European Commission

Database of European Forest Insect & Disease Disturbances – DEFID2

Protocol for data contribution

Version 2.1 | August 2023

Coordinators:
Pieter S. A. Beck, European Commission, Joint Research Centre, Ispra, Italy
Giovanni Forzieri, University of Florence, Italy
Alessandro Cescatti, European Commission, Joint Research Centre, Ispra, Italy

Advisory board:
Henrik Hartmann, Max-Planck-Institute for Biogeochemistry and Julius Kühn-Institute, Germany
Andrea Battisti, University of Padova, Italy
Constantin Nistor, University of Bucharest, Romania

Ips typographus outbreak in Lower Saxony due to climate change/abandoned spruce plantations. Bienenfreund2018 / CC BY-SA.

Joint Research Centre

1 Rationale and aim

European forests provide a set of fundamental services that contribute to climate mitigation and human well-being. At the same time, forests are vulnerable systems because the long life-span of trees limits their ability to rapidly adapt to drastic environmental changes. Natural disturbances – large pulses of tree mortality that originate from climate-related abiotic and biotic agents such as fires, strong winds or insect and disease outbreaks – are drivers of many ecological processes and can impact the provision of forest services and products, particularly if they are exacerbated by climate change. Of particular concern are insect and disease disturbances as they affect tens of millions of hectares annually, particularly in temperate regions in the Northern Hemisphere. Impacts associated with the ensuing forest cover losses are expected to rise drastically with global warming and key forest ecosystem services, such as carbon sequestration and the supply of wood products, could be considerably compromised in the near future. Examples from other countries, such as Canada and the United States, show that biotic disturbances driven by the rapidly changing climate may substantially alter the state of forests to the point that they become carbon sources instead of carbon sinks.

Despite the impact of insect and disease outbreaks, spatially detailed reference records of such disturbances are, traditionally, hard to find, including in Europe. This hampers the research on changing biotic disturbance regimes on several fronts; reference data are critical to better use Earth Observation for detecting and monitoring pest outbreaks, or for incorporating these phenomena in Earth system models, for example. To address this, **the Database of European Forest Insect & Disease Disturbances (DEFID2[1]) collects and makes available spatially explicit records of insect and disease outbreaks in European forests**.

DEFID2 is a joint and voluntary effort among scientists to share and harmonize their geospatial observations of insect and disease disturbances. As of August 2023, scientists from 30 different European research institutes have contributed their data to DEFID2 and the database currently comprises more than 650,000 disturbance records, from eight different countries. The records were acquired through both ground surveys and photogrammetric techniques and are mapped as geospatial polygons or points. Where available, DEFID2 provides for individual records information on the severity and patterns of damage symptoms, agents, host tree species, climate-driven trigger factors, silvicultural practices, and eventual sanitary interventions. For every data record, the name, affiliation, and contact details of the data contributor are also provided in the database. The DEFID2 database is publicly available (under a CC-BY 4.0 license) in both full and simplified formats, in sqlite and geopackage format, respectively. Furthermore, a dedicated R software package defid2R makes it easy to download, explore, and use the data.

DEFID2 is coordinated by the JRC[2] who maintains the database, assists data contributors, and ensures that new data contributions are integrated into future versions of the DEFID2 database.

The DEFID2 community welcomes and encourages new data contributions and this protocol serves as guidance for those interested in contributing insect and disease disturbance data to the DEFID2 initiative. Any further questions and expressions of interest can be addressed to JRC-DEFID2@ec.europa.eu as the JRC offers technical assistance to help data providers with data restructuring, reformatting, or harmonization.

---

[1] Forzieri, G., Dutrieux, L. P., Elia, A., Eckhardt, B., Caudullo, G., Álvarez Taboada, F., Andriolo, A., Bălacenoiu, F., Bastos, A., Buzatu, A., Castedo Dorado, F., Dobrovolný, L., Duduman, M.-L., Fernandez-Carillo, A., Hernández-Clemente, R., Hornero, A., Ionut, S., Lombardero, M. J., Junttila, S., Lukeš, P., Marianelli, L., Mas, H., Mlčoušek, M., Mugnai, F., Netoiu, C., Nikolov, C., Olenici, N., Olsson, P.-O., Paoli, F., Paraschiv, M., Patočka, Z., Pérez-Laorga, E., Quero, J. L., Rüetschi, M., Stroheker, S., Nardi, D., Ferencik, J., Battisti, A., Hartmann, H., Nistor, C., Cescatti, A., Beck, P. S. A., 2023. The Database of European Forest Insect and Disease Disturbances: DEFID2. *Global Change Biology*, DOI: 10.1111/gcb.16912

[2] The JRC conducts research on a range of forest-related topics (https://forest.jrc.ec.europa.eu/en/) and is currently developing the EU Observatory on Deforestation, Forest Degradation, Changes in Forest Cover and Associated Drivers, as part of EU actions to protect and restore the world's forests. Its contribution to DEFID2 fits in the Observatory's commitment to develop Earth-Observation-based monitoring tools for forests (COM/2021/572 final).

2 Protocol for data contribution

In the following lines, we described the DEFID2 common protocol that data providers are encouraged to follow in order to assure consistency. To facilitate this, **the JRC offers technical assistance to help data providers with data restructuring, reformatting, or harmonization**.

- **Contact point.** Questions and data can be sent to JRC-DEFID2@ec.europa.eu.
- **Contents and acquisition methods.** The protocol covers information on the insect(s), the pathogen(s), the host(s), and the spatial extents of the forest areas affected by insect disturbances. Disturbed areas can be derived from ground surveys, visual interpretation of aerial or satellite imagery or some form of automatic classification algorithms of remote sensing data.
- **Spatial and temporal coverage**: records of interests should geographically fall within geographic Europe, European Russia, Northern Africa and Middle East in the 1981-to-present period.
- **Geographical format.** Each record of insect disturbance should be represented as a spatial feature in a GIS-compatible vector format such as shapefile (.shp), GeoJSON (.json), KML, GML, Geopackage (.gpkg), etc. The preferred spatial geometries for DEFID2 are "exact polygons" that precisely outline the disturbed forest areas. However, the database also accommodates geometries of the form "exact point" (being a geographical point geometry located exactly within a disturbed forest patch), "substitute polygon" (meaning the disturbance occurred within the provided polygon geometry but does not cover its entire extent) and "substitute point" (which is a point geometry with a loose spatial relation to the disturbance event). The spatial reference system should be preferably in geographical coordinates, i.e. latitude and longitude, in WGS84 (EPSG:4326). Other systems are accepted if information about the used spatial reference are provided (preferably an EPSG code). For each feature, information about the disturbance should be provided in the attributes of the vector file using specific fields and data formats (see Table 2). In case the same disturbance event is monitored through time, the evolution of the extents of the disturbed forest area should preferably be tracked by separate features, one for each observation date/year. In case data are provided in raster format (e.g., output of classification of remote sensing data), the JRC will transform the data into vector format using conventional segmentation tools.
- **Descriptive attributes.** The selection of the descriptive attributes to be reported in the vector file is inspired by the National Insect and Disease Survey (IDS, http://foresthealth.fs.usda.gov) database of the United States Department of Agriculture (USDA). In summary, for each disturbed forest feature, four different damage types can be recorded: discoloration, defoliation, dieback, mortality (see Table 1). Each damage type can be characterized by a different severity/pattern of damage. Furthermore, each forest feature can be characterized by one or more agents, by one or more affected host tree species, by the two dominant climate-driven triggering factors, by silvicultural practices and eventual sanitary interventions.

Table 1. Damage types and descriptions according to the National Insect and Disease Survey (USDA Forest Service 1999).

Damage type	Description
Discoloration	Foliage is a colour other than green, such as yellow, red, purple, black, or brown.
Defoliation	Damage that results in physical or functional removal of foliage, partially or wholly by some agent.
Dieback	Distal portions of branches in upper part of crown killed, with dead or dying (yellow, red, or brown) foliage. Non-distal portions of these branches have visible green, live foliage.
Mortality	Standing dead trees.

The descriptive attributes are grouped into four sets (Table 2):

- Set 1: information about the contributor and the data source.
- Set 2: key information about the disturbance.
- Set 3: complementary information mostly related to occurrences characterized by multiple agents or multiple hosts, climate-driven triggering factors and silvicultural practices.
- Set 4: qualitative assessment of the damage.

Table 2 provides names and descriptions of each attribute field.

Table 2. List of descriptive attributes. The short field names are there to accommodate those who wish to contribute data in the form of ESRI's shapefiles, which do not allow for field names with more than 10 characters.

Set	Field name	Short field name	Description	Field Type	Values and formats
1	country	COUNTRY	Country of damage	Text	Country name
1	data_provider	PROVIDER	Data provider(s) responsible for the data	Text	Name1 Surname1, Name2 Surname2, …
1	email	EMAIL	Contact email(s)	Text	email1, email2, …
1	affiliation	AFFILIAT	Affiliation(s) of data provider(s)	Text	Name and address of university, research institute, or agency
1	source	SOURCE	Original source of the data	Text	e.g. website of the institution, paper published, geoportal, etc.
2	survey_start_date	SURVEYSTA	Start date of damage survey[1]	Date	YYYY-MM-DD
2	survey_end_date	SURVEYEND	End date of damage survey[1]	Date	YYYY-MM-DD
2	survey_method	METHOD	Data acquisition method	Text	<Null> = No Data Aerial photointerpretation Satellite photointerpretation Remote sensing classification Field surveys
2	agents	AGENTS	List of disturbance agents	Text	Scientific names separated by commas (e.g.: Ips typographus, Thaumetopoea pityocampa)
2	hosts	HOSTS	List of affected hosts	Text	Scientific names separated by commas (e.g.: Picea abies, Pinus halepensis)
2	symptoms	SYMPTOMS	Comma-separated list of symptoms	Text	Discolouration, Defoliation, Dieback, Mortality
2	severity_discoloration	SEV_DIS	Discoloration severity	Integer	<Null> = No Data 1 = Low (≤ 50% discoloration) 2 = High (> 50% discoloration)
2	severity_defoliation	SEV_DEF	Defoliation severity	Integer	<Null> = No Data 1 = Low (≤ 50% defoliation) 2 = High (> 50% defoliation)
2	severity_dieback	SEV_DIE	Dieback severity	Integer	<Null> = No Data 1 = Low (≤ 50% dieback) 2 = High (> 50% dieback)
2	severity_mortality	SEV_MOR	Mortality severity	Integer	<Null> = No Data 1 = Marginally affected (killed trees ≤ 20%) 2 = Moderately affected (20% < killed trees ≤ 40%) 3 = Substantially affected (40% < killed trees ≤ 60%) 4 = Highly affected (60% < killed trees ≤ 80%)

					5 = Totally affected (80% < killed trees ≤ 100%)
3	trigger_primary	TRIGGER1TY	Primary trigger of damage	Text	<Null> = No Data Drought Heatwave Wind/Storm Fire Snow/Ice Pest/Disease
3	trigger_secondary	TRIGGER2TY	Secondary trigger of damage	Text	<Null> = No Data Drought Heatwave Wind/Storm Fire Snow/Ice Pest/Disease
3	trigger_prim_date_start	TRIGG1STA	Start date of primary trigger of damage[1]	Date	<Null> = No Data YYYY-MM-DD
3	trigger_prim_date_end	TRIGG1END	End date of primary trigger of damage[1]	Date	<Null> = No Data YYYY-MM-DD
3	trigger_secon_date_start	TRIGG2STA	Start date of secondary trigger of damage[1]	Date	<Null> = No Data YYYY-MM-DD
3	trigger_secon_date_end	TRIGG2END	End date of secondary trigger of damage[1]	Date	<Null> = No Data YYYY-MM-DD
3	silvicultural_system	FM_SILVITY	Type of silvicultural system of the damaged forest stand	Text	<Null> = No Data Clear cut Shelterwood Selective logging None
3	sanitary_intervention	FM_SANITY	Type of sanitary intervention	Text	<Null> = No Data Clear cut Selective logging None
3	sanitary_interv_ date_start	FM_SAN_ST	Start date of sanitary intervention[1]	Date	<Null> = No Data YYYY-MM-DD
3	sanitary_interv_ date_end	FM_SAN_EN	End date of sanitary intervention[1]	Date	<Null> = No Data YYYY-MM-DD
4	pattern_defoliation	PATTERN_DE	Pattern of defoliation damage	Text	<Null> = No Data High-contiguous = Host species accounts for > 50% of the stand and

					the damage is contiguous High-patchy = Host species accounts for > 50% of the stand and the damage is patchy Low-contiguous = Host species accounts for ≤ 50% of the stand and the damage is contiguous Low-patchy = Host species accounts for ≤ 50% of the stand and the damage is patchy
4	pattern_discoloration	PATTERN_DI	Pattern of discoloration damage	Text	<Null> = No Data High-contiguous = Host species accounts for > 50% of the stand and the damage is contiguous High-patchy = Host species accounts for > 50% of the stand and the damage is patchy Low-contiguous = Host species accounts for ≤ 50% of the stand and the damage is contiguous Low-patchy = Host species accounts for ≤ 50% of the stand and the damage is patchy
4	pattern_mortality	PATTERN_MO	Pattern of mortality damage	Text	<Null> = No Data High-contiguous = Host species accounts for > 50% of the stand and the damage is contiguous High-patchy = Host species accounts for > 50% of the stand and the damage is patchy Low-contiguous = Host species accounts for ≤ 50% of the stand and the damage is contiguous Low-patchy = Host species accounts for ≤ 50% of the stand and the damage is patchy
4	pattern_dieback	PATTERN_DB	Pattern of dieback damage	Text	<Null> = No Data High-contiguous = Host species accounts for > 50% of the stand and the damage is contiguous High-patchy = Host species accounts for > 50% of the stand and the damage is patchy Low-contiguous = Host species accounts for ≤ 50% of the stand and the damage is contiguous Low-patchy = Host species accounts for ≤ 50% of the stand and the damage is patchy

1 = If the precise date is known, the start date should equal the end date. If the date is unknown during a month, season, or year, the start date should be the first day of that month/season/year, and the end date should be the last day of that month/season/year.

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

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union