5). The resulting merged 6-hourly time-series is often the basis for METEO's data deliveries to the CEMS. Where original data in higher resolution is available, 6-hourly totals are accumulated from the corresponding time-series, e.g. with 15 minutes temporal resolution.

Daily precipitation totals have to be provided as aggregated values from 6 UTC to 6 UTC of the following day and 12 UTC to 12 UTC of the next day (Figure 5). Whereas the daily totals at 6 UTC can be retrieved from the synoptic observations at 6 UTC and the higher temporal resolution time series, the daily totals at 12 UTC can only be computed from the higher time-resolution time-series.

As not all data providers deliver minimum and maximum temperatures according to the definitions of the CEMS, the required values are computed from the delivered instantaneous temperature data (Figure 6). The minimum temperature is the lowest temperature between 18 UTC and 6 UTC of the next day, whereas the maximum temperature is calculated from the observations taken between 6 UTC and 18 UTC. Also, the 6-hourly mean air temperature is calculated from the instantaneous data.

Figure 6. Delivered instantaneous air temperature (red) and therefrom calculated daily minimum temperature (blue), daily maximum temperature (magenta) and 6-hourly mean temperature (green). The magenta and blue lines of the plot allow to quickly identify the range of expected values.



Source: METEO

Additionally, daily means of wind speed and daily accumulated totals of solar radiation are calculated.

A minimum availability of data, subsequently referred to as “coverage”, is required to compute minimum, maximum and mean values as well as aggregated totals. Precipitation totals are only computed if the aggregation period is fully covered by observational data (coverage = 100%). For all other parameters, a coverage of 87% is requested. An exception is made for six and twelve hourly means with a minimum coverage of 66% to consider also stations reporting with a temporal resolution of six hours.

4.2 Gridding: Meteorological Grids for CEMS EFAS

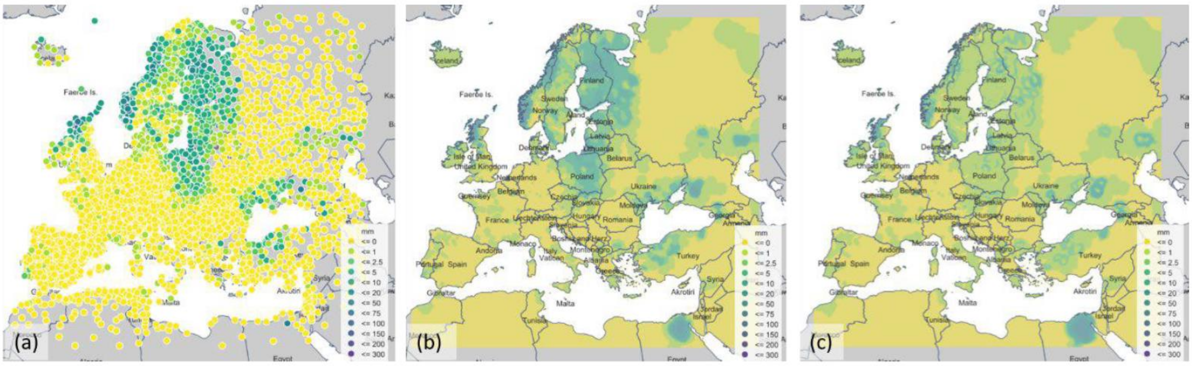
The hydrological model[3] used by EFAS requires gridded input data. Grids are generated by means of the modified SPHEREMAP (Willmott, Rowe and Philpot 1985) interpolation scheme. This is a geometric scheme, which considers the distances between the stations and the grid point as well as the clustering of stations. Additionally, an estimation of the grid reliability by means of the standard deviation (Yamamoto 2000) was implemented. This method depends on the differences between the input data and the interpolated value.

The input data and output of the modified SPHEREMAP scheme are illustrated in Figure 7 for the input stations (a), the grid itself (b) and the estimated uncertainty (c). As it can be seen, the uncertainty depends on the observed value and is higher in regions with high precipitation totals and zero in regions without precipitation. The estimated uncertainty within the grids is lower when stations are close together. A more detailed discussion about the relationship between uncertainty, measurement values, and spatial distribution of the stations network can be found in Thiemig et al. (2022).

---

[3] LISFLOOD Open Source, https://github.com/ec-jrc/lisflood-code

Figure 7. Input and output of the gridding: (a) station observations, (b) gridded data, (c) estimated uncertainty of the gridded data. The maps depict the spatial distribution of daily precipitation totals.

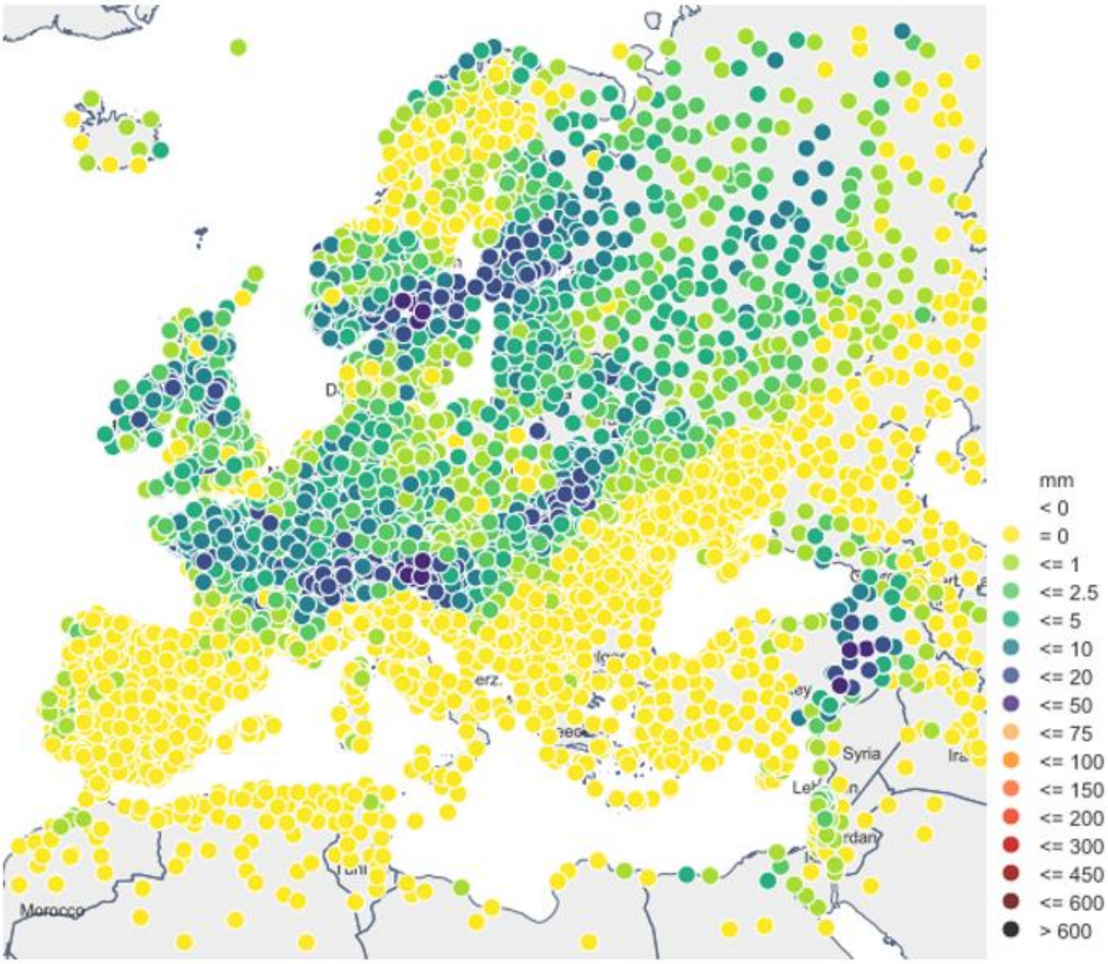


Source: METEO

4.3 Station Lists for CEMS EFFIS

The European Forest Fire Information System (EFFIS) needs observations at station level for one date in a defined region. CEMS METEO provides such a list of quality controlled post-processed data. The spatial distribution of the station data summarised in a station list is depicted in Figure 8.

Figure 8. Spatial distribution of stations summarised in a station list, as provided to EFFIS. Depicted are stations that allow the calculation of 24-hourly precipitation totals at 12 UTC (end of time stamp, as explained in section 4.1).



Source: METEO

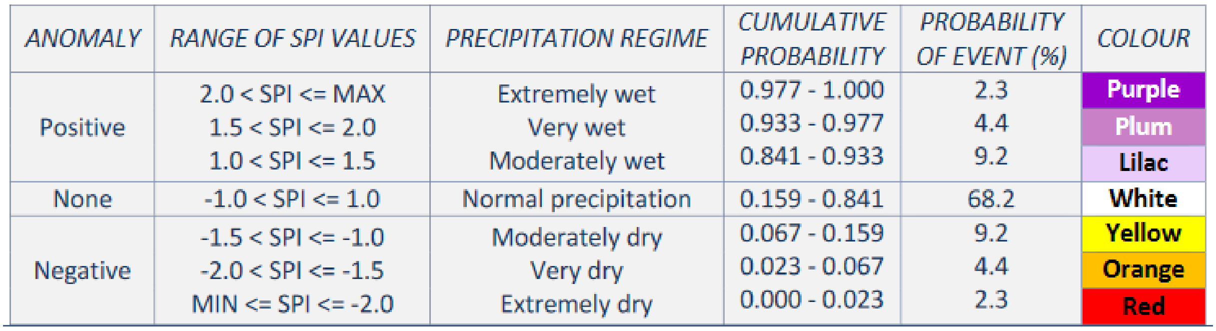
4.4 Indicators for CEMS EDO and GDO

In 2022, METEO started the computation and delivery of the Standardized Precipitation Index (SPI) and the Heat and Cold Wave Index (HCWI) to the CEMS European and Global Drought Observatory. The two indicators are described in more detail below.

4.4.1 Standardized Precipitation Index (SPI)

The Standardized Precipitation Index (SPI) is used for detecting and characterising meteorological droughts, which are defined as periods with an abnormal precipitation deficit in relation to the long-term average conditions for a region. The SPI is the most commonly used indicator worldwide and was developed by McKee et al. in 1993 and described in detail by Edwards and McKee in 1997. It compares the observed total precipitation amounts for an accumulation period of interest (e.g. 1, 3, 12 or 48 months) with the long-term historic rainfall record for that period to detect precipitation anomalies. For any given region, severe rainfall deficits (i.e., meteorological droughts) are indicated as a SPI value below –1.0, while severe excess rainfall is indicated as a SPI value above 1.0. The SPI values for any given location and accumulation period are classified into seven different precipitation regimes, as seen in Figure 9.

Figure 9. SPI classification scheme



ANOMALY	*RANGE OF SPI VALUES*	*PRECIPITATION REGIME*	*CUMULATIVE PROBABILITY*	*PROBABILITY OF EVENT (%)*	*COLOUR*
Positive	2.0 < SPI <= MAX	Extremely wet	0.977 - 1.000	2.3	**Purple**
	1.5 < SPI <= 2.0	Very wet	0.933 - 0.977	4.4	**Plum**
	1.0 < SPI <= 1.5	Moderately wet	0.841 - 0.933	9.2	**Lilac**
None	-1.0 < SPI <= 1.0	Normal precipitation	0.159 - 0.841	68.2	**White**
Negative	-1.5 < SPI <= -1.0	Moderately dry	0.067 - 0.159	9.2	**Yellow**
	-2.0 < SPI <= -1.5	Very dry	0.023 - 0.067	4.4	**Orange**
	MIN <= SPI <= -2.0	Extremely dry	0.000 - 0.023	2.3	**Red**

Source: European Commission, 2020, EDO INDICATOR FACTSHEET- Standardized Precipitation Index (SPI), page 4.
https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1101

4.4.1.1 *Computation of SPI*

The statistical distribution of precipitation amounts for defined intervals (e.g. monthly, 3 months, etc.) over a long time series (at least 30 years) can be effectively represented by a continuous probability distribution with two parameters, known as the "gamma distribution". To calculate the SPI for an observed rainfall accumulation for a time period of interest (e.g. 1, 3, 12 or 48 months), the two parameters (i.e. shape and scale) of the gamma distribution are first fitted to the non-zero frequency distribution of historical rainfall accumulations for all years in the available time series, using one of two alternative approximations of the "maximum likelihood estimators" for the gamma distribution developed by Thom (1958) and Greenwood and Durand (1960). For each observed rainfall accumulation, the cumulative probability is then derived based on the parameters of the gamma distribution and using the algorithms provided by Press et al. (1992). After adjusting for zero probability of rainfall accumulation, the cumulative probability of observed rainfall is then transformed into the standard normal random variable Z with mean zero and variance one, using an approximation described by Abramowitz and Stegun (1965). This transformed value is the SPI.

For METEO, the long-term gamma distribution is based on data from the reference period 1991-2020. The data included in the SPI-48 (4 years SPI) are thus based on monthly rainfall input data from 1988 until 2020. The input data are METEO’s operational precipitation grids.

As the SPI values are in units of standard deviation from the long-term mean, the indicator can be used to compare precipitation anomalies for any geographic location and any time-scale. The name of the indicator is usually modified to include the accumulation period. Thus, SPI-3 and SPI-12, for example, refer to accumulation periods of three and twelve months. Since the SPI can be calculated over different precipitation accumulation periods, the resulting different SPI indicators allow for estimating different potential impacts of a meteorological drought:

- **SPI-1 to SPI-3:** can be used as an indicator for immediate impacts such as reduced soil moisture, snowpack and flow in smaller creeks.
- **SPI-3 to SPI-12**: can be used as an indicator for reduced stream flow and reservoir storage.
- **SPI-12 to SPI-48:** can be used as an indicator for reduced reservoir and groundwater recharge.

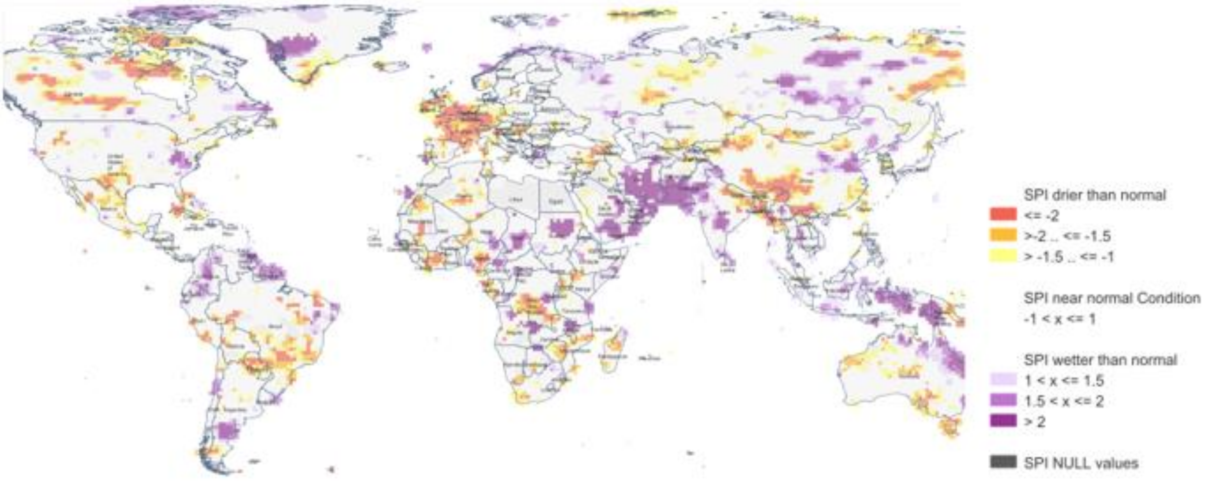
In order to get a full picture of the potential impacts of a drought, the SPI should be calculated and compared for different accumulation periods.

4.4.1.2 Results

In 2022, METEO delivered the global SPI based on the GPCC's (Global Precipitation Climatology Centre) "First Guess" Monthly Product. This product is calculated for the previous month and the month before and it uses monthly precipitation data based on SYNOP messages of approx. 7,000 stations arriving at DWD. After 3 months, the global grids are recalculated with corrected values and the SPI product is recalculated as the "Monitoring" Monthly product. One example of results is shown in Figure 10.

The computation of the SPI for the European domain using METEO's precipitation grids has also been tested during the year 2022. However, the analysis of the results highlighted some inaccuracies (possibly due to the spatial distribution of the stations). For this reason, the SPI for the European domain using METEO's precipitation grids has not been implemented operationally.

Figure 10. SPI Global 1 month, July 2022



Source: METEO

4.4.2 Heat and Cold Wave Index (HCWI)

The Heat and Cold Wave Index (HCWI) is used for detecting and characterising periods of extreme-temperature anomalies (i.e. heat and cold waves), as they can have strong impacts on human activities and health.

The HCWI shows the duration in days of detected heat/cold waves and is computed for each location (grid-cell), using the methodology developed by Lavaysse et al. in 2018. For this, the threshold values of daily minimum and maximum temperature (TN and TX) that characterise a heat or cold wave are computed from TN and TX for that calendar day during a 30-year baseline period (1991-2020, same as SPI).

4.4.2.1 Computation of HCWI

For heat waves, the daily threshold values for TN and TX are defined as the 90th percentile of all temperature values in an 11-day window centred on that day, for all 30 years in the baseline period (330 temperature values). For cold waves, the daily threshold values for TN and TX are defined as the 10th percentile of all temperature values in an 11-day window centred on that day, for all years in the baseline period (330 temperature values). A heat or cold wave is detected when there are at least three consecutive days with both TN and TX above (for heatwaves) or below (for cold waves) their daily threshold values.

METEO’s HCWI indicates directly where and when a heat or cold wave has occurred. By definition, the minimum duration of a wave is three days, increasing to fourteen days and longer.

Other heat and cold wave-related data layers available via METEO are:

- Normalised temperature anomaly which is the difference of the observed daily maximum temperature minus the daily long-time average maximum temperature of the 11-day-window centred at the observed day of the year (period 1991-2020) divided by the standard deviation of the long-time average daily maximum temperature of the same 11-day-window. A value of “1” means, the temperature anomaly is 1 standard deviation above the normal value.
- Extreme temperature gives the highest daily maximum temperature observed since beginning of (i.e. during) the heat wave. In case of a cold wave, this is the lowest daily minimum temperature observed since beginning of (i.e. during) the cold wave.

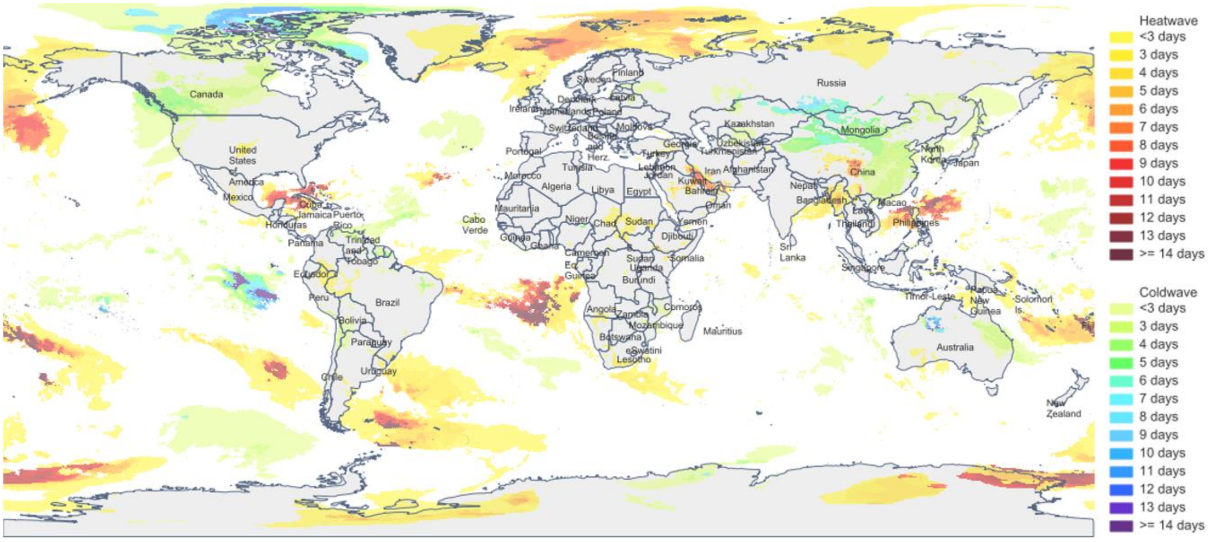
Finally, METEO’s data files include two other parameters: start and end date of heat or cold wave. These values are available for every grid point (and each day).

4.4.2.2 Results

In 2022, METEO delivered the HCWI for the global domain using ERA5[4] reanalysis data as input. One example of results can be seen in Figure 11: the HCWI indicator displays the duration of a heat or cold wave in order of increasing magnitude, with colour scales ranging from yellow to dark red (for heat waves) and blue to dark purple (for cold waves). Figures 12 and 13 show an example of two other global products: the global normalised temperature anomaly and the global extreme temperature.

Finally, it is here noted that the computation of the HCWI for the European domain using METEO’s minimum and maximum temperature grids has been tested during the year 2022. However, the analysis of the results highlighted some inaccuracies presumably due to the spatial distribution of the stations. This issue requires further analysis. Therefore, the computation HCWI for the European domain using METEO’s minimum and maximum temperature grids has not yet been implemented into operations.

Figure 11. HCWI Global, 02/12/2022

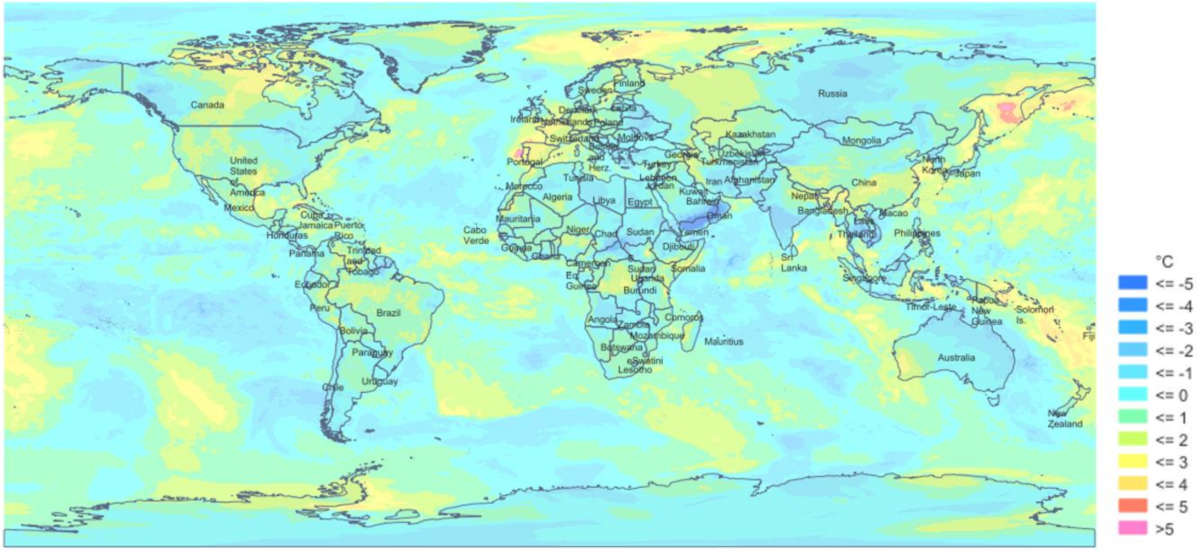


Source: METEO

---

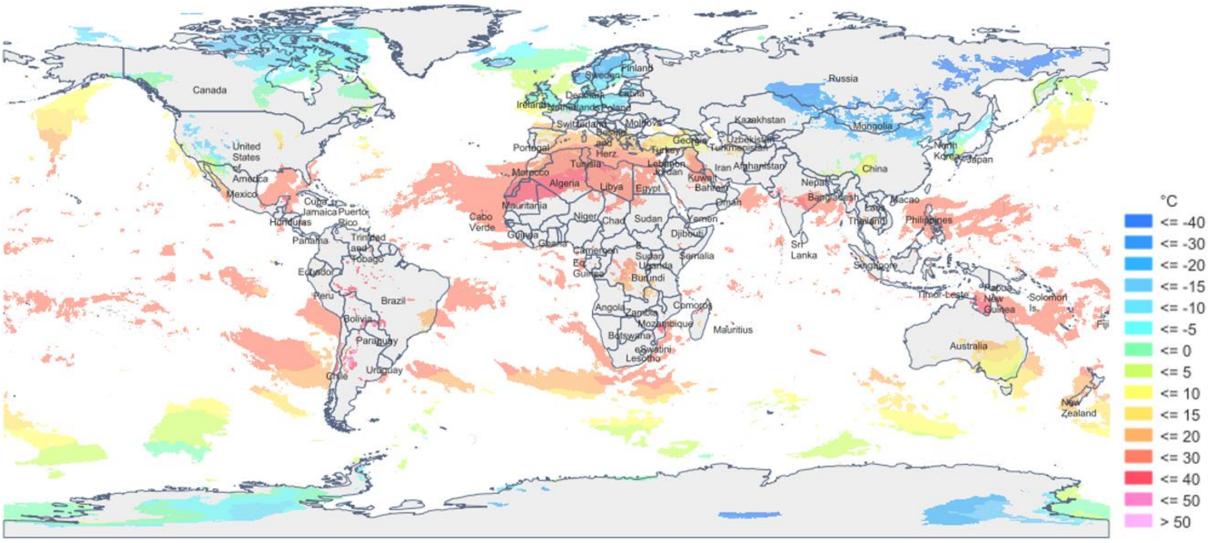
[4] https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5

Figure 12. Normalised temperature anomaly Global, 13/07/2022



Source: METEO

Figure 13. Extreme temperature Global, 15/12/2022



Source: METEO

5 Gap Analysis

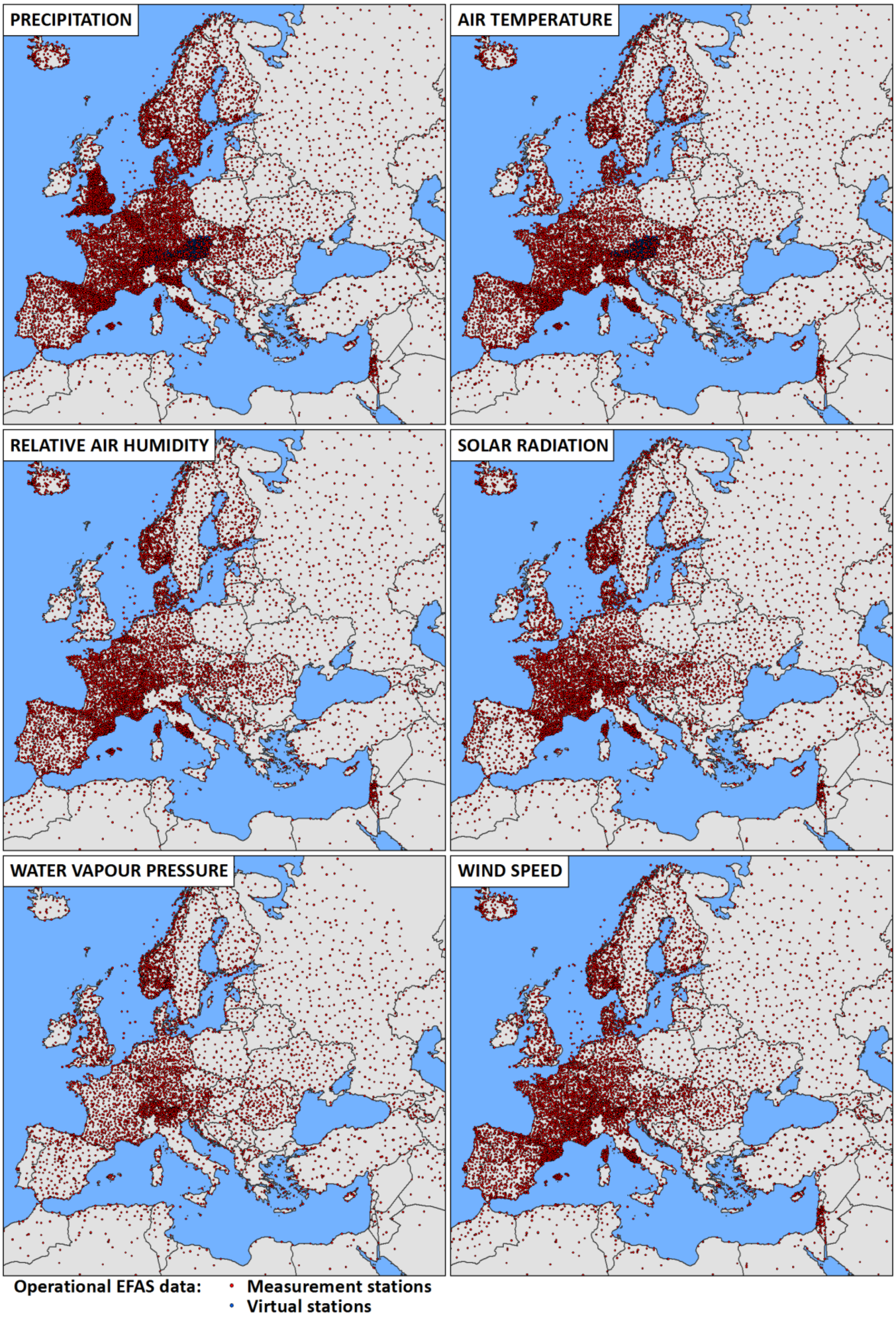
5.1 Gap Analysis

Even if the interpolation scheme estimates data in un-probed regions, the reliability of grids is higher and uncertainty is lower if the data density is high and the input stations are homogeneously distributed in space. Station density varies largely within the EFAS domain and the data density depends on the parameter as not all stations measure all parameters nor all data providers deliver all parameters at all stations. Meteorological services operate many sensors at a station but the station density is sometimes low, even if it is sufficient for the legal task of the specific service. On the other hand, hydrological services operate often a station network with a higher station density but they focus mainly on precipitation with only a few stations observing temperature, wind or solar radiation. As currently not all existing meteorological and/or hydrological services are contributing with their data to the CEMS METEO database, the spatial distribution of available stations and parameters is very inhomogeneous. This section aims to detect gaps in the real-time data availability to give some advice for future data collection activities.

The base information of the available data is fetched from the World Meteorological Organization (WMO) Global Telecommunication System (GTS). These data (referred to as “Deutscher Wetterdienst (Global)” in Table 1), exchanged from the national hydrological/meteorological services, cover the whole EFAS-Domain but are not spatially homogeneous as some countries exchange more data than others. This data exchange, on the other hand, covers nearly all required parameters. Such data can be densified by additional deliveries from national services based on agreements between the national data providers and the Copernicus program. As add-on, the deliveries from national meteorological services offer a redundant data delivery, while hydrological services provide a high density of mainly precipitation stations for certain administrative regions or catchment areas.

Figure 3 depicts the spatial distribution of the stations that delivered at least one parameter in real time in 2022. In this Figure the station density in Algeria, Azerbaijan, Belarus, Bulgaria, Egypt, Greece, Iraq, Iran, Ireland, Jordan, Libya, Lithuania, Morocco, Poland, Russia, Saudi Arabia, Tunisia, Türkiye and in some parts of Italy is low when compared to the other countries of the EFAS domain. Figure 14 complements Figure 3 and it shows the spatial distribution of the stations that delivered one specific parameter in 2022. When focusing on precipitation, a comparatively low station density occurs additionally in Ireland. Solar radiation is a parameter not observed and distributed at many stations. Countries with the highest station density delivering this parameter are Belgium, Denmark, Northern and West-central Italy, Israel, France, Germany, the Netherlands, Portugal, North-eastern Spain, Hungary, Romania, Switzerland, Tunisia, and the United Kingdom. In Spain, the river authority for the Ebro River and the Servei Meteorològic de Catalunya provide solar radiation data with a higher coverage than the other parts of this country. A medium solar radiation station density is in Armenia, Estonia, Iran, Latvia, Lithuania, Russia and Türkiye.

Figure 14. Spatial distribution of active stations providing data for the parameter noted above the map. Upper row: precipitation (left) and air temperature (right). Middle row: relative air humidity (left) and solar radiation (right). Lower row: water vapour pressure (left) and wind speed (right).



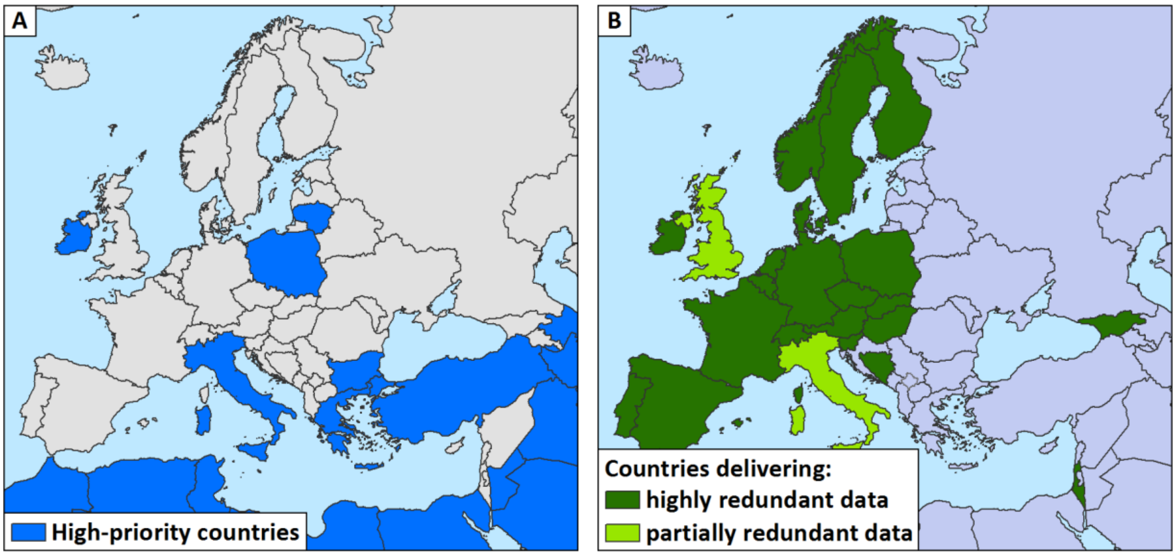
Source: METEO

5.2 Proposal for future data collection strategy

METEO aims to increase the spatial density of the stations network. The gap analysis of the previous section was complemented by pragmatic considerations (e.g. expected easiness of establishing a contact with the local data providers) to define the list of countries for which stations should be included with a high priority . These countries are in alphabetic order Algeria, Azerbaijan, Bulgaria, Egypt, Greece, Iraq, Iran, Ireland, Italy, Jordan, Libya, Lithuania, Morocco, Poland, Saudi Arabia, Tunisia and Türkiye (Figure 15 left)

Multiple data deliveries are currently implemented as a safety feature (in case of supplier failure) for countries marked in dark green in Figure 15 (right). In alphabetic order these are Austria, Bosnia and Herzegovina, Czech Republic, Finland, Georgia, Germany, Hungary, Ireland, Luxembourg, Kosovo, the Netherlands, Norway, Poland, Portugal, Spain, Slovakia, Slovenia, Sweden and Switzerland, and due to the integration of regional or hydrological data providers regional in Italy, United Kingdom of Great Britain and Northern Ireland. These countries are marked in light green in Figure 15 (right).

Figure 15. Map of the EFAS-Domain with (A) countries proposed to give a high priority to establish a data provision to METEO and (B) countries with redundant data deliveries (dark green) and redundant data deliveries for parts of the country (light green) to METEO



Source: METEO

Regarding the measured parameters, for the integration of additional stations, the highest priority should be given to increase the amount of available precipitation observations and secondly to a more homogeneous delivery of solar radiation data. This priority list has been defined according to the following reasoning: precipitation is of high importance in the hydrological modelling because it is a critical input variable with high spatial and temporal variability; solar radiation is currently provided by a limited and inhomogeneous distributed number of stations (section 5.1).

Additionally, the calculation of not provided parameters from available data should be implemented, e.g. water vapour pressure can be calculated from dew point temperature and air temperature values.

6 New Developments

This section deals with the innovations already set up in 2022, the data validation framework and new seasonal and geographical thresholds for certain variables, which will officially go into operation in 2023. Furthermore, a brief outlook is given on the innovations within 2023: Update of METEO grids, Computation of global SPI using ERA5 data and the usage of a new open source interpolation scheme.

6.1 Data Validation Framework

A new data validation framework was implemented which can be used manually at any time or set up to run several tasks automatically in given schedules. The framework allows application of a variety of data validation algorithms. The following principles apply:

1. Consistency-type rules: check validated time series for conditions such as whether the timeseries are up-to-date, incomplete or if there are any gaps in the data.
2. Validation-type rules: check whether individual data points in a time series conform to the algorithmic parameters detailed in the rule configuration. Data points that fail to conform are flagged as "incidents" and are tagged with a comment and, if requested, are also marked with a designated quality code. The validation rules can be split further into different groups: single target rules, comparison rules and spatial rules.
    (a) Single target validation type rules are rules which only use data points from a single time series (the target time series) to validate the data, e.g. threshold, rate-of-change, range, etc.

    (b) Comparison-type rules make use of other time series to validate data of a target time series. These can be generally split up into two groups: Inner Consistency and 'Outer' Comparison-type rules. Inner Consistency type rules only use data from time series located at the same station as the target. A common example of this might be a check to ensure that if a precipitation sensor records rainfall, then the humidity sensor at the same station should measure at least 50% relative humidity. An 'Outer' Comparison type rule will use time series from outside the target station in the validation. The key requirement for outer comparison type rules is that the relationships between the stations or time series must always be explicitly defined. This differentiates them from Spatial Type rules which select their neighbours based on spatial proximity alone.

    (c) Spatial type rules make use of the spatial component of a measurement network to validate data. For this type of rules some additional considerations and parameters need to be thought of. In general, neighbour identification operates in three steps. The first step is to identify all possible neighbouring stations within a given search radius. In the second step, this list of neighbour candidates is reduced based on any disqualifying characteristics, e.g. stations where the elevation difference to the target station exceeds the maximum elevation difference parameter. Likewise, inactive stations with no data within the validation period are removed. In the third step, the remaining list of neighbours is sorted to prioritise and to be trimmed until the number of selected neighbours equals the group size parameter. Prioritization here is done by either orientation, quadrant or distance mode where orientation mode tries to select stations from each of the four cardinal directions and quadrant mode tries the same for each of the four quadrants around the target station. Distance mode ignores these restrictions and selects stations only on their proximity to the target station. This procedure is run for every station in the target data set.
3. Correction-type rules: correct suspect values identified by one of the preceding rules or fill gaps. Different algorithms for correction or gap filling are available, e.g. use value from nearest neighbour or interpolate from values at the edges of a gap

It is intended to make first use of the data validation framework by applying the following rules in order to improve validation of precipitation data:

1. Spatial Comparison: checks data at a target station against a range of values from neighbouring stations. The range of values from these neighbouring stations form an upper and lower boundary. Additional

delta values can be defined in the rule configuration to widen the acceptable thresholds. Any value from the target station outside these thresholds will result in a validation exception.

2. Spatial Zero Comparison: compares the values at a target station with data from neighbouring stations. If a time stamp at the target station records no rainfall but neighbouring stations do, then a validation exception occurs.

It is here noted that *validation exception* means that the value is not trustworthy, and further checks are required. A protocol for the optimal use of the Spatial Comparison, the Spatial Zero Comparison, and the other quality checks of the validation protocol is under investigation. This protocol will foresee the use of a combination of rules. The METEO observations database, with data from 1970 to 2021, will allow implementing a thorough testing of the protocol. The results of the testing phase will be presented in the Annual Report for the reference year 2023.

Furthermore, additional rules will be tested in order to detect issues in the data that remained undetected with the validation methods described in chapter 3.2.1. An example for such an additional rule is the "Value Sequence"-rule which checks a time series to determine the similarity between values in a sequence of data. If the values are too similar to one-another over a certain time frame, a validation exception is issued. This rule can be used to detect so-called "flatlining", where a sensor is continuously transmitting values with no or very little variation.

6.2 Seasonal and Geographical Thresholds

New seasonal and geographical thresholds were set for parameters:

- Air temperature
- Dew point temperature
- Sunshine duration
- Solar radiation
- Vapour pressure

for all climate zones within the EFAS-Domain. They were calculated from the upper and lower percentile of all available data between 1990 and 2021 within each climate zone, as this covers the more recent WMO reference period. No changes were made for relative air humidity, cloud cover and wind direction (percentiles were calculated and analysed, but no dependencies were found). The upper threshold for wind speed was increased to 50 m/s (former 45 m/s) over all climate zones equally, as this better fits to the observed data. The precipitation thresholds will not change compared to the ones given in Table 5 as precipitation itself is very variable and there is no regularity to applying any percentile thresholds. However, the previously mentioned additional quality checks (e.g. the spatial and spatial zero comparisons) will be applied to further improve the quality of the precipitation dataset.

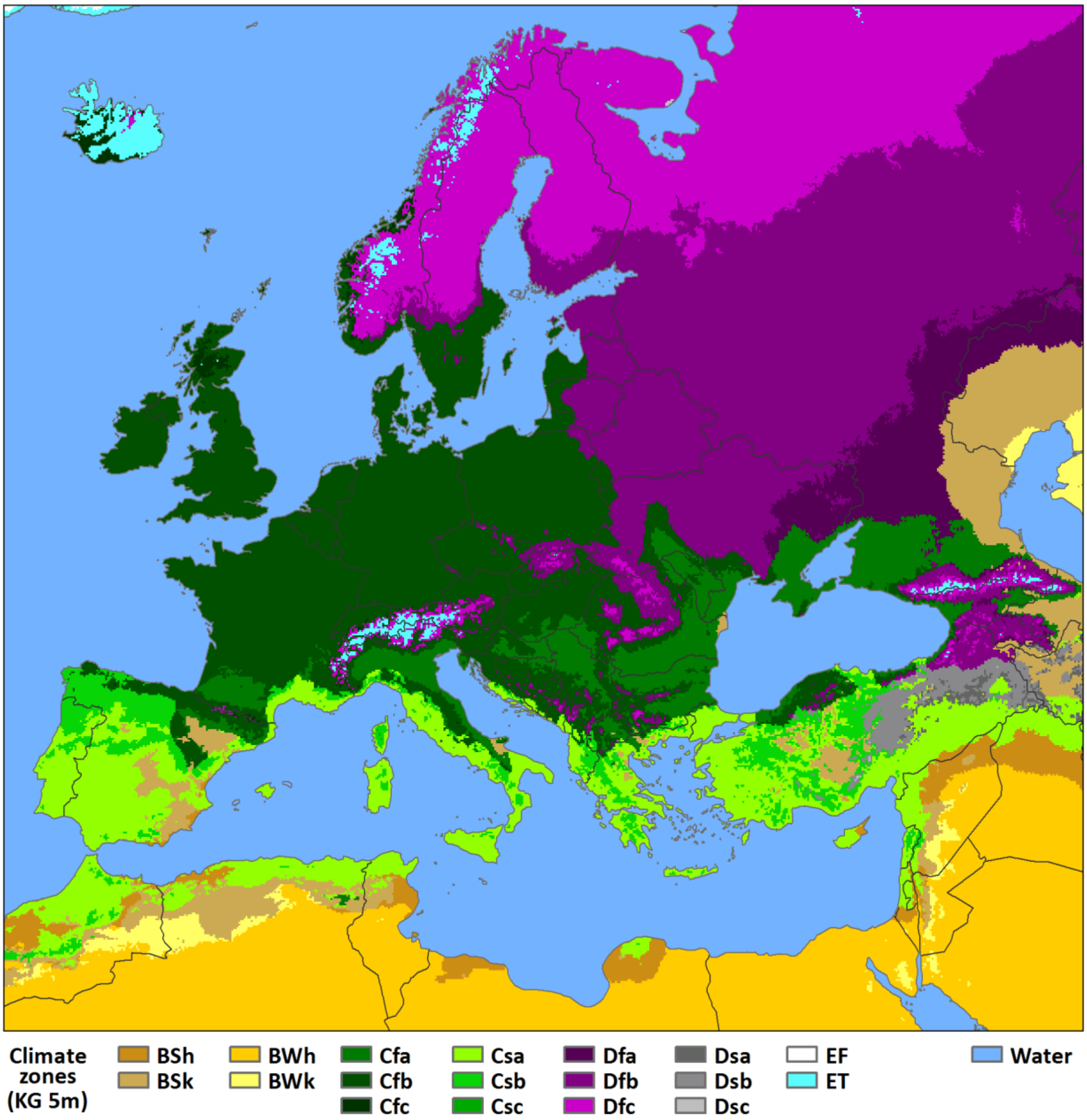
As you can see in Figure 16, the climate over the EFAS-Domain is quite diverse. The aim was to better adapt the thresholds for the five variables listed above to their location and thus make the data more representative of their climate zone.

Especially in winter, many temperature values in the south and southeast of the EFAS-Domain (e.g. Egypt, Saudi Arabia, Syria, etc.; climate zone BWh) have been erroneously rejected so far as the values were above the generally set limits. With the limits now specially adapted for more arid climates, the data can be evaluated and verified more realistically. Moreover, due to the new thresholds, values that are too high for their region and would otherwise have slipped through the quality check can now be correctly excluded for further processing (e.g. Scandinavia; climate zone Dfc).

To define an appropriate limit, several percentiles were calculated using the data in the METEO data base. For the lower limit, the 0.01%, 0.1%, 1% and 5% percentiles were calculated. The 95%, 99%, 99.9% and 99.99% percentiles were computed to define the upper limits. The distribution curve represents a daily variation and was smoothed by a seven days low-pass filter (for example Figure 17). In an expert meeting, the percentiles for each parameter and climate zone were visually checked and the one used as threshold and the safety margin were decided (Table 7).

While the thresholds are based on the past observations, the risk of eliminating real observations connected to extreme (not yet observed) meteorological events was largely mitigated by (i) identifying the optimal (according to expert judgement) percentile value for each variable and each climate zone; (ii) adding a buffer to enlarge the range of acceptable values; (iii) planning a periodic revision of the seasonal and geographical thresholds; (iv) planning a thorough testing and monitoring of the impact of the new threshold during the operational deployment in 2023. It is highlighted that the magnitude of the buffer was identified individually for each variable and each climate zone, according to a detailed analysis of all the available observations (the values are reported in Table 7).

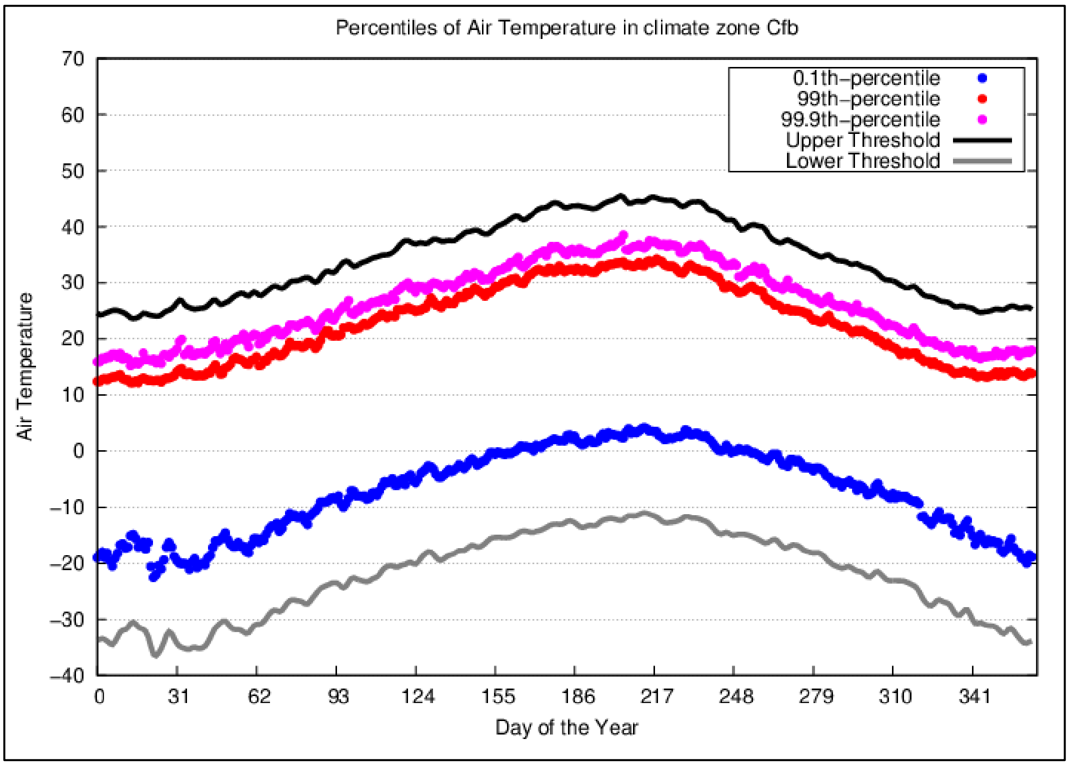
Figure 16. Climate zones within EFAS-Domain



Source: METEO with data from Kottek, M., et al., 2006.

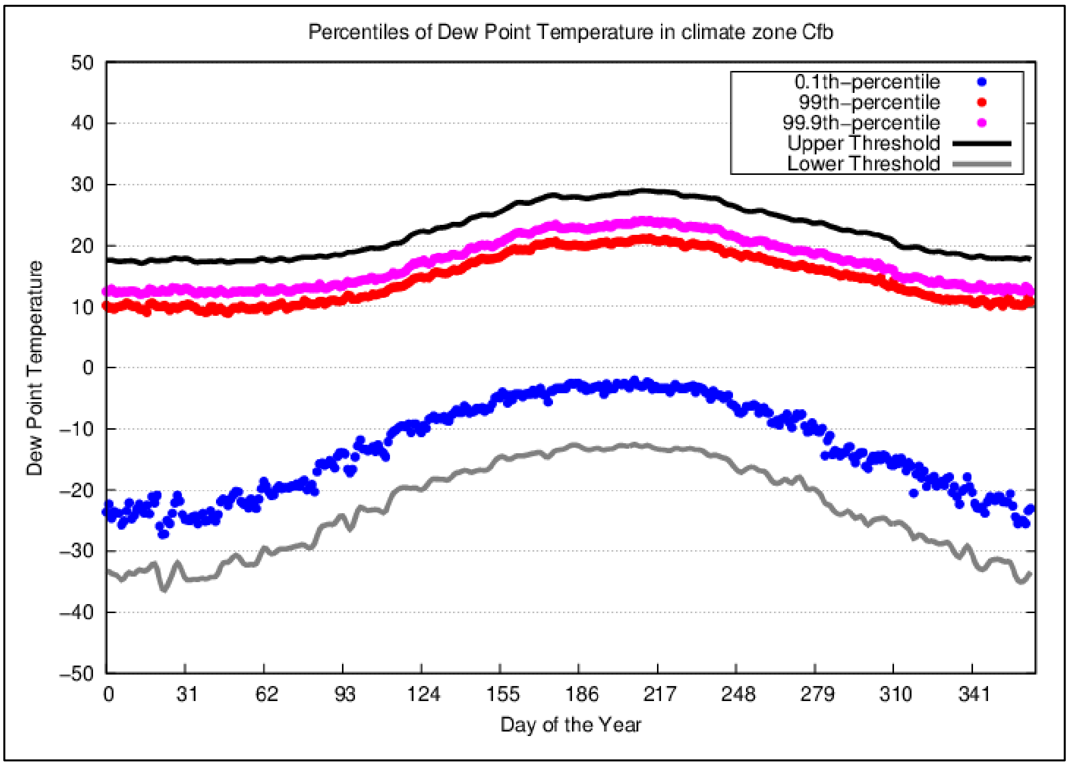
The following graphics are examples of the new thresholds for climate zone Cfb:

Figure 17. Percentiles of Air Temperature in climate zone Cfb



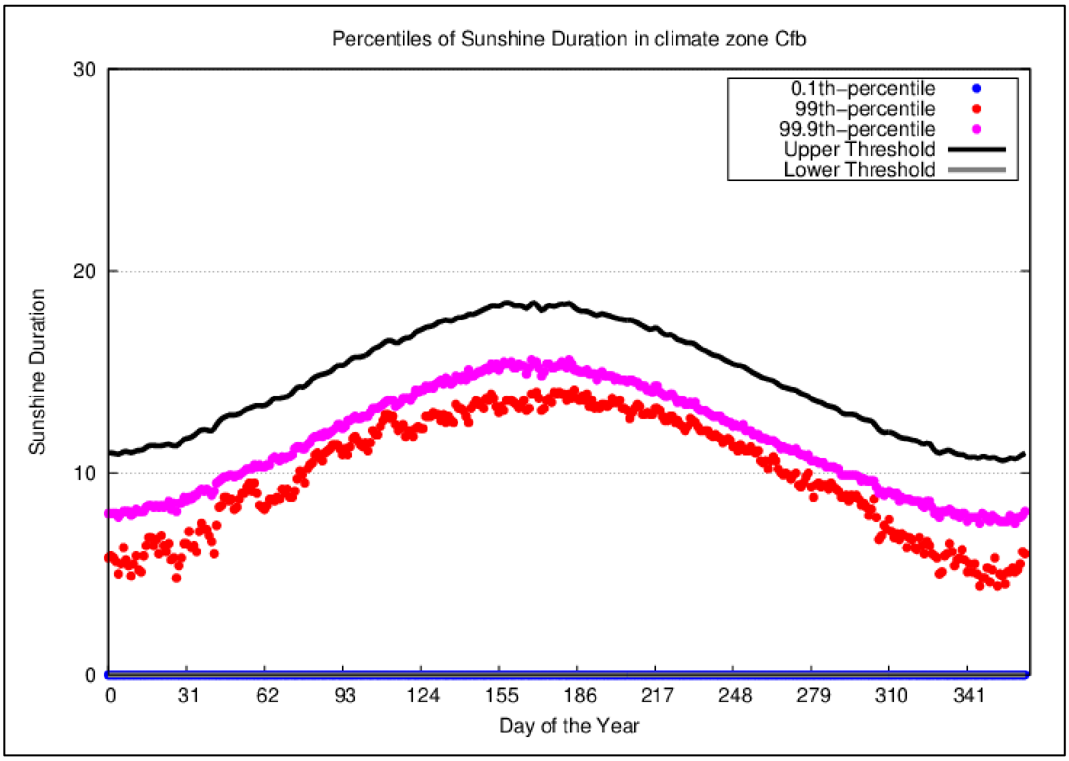
Source: METEO

Figure 18. Percentiles of Dew Point Temperature in climate zone Cfb



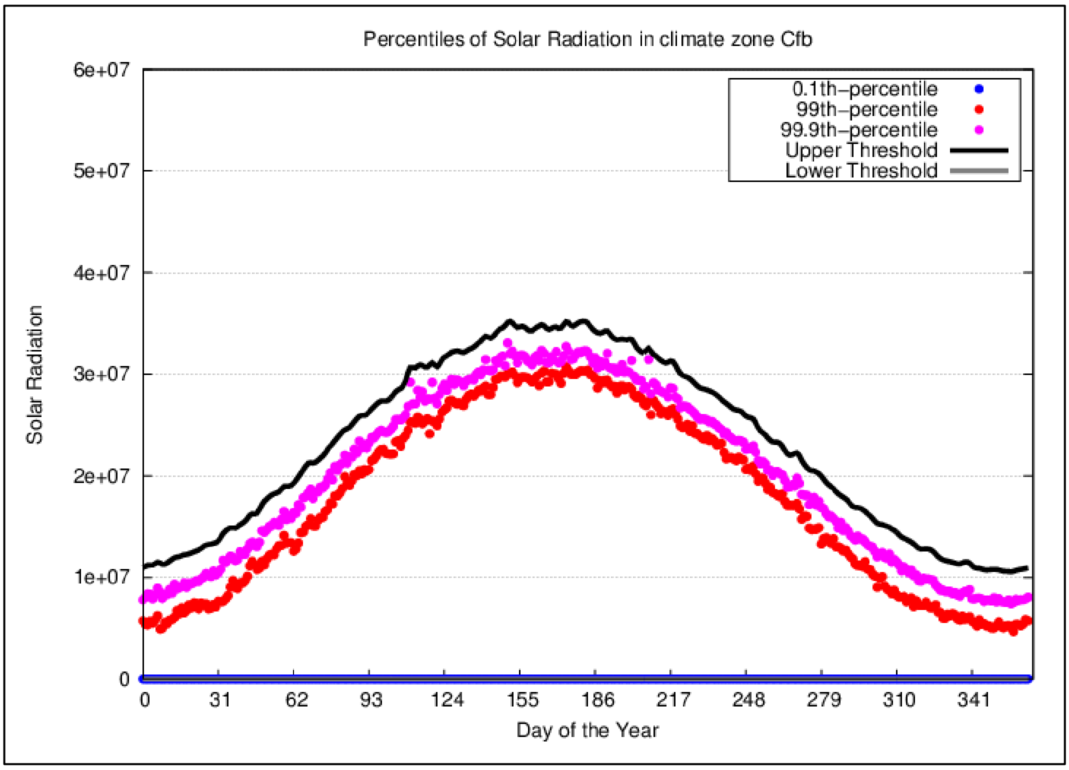
Source: METEO

Figure 19. Percentiles of Sunshine Duration in climate zone Cfb



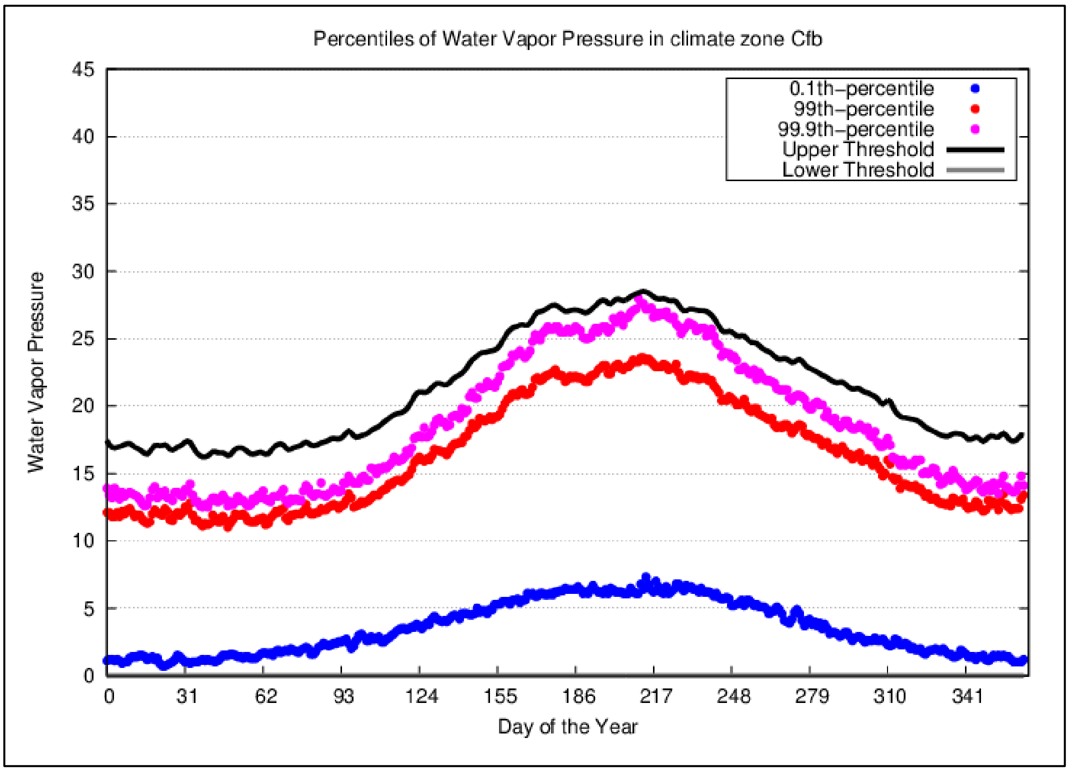
Source: METEO

Figure 20. Percentiles of Solar Radiation in climate zone Cfb



Source: METEO

Figure 21. Percentiles of Water Vapour Pressure in climate zone Cfb



Source: METEO

The new thresholds for each parameter are as follows:

Table 7. New geographical and seasonal thresholds

		Parameter				
		AT	**DT**	**SunD**	**SunRad**	**VP**
Climate zone	BWh	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=95% + 1e7	ll=0 ul=99% + 8 hPa
	BWk	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	NA	NA	ll=0 ul=99% + 8 hPa
	BSh	ll = 0.1% -10K ul = 99.9% + 5K	ll = 1% -15K ul = 99% +10K	ll=0 ul=99.9% +3h	ll=0 ul=95% + 1e7	ll=0 ul=99% + 8 hPa
	BSk	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	Csa	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99% +3h	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	Csb	ll = 0.1% -10K ul = 99.9% + 5K	ll = 1% -15K ul = 99.9% + 5K	ll=0 ul=99% +3h	ll=0 ul=95% + 1e7	ll=0 ul=99% + 8 hPa

		Parameter				
		AT	**DT**	**SunD**	**SunRad**	**VP**
	Cfa	ll = 0.1% -10K ul = 99.9% + 5K	ll = 1% -15K ul = 99% +10K	ll=0 ul=99% +3h	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	Cfb	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=99% + 3e6	ll=0 ul=99% + 8 hPa
	Cfc	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	NA	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	Dsa	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=95% + 1e7	ll=0 ul=99% + 8 hPa
	Dsb	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=95% + 1e7	ll=0 ul=99% + 8 hPa
	Dsc	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	NA	NA	NA
	Dfa	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	NA	ll=0 ul=99% + 8 hPa
	Dfb	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	Dfc	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.9% +3h	ll=0 ul=99% + 5e6	ll=0 ul=99% + 8 hPa
	ET	ll = 0.1% -10K ul = 99.9% + 5K	ll = 0.1% -10K ul = 99.9% + 5K	ll=0 ul=99.99% +3h	ll=0 ul=99% + 3e6	ll=0 ul=99% + 8 hPa

ll= lower limit
ul= upper limit
NA= not available due to low data availability; in these cases, the old thresholds apply from section 3.2.1.

Source: METEO

6.3 Foreseen developments in 2023

- Implementation of the new quality checks (see sections 6.1 and 6.2) to verify the entire historical database.

Update of METEO grids: computation of higher resolution operational grids with 1-arc-minutes (~1.4 km) spatial resolution to allow the implementation of higher resolution CEMS products. The EFAS 1 arcmin domain will replace the existing EFAS Extended domain. The EFAS 1-arcmin domain covers most parts of the EFAS extended domain; it extends even further south into North Africa and it also covers more of Saudi- rabia, Iraq and Iran. The EFAS 1-arcmin domain uses the well-know WGS84 longitude/latitude coordinates (EPSG code 4326). Extent: longitude 25.25°W - 50.25°E, latitude 22.75°N - 72.25°N. In order to avoid problems with gridding along the borders of the domain, METEO will use station data from a slightly larger region of -29.25°W - 54.25°E, 20.75°N - 74.25°N.

- Computation of global SPI using ERA5: New edition of the global SPI by replacing GPCC input data by ERA5 reanalyzing grids with a higher update cycle of 10 days.
- New Interpolation scheme: In order to be consistent with the open source spirit of EFAS, not only the hydrological forecast model OS LISFLOOD but also an open source interpolation scheme (developed by the Joint Research Centre) will be used.

7 Conclusions

This report provides an overview about the status and progress of the Copernicus Emergency Management Service (CEMS) Meteorological Data Collection Centre (METEO) in 2022.

CEMS METEO collects, quality-controls and post-processes in situ and ground-based remote sensed meteorological data to provide input data tailored to the needs of the CEMS EFAS, EDO/GDO, EFFIS. All real-time and historical data are quality-controlled in the same manner during the import to the database. Only data that passed the quality control are processed further and provided as station lists or grids including an uncertainty estimation.

Within 2022, four additional real-time data providers were integrated as well as several historical data. It is here noted that one gridded dataset was integrated as a temporary solution to diminish the impact of the lack of data delivery from Ukraine (due to the war).

The status of the METEO data collection in Europe at the end of 2022 is summarized as follows: real-time data are delivered by 35 providers for 11 parameters from around 22,000 stations. On average, 9,000,000 records are being added to the database each day. Additionally, the database includes more than 31,500 stations with historical data.

During 2022 METEO undertook several activities to further improve data collection and validation, and the delivery of products to CEMS. First, a gap analysis regarding the available data allowed to define a priority list in terms of countries and parameters. Second, additional quality checks have been developed and planned. The deployment of these quality checks into operation is planned for the year 2023. It is expected that these additional quality checks along with the inclusion of new stations will further improve the quality and accuracy of the collection and monitoring of the meteorological variables in the European domain.

References

Arnal, L., Asp, Baugh, de Roo, Disperati, Dottori, Garcia, GarciaPadilla, Gelati, Gomes, Kalas, Krzeminski, Latini, Lorini, Mazzetti, Mikulickova, Muraro, Prudhomme, Rauthe-Schöch, Rehfeldt, Salamon, Schweim, Skoien, Smith, Sprokkereef, Thiemig, Wetterhall, Ziese, EFAS upgrade for the extended model domain – technical documentation, EUR 29323 EN, *Publications Office of the European Union*, Luxembourg, 2019, ISBN 978-92- 79-92881-9, doi: 10.2760/806324, JRC111610.

Abramowitz, M. and I. A. Stegun (eds), 'Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables', *NBS Applied Mathematics Series,* Vol. 55, National Bureau of Standards, Washington, DC, 1964.

Balsamo, G., Albergel, C., Beljaars, A., Boussetta, S., Brun, E., Cloke, H., Dee, D., Dutra, E., Muñoz-Sabater, J., Pappenberger, F., de Rosnay, P., Stockdale, T., and Vitart, F., 'ERA-Interim/Land: a global land surface reanalysis data set', *Hydrol. Earth Syst. Sci*., Vol. 19, 2015, pp. 389–407.

Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-.-J., Park, B.-.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-.-N. and Vitart, F. (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc*., 137: 553-597. https://doi.org/10.1002/qj.828

Edwards, D.C. and T.B. McKee, 'Characteristics of 20th Century Drought in the United States at Multiple Time Scales', *Climatology Report*, Vol. 97, No 2, Colorado State University, Fort Collins, 1997.

Greenwood, J.A. and D. Durand, 'Aids for fitting the gamma distribution by maximum likelihood', *Technometrics*, Vol. 2, No 1, 1960, pp. 55-65.

'Heat and Cold Wave Index (HCWI)', *EDO INDICATOR FACTSHEET*, 2018, Copernicus European Drought Observatory, European Comission, 2018.

Isotta, F., C. Frei, V. Weilguni, M. Lassegues, P. Tadic, B. Rudolf, V. Pavan, C. Cacciamani, et al., 'The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data.', *International Journal of Climatology*, Vol. 34, 2014, pp. 1657-1675.

Kottek, M., J. Grieser, C. Beck, B. Rudolf, and F. Rubel, 'World Map of the Köppen-Geiger climate classification updated', *Meteorol. Z.,* Vol. 15, No 3, 2006, pp. 259-263.

Lavaysse, C., C. Cammalleri, A. Dosio, G. van der Schrier, A. Toreti, and J. Vogt, 'Towards a monitoring system of temperature extremes in Europe', *Natural Hazards and Earth System Sciences*, Vol. 18, 2018, pp. 91-104.

'Manual on Codes Volume I.1 Annex II to the WMO Technical Regulations Part A – Alphanumeric Codes', WMO No.306, Part I.1 Part A, World Meteorological Organization, Geneva, Switzerland, 2019.

McKee, T.B., N.J. Doesken and J. Kleist, 'The relationship of drought frequency and duration to time scale', *Proceedings of the Eighth Conference on Applied Climatology*, Anaheim, California, 17–22 January 1993. Boston, American Meteorological Society, pp. 179–184.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery., *Numerical Recipes in C: The Art of Scientific Computing*. 2nd Edition. Cambridge University Press, 1992, 994p.

'Standardized Precipitation Index (SPI)', *EDO INDICATOR FACTSHEET*, 2020, Copernicus European Drought Observatory, European Comission, 2020.

Sideris, I. V., M. Gabella, R. Erdin, and U. and Germann, 'Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland.' *Q. J. Roy. Meteor. Soc*., Vol. 140, 2014, pp. 1097–1111.

Spinoni, J., S. Szalai, T. Szentimrey, M. Lakatos, Z. Bihari, A. Nagy, Á. Németh, et al., 'Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables."', *International Journal of Climatology*, Vol. 35, 2015, pp. 1322–1341.

Spinoni, J., Szalai, S., Szentimrey, T., Lakatos, M., Bihari, Z., Nagy, A., Németh, Á., Kovács, T., Mihic, D., Dacic, M., Petrovic, P., Kržic, A., Hiebl, J., Auer, I., Milkovic, J., Štepánek, P., Zahrad- ˇ nícek, P., Kilar, P., Limanowka, D., Pyrc, R., Cheval, S., Birsan, M.-V., Dumitrescu, A., Deak, G., Matei, M., Antolovic, I., Nejedlík, P., Štastný, P., Kajaba, P., Bochnícek, O., Galo, D., Mikulová, K., Nabyvanets, Y., Skrynyk, O., Krakovska, S., Gnatiuk, N., Tolasz, R., Antofie, T., and Vogt, J.: Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables, *Int. J. Climatol*., 35, 1322–1341, https://doi.org/10.1002/joc.4059, 2015

Thiemig, V., G. N. Gomes, J. O. Skøien, M. Ziese, A. Rauthe-Schöch, E. Rustemeier, K. Rehfeldt, et al.,'EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe.', *Earth Syst. Sci. Data*, Vol. 14, 2022, pp. 3249–3272.

Thom, H.C.S. 'A Note on the Gamma Distribution', *Monthly Weather Review*, Vol. 86, No 4, Office of Climatology, U.S. Weather Bureau, Washington, D. C.,1958.

Willmott, C., C. Rowe, and W. Philpot, 'Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring.' *The American Carthographer,* Vol. 12, 1985, pp. 5-16.

Yamamoto, J., 'An Alternative Measure of the Reliability of Ordinary Kriging Estimates.', *Mathematical Geology,* Vol. 32, 2000, pp. 489-509.

Ziese, M., Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schneider, U., *GPCC First Guess Product at 1.0°: Near Real-Time First Guess monthly Land-Surface Precipitation from Rain-Gauges based on SYNOP Data*, GPCC c/o Deutscher Wetterdienst, Offenbach a. Main, 2011.

List of abbreviations and definitions

%	Percent
°C	Degree Celsius
API	Application Programming Interface
AT	Air temperature
BWh	Hot desert climate zone
CarpatClim	Historic dataset (1961-2010) for climate data in the Carpathian region.
CEMS	Copernicus Emergency Management Service
Cfb	Temperate oceanic climate zone
ClCo	Cloud cover
Danube	Historic dataset (1990-2003) provided by JRC
Dfc	Subarctic climate zone
DMI	Danish Meteorological Institute, Denmark
DT	Dew point temperature
DWD	Deutscher Wetterdienst (German weather service)
DWD Climatic	Historic dataset (1990-2005) provided by JRC
ECA	Historic dataset (varying time periods) provided by JRC
EDO	European Drought Observatory
EFAS	European Flood Awareness System
EFFIS	European Forest Fire Information System
EMO-5	High-resolution multi-variable gridded meteorological dataset for Europe (1990-2019)
ERA5	ECMWF atmospheric reanalysis of the global climate (1940-present)
ERA-Interim-land	Global land surface reanalysis data set (1979–2010)
Euro Synop	Historic dataset (2003-2009) provided by JRC
EURO4M-APGD	Alpine precipitation grid dataset (1971–2008)
Evap	Evaporation
ftp	File Transfer Protocol
GDO	Global Drought Observatory
GPCC	Global Precipitation Climatology Centre
GTS	Global Telecommunication System
HCWI	Heat and Cold Wave Index
HIC	Hydrological Information Centre, Belgium
ICON	Icosahedral Nonhydrostatic numerical weather prediction model
JRC	Joint Research Centre
ll	Lower limit
MARS	Monitoring Agricultural Resource System
METEO	Meteorological Data Collection Centre
MeteoConsult	Historic dataset (2007-2015) provided by JRC

min	Minutes
mm	Millimetres
NA	Not available
Precip	Precipitation
ReAiHu	Relative Air humidity
SAIH	Automatic System of Hydrological Information, Spain
SPI	Standardized Precipitation Index
SPW	Service public de Wallonie, Belgium
SunD	Sunshine duration
SunRad	Solar radiation
SYNOP	Synoptic observations
TN	Daily minimum temperature
TX	Daily maximum temperature
ul	Upper limit
UTC	Universal Time Coordinated
VMM	Flanders Environment Agency, Belgium
VP	Water vapour pressure
WDir	Wind direction
WISKI	Water Information Systems by KISTERS
WMO	World Meteorological Organization
WSpeed	Wind speed
Z	Standard normal random variable

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

— at the following standard number: +32 22999696,

— via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

Science for policy

The Joint Research Centre (JRC) provides independent, evidence-based knowledge and science, supporting EU policies to positively impact society

EU Science Hub
joint-research-centre.ec.europa.eu

@EU_ScienceHub
EU Science Hub - Joint Research Centre
EU Science, Research and Innovation
EU Science Hub
@eu_science