





Copernicus Emergency Management Service

The CEMS Meteorological Data Collection Centre

Annual Report 2021

Para (

Prepared by the CEMS METEOROLOGICAL DATA COLLECTION CENTRE







This publication is a Technical report by 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

JRC129076

Ispra: European Commission, 2022

© European Union, 2022



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 (<u>https://creativecommons.org/licenses/by/4.0/</u>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

How to cite this report: Rehfeldt, K., Schirmeister, Z., Pichon, D., Rauthe-Schöch, A., Schweim, C., Walawender, J., Ziese M., Gomes, G., Thiemig, V. and Salamon, P., The CEMS Meteorological Data Collection Centre – Annual report 2021, European Commission, 2022, JRC129076

Contents

	Acknowledgements	1
1	Description of CEMS METEO	3
2	Data Providers and Provision	4
	2.1 Data Providers, Parameter and Time Resolution	4
	2.2 New Data Providers in 2021	8
3	Database	10
	3.1 Data Flow	10
	3.2 Quality Control	10
	3.2.1 Quality control statistics	12
	3.3 Database Statistics	12
4	Post-Processing	15
	4.1 Aggregation and Disaggregation, Calculation of Minimum, Maximum and Mean	15
	4.2 Gridding: Meteorological grids for CEMS EFAS	16
	4.2.1 Operational grids	16
	4.2.2 Historical grids: EMO-5 (v1)	17
	4.3 Station Lists for CEMS EFFIS	18
	4.4 Indicators for CEMS EDO	18
5	Gap Analysis	19
	5.1 Gap Analysis	19
	5.2 Proposal for future data collection strategy	21
6	Summary	22
Lis	st of figures	23
Lis	t of tables	24
Re	ferences	25

Acknowledgements

The EFAS team, and particularly the MDCC, 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 the MDCC, 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.

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 the CEMS EFAS (European Flood Awareness System), EFFIS (European Forest Fire Information System) and EDO (European Drought Observatory). By the end of 2021, real-time data delivered by 30 providers for 11 parameters from more than 23,000 stations are collected. On average, 7,000,000 records are being added to the database each day. Additionally, the database comprises more than 44,000 stations with historical data.

All data, real-time and historical, are quality-controlled in the same manner during the import to the database. Only the data that passed the quality control are processed further as reliable input data for the CEMS. This post-processing includes the calculation of minimum, maximum and mean values as well as the aggregation of totals over different accumulation periods. Data were provided as station lists or grids including an uncertainty estimation to the CEMS. The grids are generated by means of a modified SPHEREMAP scheme through interpolation from the quality-controlled station data.

Within 2021, six additional real-time data providers were integrated as well as several historical data. The configuration of the database was improved to use more of the existing data during the product generation.

Furthermore, a gap analysis was done regarding the availability of data with a proposal for the future data collection strategy and extended use of existing data.

1 Description of CEMS METEO

CEMS METEO, the Meteorological Data Collection Centre, was established to provide application tailored quality- controlled surface meteorological 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 EFAS (European Flood Awareness System), EFFIS (European Forest Fire Information System) and EDO (European Drought Observatory). To fulfil this task, CEMS METEO collects in situ meteorological observations as well as ground-based remotesensed data like radar observations of precipitation. The data are received from many data providers using various channels, e.g. ftp-server, APIs or via email, and different file formats. All received data are quality controlled and integrated into a database. The data is post-processed to the needs of the CEMS applications. This comprises the calculation of minimum, maximum, and mean values, the aggregation and disaggregation of totals. Depending on the component, the data are either provided as station data or gridded fields.

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 2021, 30 data providers deliver real-time data to CEMS METEO. The data providers and delivered parameters are summarised in the Table 1.

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

Name	Parameter										
	CICo	DT	Evap	Precip	ReAiHu	SunRad	SunD	АТ	٨P	WDir	WSpeed
Agencia Estatal de Meteorología (Spain)	-	x	-	x	x	x	х	х	-	х	x
Deutscher Wetterdienst (Germany)	х	x	-	x	-	-	x	x	-	x	x
Agenzia Regionale per la Prevenzione e l'Ambiente dell'Emilia-Romagna (Italy)	-	-	-	x	-	-	-	x	-	-	-
Slovenian Environment Agency	-	-	-	x	x	x	-	x	-	х	х
Czech Hydro- Meteorological Institute	-	-	-	x	-	-	-	х	-	-	-
Deutscher Wetterdienst (global)	х	x	-	x	x	x	-	x	-	x	х
Environment Agency (England)	-	-	-	x	-	-	-	-	-	-	-
Federal Hydrometeorological Institute (Bosnia-Herzegovina)	-	-	-	x	-	-	-	х	-	-	-
Finnish Meteorological Institute	х	x	-	x	x	-	-	x	-	x	х
Hydrological Information Centre (HIC)- Flanders Hydraulics Research (Belgium)	-	-	-	x	-	-	-	-	-	-	-
Hungarian Meteorological Service	-	-	-	x	х	x	-	х	-	х	x
Institute of Meteorology and Water Management (Poland)	х	x	-	x	-	x	x	x	-	x	x
Israel Meteorological Service	х	x	x	x	x	x	х	х	-	х	x

Institute for Ocean and Atmosphere (Portugal)	-	x	-	x	x	x	-	x	-	x	x
Kosovo Hydrometeorological Institute	-	-	-	х	х	x	-	х	-	х	х
Royal Netherlands Meteorological Institute	x	x	-	x	x	x	x	х	-	x	x
Agenzia Regionale Di Protezione Civile (Lazio, Italy)	-	-	-	x	x	x	-	х	-	x	x
National Environmental Agency (Georgia)	x	x	-	x	-	-	-	х	-	x	x
MARS ¹ (global)	x	-	-	x	-	x	-	х	x	-	x
MeteoLux (Luxembourg)	x	x	-	x	x	-	х	х	-	х	х
MeteoSchweiz (Switzerland)	х	x	x	x	x	x	x	x	-	x	x
Met Éireann (Ireland)	-	-	-	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	-	-	-
Automatic System of Hydrological Information (SAIH) for the Ebro river basin (Spain)	-	-	-	х	x	x	-	х	-	х	x
Slovak Hydro-Meteorological Institute	-	-	x	x	-	-	-	x	-	-	-
Servei Meteorològic de Catalunya (Spain)	-	-	-	x	x	x	-	х	-	x	х
Swedish Meteorological and Hydrological Institute	x	-	-	x	x	x	x	x	-	x	x
Zentralanstalt für Meteorologie und Geodynamik (Austria)	-	-	-	x	-	-	-	x	-	-	-

¹ Monitoring Agricultural ResourceS system, JRC, European Commission

Parameter abbreviation	Parameter description
CICo	Cloud cover
DT	Dew point temperature
Evap	Evaporation
Precip	Precipitation
ReAiHu	Relative Air humidity
SunRad	Solar radiation
SunD	Sunshine duration
AT	Air temperature
VP	Water vapour pressure
WDir	Wind direction
WSpeed	Wind speed

Table 2: Parameter abbreviation and description.

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. 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 and DWD (global). All the historical data integrated into the CEMS METEO database are summarised in Table 3.

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

Name	Parameter type										
	CICo	DT	Evap	Precip	ReAiHu	SunRad	SunD	АТ	VP	WDir	WSpeed
Agenzia Regionale per la Prevenzione e l'Ambiente dell'Emilia-Romagna (Italy)	-	-	-	x	x	х	-	х	-	x	x
Automatic System of Hydrological Information (SAIH) for the Ebro river basin (Spain)	-	-	-	х	-	-	-	х	-	-	-

CarpatClim	-	-	-	х	-	-	-	-	-	-	-
Czech Hydro-Meteorological Institute	-	-	-	x	-	-	-	х	-	-	-
Danube	-	-	-	х	-	-	х	х	-	-	-
Deutscher Wetterdienst (Germany)	х	x	-	x	-	-	х	x	-	х	х
DWD Climatic (Germany)	х	-	-	х	х	-	х	х	-	-	-
Environment Agency (England)	-	-	-	х	-	-	-	-	-	-	-
ERA-Interim-land	-	-	-	х	-	-	-	-	-	-	-
Euro Synop	х	х	-	х	-	-	-	х	-	х	х
EURO4M-APGD	-	-	-	х	-	-	-	-	-	-	-
European Climate Assessment and Dataset (ECAD)	-	-	-	x	-	-	-	-	-	-	-
Finnish Meteorological Institute	х	х	-	х	х	-	-	х	-	х	х
Hellenic National Meteorological Service	x	-	-	x	x	-	х	х	x	-	x
Hungarian Meteorological Service	x	х	-	x	x	-	х	х	-	x	x
Hungarian Meteorological Service	x	х	-	х	-	-	х	х	-	x	x
Institute for Ocean and Atmosphere (Portugal)	-	х	-	х	x	x	-	х	-	x	x
Institute of Meteorology and Water Management (Poland)	x	х	-	х	x	-	х	х	x	x	x
Kosovo Hydrometeorological Institute	-	-	-	х	x	x	-	х	-	х	x
MARS	-	-	-	х	-	х	-	х	х	-	х
Met Éireann	-	-	-	х	-	-	-	х	-	х	х
MeteoConsult	-	-	-	х	-	-	-	-	-	-	-
MeteoLux (Luxembourg)	х	х	-	х	х	-	х	х	-	х	х

MeteoSchweiz	х	х	х	х	х	х	х	х	-	х	х
National Environmental Agency Georgia	-	-	-	х	-	-	-	х	-	-	-
Norwegian Meteorological Institute	х	х	-	х	x	-	-	х	-	х	x
Servei Meteorològic de Catalunya (Spain)	-	-	-	x	x	x	-	x	-	x	x
Slovak Hydro-Meteorological Institute	-	-	x	х	-	-	-	х	-	-	-
Slowenian Environment Agency	-	-	-	х	х	х	-	х	-	х	х
Swedish Meteorological and Hydrological Institute	х	-	-	х	x	х	х	х	-	х	x
Zentralanstalt für Meteorologie und Geodynamik (Austria)	-	-	-	х	-	-	-	х	-	-	-

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 totals are often provided. Minimum and maximum air temperature, are mostly provided on a daily basis.

2.2 New Data Providers in 2021

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 example, for the calibration and validation of the hydrological model.

To enlarge the CEMS METEO database, these data providers were added within 2021:

- Agenzia Regionale Di Protezione Civile (Lazio, Italy)
- Hungarian Meteorological Service
- Hydrological Information Centre (HIC) Flanders Hydraulics Research (Belgium)
- Israel Meteorological Service (IMS)
- Republic Hydrometeorological Service of the Republic of Srpska (Bosnia and Herzegovina)
- Servei Meteorològic de Catalunya (Spain)

The increased station density in comparison to the previous status is shown in Figure 1. These six data providers deliver real-time (Table 1) and historical (Table 3) data.



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

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 optimized 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.

The provided data are stored in a data bank super-system 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

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 to feedback an added value on the data provision to METEO. As the quality control procedure is triggered by the import of real-time data, historical data and delayed data. The latest refers to data which are re-sent by data providers in order to replace the data already existing data in the database, and therefore are checked in the same manner. This guarantees the availability of 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 2).

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 measures, "suspect", if it is inconsistent with other parameters (e.g. dew point temperature higher than air temperature) and "rejected", if it didn't pass any 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 18UTC. 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. Bear in mind that one threshold value is used for the whole EFAS domain. Additionally, cross-validation procedure against data from other parameters at the same station is carried out.

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

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

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

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

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

Temperature data are checked against time-dependent limits taking the annual cycle into account. In winter, only data between -50°C and 25°C are used and in summer between -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".

3.2.1 Quality control statistics

Data are controlled and flagged as described in section 3.2. Table 3. 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 2021.

Parameter							
	good		suspect		rejected		
	relative (%)	absolute	relative (%)	absolute	relative (%)	absolute	
Precipitation (daily totals)	98.33	4159703	1.66	70355	0.008	345	
Precipitation (6hourly totals)	90.64	14336303	9.35	1478495	0.014	2251	
Relative air humidity	99.99	1346103	0	0	0.009	118	
Solar radiation	99.34	1590299	0.62	9995	0.040	638	
Temperature (daily maximum)	99.75	3773584	0	0	0.253	9573	
Temperature (daily minimum)	99.95	3724098	0	0	0.048	1776	
Temperature (6hourly mean)	99.75	14756239	0	0	0.25	36958	
Vapour pressure	100	1157643	0	0	0	0	
Wind speed	99.97	2803320	0.01	149	0.023	641	

3.3 Database Statistics

By the end of 2021, the CEMS METEO database contained more than 63,000 stations, of which more than 23,000 stations delivering real-time data. The other stations provided real-time data in previous times or only historical data. 1,324 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).

CEMS METEO receives approximately 60,000 data files per day. All these files are processed, leading to on average 7,000,000 data records added to the database – per day.

The spatial distribution of active 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.

Figure 3: Spatial distribution of active stations within the EFAS domain. These stations deliver real-time data of at least one parameter stored in CEMS METEO database. Maps per parameter are shown in Figure 9. The number of stations was given in brackets.

All the stations currently not delivering real-time data are classified as 'inactive'. Even if these stations don't 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 for historical periods (e.g. EMO-5). EMO-5 is a high-resolution multi- variable gridded meteorological data set for Europe, which is described in more detail in section 4.2.2. 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 within the EFAS domain. These stations deliver currently no realtime data, but data for former periods. The number of stations was given in brackets.

4 Post-Processing

The received data are post-processed to fulfil the needs of the various CEMS components. The postprocessing 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 comprises:

- Calculations of 6-hourly means/totals, except for solar radiation and water vapour pressure;
- Calculations of daily minimum, maximum and mean values, expect for water vapour pressure;
- Calculation of daily totals for precipitation and solar radiation;
- With 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: Scheme of aggregation and disaggregation 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

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 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 (World Meteorological Organization 2019). 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).

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

A minimum availability of data is required to compute minimum, maximum and mean values as well as aggregated totals – further referred as coverage. 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

4.2.1 Operational grids

The hydrological model for 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. If multiple stations are clustered, the estimated uncertainty within the grids is lower than outside the clusters.

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.

4.2.2 Historical grids: EMO-5 (v1)

In 2021, the archive between 1990 and 2019 has been re-run to produce a consistent historical data set across the EFAS domain in 5 x 5 km spatial resolution, considering seven parameters: total precipitation, temperatures (minimum, maximum and mean), wind speed, solar radiation and water vapour pressure. CEMS refers to this data set as EMO-5 (Thiemig, Gomes, et al., EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe 2022). EMO-5 is a European high-resolution, (sub-) daily, multi-variable meteorological data set built on historical and real-time observations obtained by integrating data from 18,964 ground weather stations, four high-resolution regional observational grids, as well as one global reanalysis. The four high-resolution grids are CombiPrecip (Sideris, et al. 2014), ZAMG - INCA (Haiden, et al. 2011), EURO4M-APGD (Isotta, et al. 2014) and CarpatClim (Spinoni, et al. 2015). The global reanalysis is ERA-Interim/Land (Balsamo, et al. 2015). EMO-5 includes at daily resolution: total precipitation, temperatures (minimum and maximum), wind speed, solar radiation and water vapour pressure. In addition, EMO-5 also makes available 6-hourly precipitation and mean temperature. The raw observations from the ground weather stations underwent a set of quality controls, before SPHEREMAP (Willmott, Rowe and Philpot 1985) and Yamamoto interpolation methods (Yamamoto 2000) were applied in order to estimate for each 5x5 km grid cell the variable value and its affiliated uncertainty, respectively.

EMO-5 (version 1) covers the time period from 1990 to 2019, with a near real-time release of the latest gridded observations foreseen with version 2. As a product of Copernicus, the EU's Earth observation programme. EMO-5 dataset is free and open. accessed and can be at https://doi.org/10.2905/0BD84BE4-CEC8-4180-97A6-8B3ADAAC4D26 (Thiemig, Gomes, et al., EMO-5: A high-resolution multi-variable gridded meteorological data set for Europe (1990–2019) 2020)

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.

4.4 Indicators for CEMS EDO

METEO started with the implementation of two indicators for the CEMS EDO in the post-processing - the Standardized Precipitation Index (<u>SPI</u>) and the Heat Cold Wave Indicator (<u>HCWI</u>).

The SPI is implemented on basis of the 5km precipitation grids. It aggregates monthly precipitation amounts over 1, 3, 6, 12, 24 and 48 months. The baseline period for the estimation of the distribution parameters is 1981 to 2010.

The HCWI is also implemented on basis of the 5km daily minimum and maximum temperature grids. As for the SPI, the baseline period is 1981 to 2010 for the estimation of the daily percentile thresholds.

The regular calculation of these indicators will get operational in 2022.

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 spatially homogeneous distributed. Within the EFAS domain the stations density varies largely across the 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 focus mainly on precipitation and only at a few stations observe 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 an advice for future data collection activities.

The backbone of the available data are 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 other countries. 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 catchments.

Figure 3 depicts the spatial distribution of real-time stations. By the end of 2021, a comparatively low station density is apparent in Algeria, Azerbaijan, Belarus, Bulgaria, Egypt, Greece, Iraq, Iran, Ireland, Jordan, Libya, Lithuania, Morocco, Poland, Russia, Saudi Arabia, Tunisia, Turkey and in some parts of Italy. On a parameter level (Figure 9), a comparatively low station density for precipitation occurs additionally in Iceland. 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 county. A medium solar radiation station density is in Armenia, Estonia, Iran, Latvia, Lithuania, Russia and Turkey.

Beside the addition of new stations, the use of existing data was optimised. This was done by further improving the configuration of the calculation procedures in the post-processing in order to utilize more uncommon reporting behaviours and aggregation periods for totals (e.g. 9- or 15-hourly precipitation totals). Additionally, the calculation of non-provided parameters from provided parameters, for example the water vapour pressure from air temperature and dew point temperature, increase the amount of available data and provides a redundancy of ready data.

Figure 9: 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).

5.2 Proposal for future data collection strategy

Based on the gap analysis, a high priority should be given to add additional stations from countries marked blue in Figure 10. These are in alphabetic order Algeria, Azerbaijan, Belarus, Bulgaria, Egypt, Greece, Iraq, Iran, Ireland, Italy, Jordan, Libya, Lithuania, Morocco, Poland, Russia, Saudi Arabia, Tunisia and Turkey.

Redundant data deliveries are currently implemented for countries marked in dark green in Figure 10. In alphabetic order are these 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 10.

Figure 10: 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 p arts of the country (light green) to METEO.

Precipitation is of high importance in the hydrological modelling as critical input variable and due to its high spatial and temporal variability. Solar radiation is provided by a limited and inhomogeneous distributed number of stations. 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 and redundant delivery of solar radiation data.

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

6 Summary

This report provides an overview about the status and progress in 2021 of the Copernicus Emergency Management Service (CEMS) Meteorological Data Collection Centre (METEO).

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. All real-time and historical data are quality-controlled in the same manner during the import to the database. Only such data passed the quality control are processed further and provided as station lists or grids including an uncertainty estimation.

A gap analysis was done regarding the availability of data. A list of countries is provided, which inclusion could improve the data collection.

Within 2021, six additional real-time data providers were integrated as well as several historical data. The configuration of the database was improved to use more of the existing data.

By the end of 2021, real-time data are delivered by 30 providers for 11 parameters from more than 23,000 stations. On average, 7,000,000 records are being added to the database each day. Additionally, the database comprises more than 44,000 stations with historical data.

List of figures

Figure 4: Spatial distribution of inactive stations within the EFAS domain. These stations deliver currently no real-time data, but data for former periods. The number of stations was given in brackets.

Figure 6: Delivered instantaneous air temperature (red) and therefrom calculated daily minimum temperature (blue), daily maximum temperature (magenta) and 6-hourly mean temperature (green). 16

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... 17

List of tables

Table 1: List of active real-time data providers and delivered parameters (abbreviation). Parameter abbreviations are explained in Table 2.	4
Table 2: Parameter abbreviation and description.	6
Table 3: List of data providers with historical data and delivered parameters (abbreviation). Parameters abbreviations are given in Table 2	ter 6
Table 4: Parameters with fixed thresholds. For precipitation see Table 5	. 11
Table 5: Thresholds for precipitation depending on the aggregation interval.	. 11
Table 6: Quality control statistics of the quality codes "good", "suspect" and "rejected" for each parameter, given as relative and absolute values in 2021.	. 12

References

Balsamo, G., C. Albergel, A. Beljaars, S. Boussetta, E. Brun, H. Cloke, D. Dee, et al. 2015. "ERA-Interim/Land: a global land surface reanalysis data set." *Hydrol. Earth.*

Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. and Gruber. 2011. "The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region." *Weather Forecast* 26: 166-183. doi: https://doi.org/10.1175/.

Isotta, F., C. Frei, V. Weilguni, M. Lassegues, P. Tadic, B. Rudolf, V. Pavan, C. Cacciamani, et al. 2014. "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* (34): 1657-1675. doi:doi: 10.1002/joc.3794.

Sideris, I. V., M. Gabella, R. Erdin, and U. and Germann. 2014. "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.* 140: 1097–1111. doi: https://doi.org/10.1002/qj.2188.

Spinoni, J., S. Szalai, T. Szentimrey, M. Lakatos, Z. Bihari, A. Nagy, Á. Németh, et al. 2015. "Climate of the Carpathian Region in the period 1961–2010: climatologies and trends of 10 variables." *Int. J. Climatol* 35: 1322–1341. doi:https://doi.org/10.1002/joc.4059.

Thiemig, V., G. N. Gomes, J. O. Skøien, M. Ziese, A. Rauthe-Schöch, E. Rustemeier, K. Rehfeldt, et al. 2020. "EMO-5: A high-resolution multi-variable gridded meteorological data set for Europe (1990–2019)." (European Commission, Joint Research Centre (JRC) [data set], https://doi.org/10.2905/0BD84BE4-CEC8-4180-97A6-8B3ADAAC4D26). European Commission, Joint Research Centre (JRC) [data set], https://doi.org/10.2905/0BD84BE4-CEC8-4180-97A6-8B3ADAAC4D26.

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

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

World Meteorological Organization. 2019. 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). Geneva, Switzerland.

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

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 (<u>european-union.europa.eu/contact-eu/meet-us_en</u>).

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 (<u>european-union.europa.eu</u>).

EU publications

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

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 (<u>eur-lex.europa.eu</u>).

Open data from the EU

The portal <u>data.europa.eu</u> 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.

The European Commission's science and knowledge service Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

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

- 🥑 @EU_ScienceHub
- **f** EU Science Hub Joint Research Centre
- in EU Science, Research and Innovation
- EU Science Hub

O EU Science