IPM Decisions

D4.2 Data set for DSS evaluation collected: septoria tritici blotch and grapevine downy mildew

Grant agreement number: 817617

Start date of project: 01/06/2019 **End date of project:** 31/05/2024

Deliverable due date: 01/10/2020 **Date of delivery:** 28/09/2020

Classification: Public

Version History

Version number	Implemented by	Reason
1.0	Niels Holst	Deliverable 4.2 – septoria tritici blotch and grapevine downy mildew data sets.

List of Authors

Company	Author(s)	Contribution
AU	Niels Holst	Author

List of Reviewers

Company	Reviewer	Contribution
ADAS	Mark Ramsden	Formatting and checking consistency with other deliverables

Contents

1 Public Summary

Data from fungicide trials in wheat and grapes were kindly provided by BASF and Corteva. The data will be used to assess decision support systems providing guidance on the control of the diseases *Septoria tritici blotch* in wheat and *grape downy mildew* in grapes.

2 Executive Summary

The collected data sets on wheat (covering 12 countries over 5 years) and grapes (covering 8 countries over 7 years) will be used within WP4 to test the validity of decision support models for *Septoria tritici blotch* and *grape downy mildew* (D4.3) and to evaluate the economic and environmental benefits of same models (D4.13). The data sets include data on both treated and untreated trial plots. Parts of these data sets will later be made publicly available for DSS validation (D4.8).

Horizon 2020

3 Septoria tritici blotch (STB)

3.1 Data delivered from BASF

Data were received in the form of one Excel file.

3.1.1 Observations

We retrieved 1,961 records from the Excel file which were put in an R data frame with 12 columns (Table 1).

Table 1. Record layout of data set on septoria tritici blotch obtained from BASF.

Record field	Description
TrialId	Identifier for the trial.
Country	Country of observation.
Year	Year of observation.
Latitude	Latitude of observation.
Longitude	Longitude of observation.
Treatment	Treatment code.
Date	Date of observation.
Daft	No. of days after treatment.
GsFrom	Min. crop growth stage observed.
GsTo	Max. crop growth stage observed.
Organism	Pest organism observed.
Value	Value observed (% infection).

Septoria tritici blotch (STB) was the most frequently recorded disease organism, making 1,163 records out of the total 1,961 records. The STB records covered 12 countries from 2017 to 2019 (Table 2).

Table 2. Number of Septoria tritici blotch (STB) records from each country and year in the BASF data set (total = 1,163).

Country	Year		
	2017	2018	2019
Austria	16	0	0
Bulgaria	0	0	80
Czech Republic	0	0	23
Denmark	0	0	26
France	75	15	68
Germany	114	24	44
Ireland	0	0	47
Netherlands	35	16	0
Poland	54	49	109
Romania	0	0	48
Slovakia	0	0	33
United Kingdom	90	47	150

3.2 Data delivered from Corteva

Data were received in the form of five Excel files, each with several sheets detailing field trial observations and geographical information for each field site.

3.2.1 Observations

We retrieved 112,270 records from the Excel files, which were put in an R data frame with 12 columns (Table 3). The records comprised 58,477 observations on diseases and 53,793 observations on the crop (58,477 + 53,793 = 112,270).

Table 3. Record layout of data set on septoria tritici blotch obtained from Corteva.

Record field	Description
Field	Identifier for the trial.
Country	Country of observation.
Year	Year of observation.
Latitude	Latitude of observation.
Longitude	Longitude of observation.
Treatment	Treatment code, either UNTREATED or otherwise.
Date	Date of observation.
Daft	No. of days after treatment with TreatmentCode.
TreatmentCode	Code for latest treatment, A, B, C or D.
GsFrom	Min. crop growth stage observed.
GsTo	Max. crop growth stage observed.
Disease	Name of disease (see Table 4); missing for Crop=FALSE.
Crop	TRUE (for crop observations) or FALSE (for disease observations).
UserNote	Free-form supplementary information.
SamplingUnit	Unit sampled, such as leaf number or whole plot.
EvaluationType	Kind of observation, such as INFECT for pests and YIELD_STD for crop.
Value	Value observed for pest or crop.

Septoria tritici blotch (STB) was the most frequently observed disease (Table 4).

Table 4. Frequency of diseases observed in the septoria tritici blotch data set obtained from Corteva (total = 58,477).

Disease	Frequency
Septoria tritici blotch	45,384
Yellow rust	6,857
Brown rust	2,435
Powdery mildew	2,199
Tan spot	1,060
Eyespot	266
Unknown	120
Fusarion head blight	60
Snow mould	48
Rust	48

The *EvaluationType* most relevant for our work is INFECT, which reports percentage severity of the disease, and YIELD_STD which reports yield (t/ha) standardised to 84% grain moisture content.

The STB records covered 7 countries from 2014 to 2018 (Table 5).

Table 5. Number of STB records from each country and year in the data set from Corteva (total = 45,384).

Country	Year				
	2014	2015	2016	2017	2018
Denmark	759	253	0	331	124
France	1,903	1,864	2,375	2,342	3,482
Germany	1,972	3,064	2,144	1,584	1,380
Ireland	170	561	872	268	482
Poland	116	364	816	364	160
Sweden	0	436	0	0	0
United Kingdom	3,674	3,424	5,491	2,374	2,235

4 Grapevine downy mildew (GDM)

4.1 Data delivered from BASF

Data were received in the form of two Excel files, one with field trial observations and one with specifications for the field trial locations, including geographical coordinates.

4.1.1 Observations

We retrieved 411 field trial records from the Excel file, which were put in an R data frame with 12 columns (Table 6).

Table 6. Record layout of data set on grapevine downy mildew obtained from BASF.

Record field	Description
Field	Identifier for the trial.
Country	Country of observation.
Year	Year of observation.
Latitude	Latitude of observation.
Longitude	Longitude of observation.
Date	Date of observation.
GsFrom	Min. crop growth stage observed.
GsTo	Max. crop growth stage observed.
Method	P%FREQ or P%INF.
PlantPart	LEAF, RACEME or FRUIT.
Untreated	Value observed in the untreated plots.
Treated	Value observed in the treated plots.

The *PlantPart* column refers to leaves (LEAF), inflorescences (RACEME) and grape clusters (FRUIT). The disease (GDM) was assessed by two methods (*Method* column), designated P%FREQ and P%INF:

- P%INF: Intensity of attack was obtained as a visual estimation of the percentage of each plant part (leaf, raceme or grape cluster) affected by disease.
- *P%FREQ*: Frequency of attack was obtained as the number of infected plant parts (leaf, raceme or grape cluster). The frequency was expressed as a percentage of the number sampled.

The GDM records covered 8 countries and 7 years between 2012 and 2019 (Table 7).

Table 7. Number of grapevine downy mildew (GDM) records from each country and year in the data set from BASF (total = 411).

Country	Year						
	2012	2013	2015	2016	2017	2018	2019
France	7	5	14	4	55	0	73
Germany	0	2	6	3	8	2	0
Greece	53	0	32	0	14	0	0
Hungary	0	0	9	0	0	0	0
Italy	0	0	0	0	8	0	0
Portugal	0	0	24	0	24	0	0
Slovakia	0	0	40	0	0	0	0
Spain	0	0	12	0	16	0	0

5 Weather data

Weather data were obtained for every field trial (location and year) found in the three data sets above (1.1, 1.2, 2.1). We used publicly available data from the Copernicus server (www.copernicus.eu). The Copernicus *reanalysis-era5-land* data set provided weather data on a spatial resolution of 0.1 degree and a time resolution of 1 hour. The downloaded variables comprised

- 2m_dewpoint_temperature
- 2m_temperature
- surface_pressure
- surface_solar_radiation_downwards
- total_precipitation

From this the relative humidity of the air (RH) was computed as

$$RH = 100\frac{svp(\text{2m\_dewpoint\_temperature})}{svp(\text{2m\_temperature})}$$

where

$$svp(T) = 133.32\exp\left(\frac{1.0887\,(T+T_0)-276.4}{0.0583\,(T+T_0)-2.1938}\right)$$

with $T_0 = 273.15$ K.

All weather data was collected in an R data frame with columns containing the variables detailed above and, in addition, columns for latitude, longitude, date (dd/mm/yyyy) and hour of the day (1..24). A total of 101 location×year weather data sets were collated in the weather data file holding a total of 884,904 hourly records.

6 Data availability

The data were collated in four R binary files which can be read by the R *load* function. Each file contains an R data frame with the same name as the file. The files are all stored confidentially by WP4. However, parts of data sets are expected to be made public later on to fulfill deliverable D4.8.

6.1 septoria_basf.Rdata

An R data frame (*septoria_basf*) with 12 columns and 1,961 records containing the STB data set collated from the data obtained from BASF.

6.2 septoria_corteva.txt

An R data frame (*septoria_corteva*) with 12 columns and 112,270 records containing the STB data set collated from the data obtained from Corteva.

6.3 gdm_basf.Rdata

An R data frame (*gdm_basf*) with 12 columns and 411 records containing the GDM data set collated from the data obtained from BASF.

6.4 weather.Rdata

An R data frame (*weather*) with 11 columns containing latitude, longitude, date, hour and seven climatic variables on an hourly time resolution, making a total of 884,904 records.