
USING MAIZE FINDER

Jens Riis-Jacobsen



MaizeFinder

 CIMMYT_{MR}

Welcome to MaizeFinder	3
What you can do with MaizeFinder.....	3
How to install MaizeFinder	Error! Bookmark not defined.
Quick start tutorial	Error! Bookmark not defined.
Exercise 1: Find information on a specific maize variety.....	Error! Bookmark not defined.
Exercise 2: Find maize with specific characteristics	Error! Bookmark not defined.
Exercise 3: Systematic comparison of different maize varieties .	Error! Bookmark not defined.
Exercise 4: Find maize variety that matches local conditions	Error! Bookmark not defined.
The MaizeFinder information system.....	4
Overview of MaizeFinder	4
Head-to-head analysis.....	7
Introduction to head-to-head analysis	7
Details on head-to-head functionality in MaizeFinder	8
Queries in MaizeFinder	12
Introduction to Queries in MaizeFinder.....	12
Specific Query functionality	12
GIS in MaizeFinder	20
Background on GIS.....	20
Using GIS in MaizeFinder	20
Reports in MaizeFinder	31
Background on reports	31
Specific report functionality	31
Tables and CSV files.....	34
Introduction to the data table and CSV files.....	34
Specific table and CSV functionality.....	34
ANNEX 1 QUERY CRITERIA DEFINITIONS	ERROR! BOOKMARK NOT DEFINED.

Welcome to MaizeFinder

What you can do with MaizeFinder

MaizeFinderPC is designed to help maize breeders and other users of information from maize breeding research to find information they need quickly and easily. It provides very powerful search facilities, which enables you to combine different criteria in order to quickly extract the most relevant data. It also permits you to select exactly the types of data you would like to include in the output.

MaizeFinderPC includes a simple GIS which enables you to easily visualize the different locations which supplied the information included in your query. It also gives you the possibility to combine the information on the trial sites with other types of information contained in GIS layers such as soil maps or maps with maize mega environments. The GIS tools can help you to quickly determine how relevant the data are for your purposes.

MaizeFinder can also help you to systematically compare different maize varieties. The tool for this is head-to-head analysis where two materials are compared by collecting all the data from the trials where the materials have been grown at the same time and with the same management. This way you can combine data across different years and statistically test in what aspects two different materials vary.

MaizeFinder can provide you with many different answers to a particular question, but do not expect it to provide *the* answer. Plant breeding is a game of probabilities, and hopefully the easy data access in MaizeFinder can improve your chances of making better decisions, however, it is still based on probabilities and not certainty. For the best use of MaizeFinder try to frame your question in different ways and make sure to obtain different types of outputs. Then use your professional experience to interpret the data.

The above-mentioned functionality can be combined in many different ways to answer many different questions. The following are some examples, which were considered when designing MaizeFinder

1. Find maize varieties to match local conditions

1.1 A seed-distribution agency is searching for the best stress tolerant, responsive open pollinated varieties or hybrids currently available in order to distribute quality maize seeds to farmers in a specific region.

1.2 A private seed company wants to target their marketing of maize varieties or hybrids towards regions where the varieties will perform well, i.e., provide high yields under the stress factors that are present in farmers' fields.

2. Assess performance of new varieties under different conditions

2.1 A breeder in National Agricultural Research Program (NARP) is searching for the best stress tolerant, responsive open pollinated varieties or hybrids for testing in National Maize Evaluation Trials.

2.2 A breeder in NARP needs information on performance of a variety when stressed by drought or nitrogen shortage.

2.3 A private seed company wants to compare the performance of the company's varieties to other varieties under different environmental conditions and different types of management.

2.4 A breeder in a private seed company is searching for germplasm with potential for commercialization.

3. Search for germplasm

3.1 A breeder in NARP or in a private seed company is searching for germplasm with specific traits to use his or her in breeding program.

3.2 A breeder would like to compare the general combining ability of two elite lines in hybrid combinations under similar management conditions

4. Compare different maize varieties

4.1 A NARP breeder needs for trial data for a new variety in order to compare it with existing varieties and shorten the approval process.

You can explore these examples in more detail in the Quick start tutorials bellow.

The MaizeFinder information system

This chapter presents an overview of MaizeFinder and its relation to Fieldbook.

Overview of MaizeFinder

The International Center for Maize and Wheat Improvement, CIMMYT, is a center of excellence for maize and wheat breeding with a strong focus on producing improved germplasm. In addition to improved varieties, CIMMYT produces large volumes of research information. National agricultural research programs, universities, NGO's, private companies and other users of the information are faced with the task of identifying which parts of this information are relevant to them in order to be able to use the improved germplasm effectively.

To help the CIMMYT staff and its collaborators in their work on maize breeding, and also as a tool to access information generated during the breeding process, a series of applications have been developed as shown in figure 1.

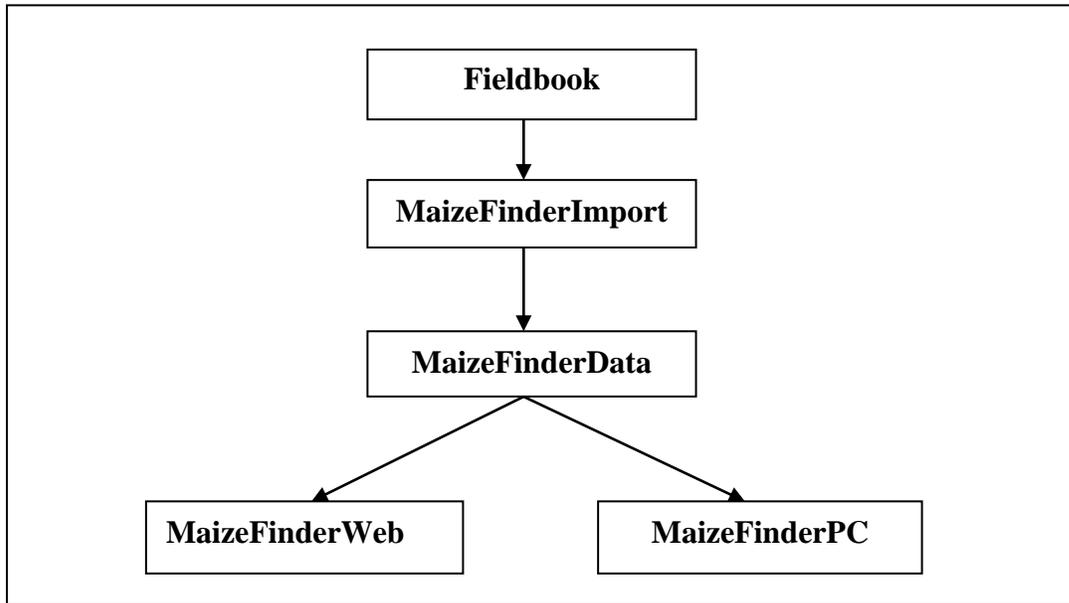


Figure 1: Overview of MaizeFinder

Fieldbook is widely used by the CIMMYT Maize Program and its collaborators in the maize breeding fieldwork and it performs a range of useful functions to the users regarding seed preparation, management of nurseries and trials, inventory, and seed shipment. *Fieldbook* is an application in Excel and this presents a series of limitation for the management of large volumes of related data. To remedy this limitation and provide better access to the data *MaizeFinder* was developed as a database application. In relation to *MaizeFinder* *Fieldbook* serves as the data capture tool. *Fieldbook* is freely available to all users.

MaizeFinderImport is a tool that is used for importing the data captured in *Fieldbook* into the *MaizeFinder* database. It permits the data curator to perform a series of quality checks and cleanup procedures on the data and afterward import the data into the database. *MaizeFinder import* is developed in Access. The document, *Data import to MaizeFinder*, describes *MaizeFinderImport* in detail. *MaizeFinderImport* is used by CIMMYT to import data, which subsequently is accessible in *MaizeFinder*.

If you would like to have your maize data included in MaizeFinder you should contact CIMMYT to obtain information on the requirements for the data. If your data fulfills these criteria they can be included in the next distribution of MaizeFinderData.

MaizeFinderData is an Access database designed to store information related to maize germplasm, and trials and nursery series performed with this germplasm. It can store

germplasm data and performance data on different levels such as plot, master calculation, experiment, and trial series. The database is described in detail in the document Data model for MaizeFinder. The database is distributed jointly with MaizeFinderPC. It is made in Access 2000 and you can use this to query it. To avoid conflicts with the MaizeFinderPC application it is recommended that you should copy MaizeFinderData before you start manipulating it in Access.

MaizeFinderWeb is an internet based GIS and database application that gives the users the opportunity to perform complex queries on the data from MaizeFinderData, and view the results of the queries in a series of reports and also on a map. A prototype of the application can be visited on: <http://192.100.189.157/maizeinfo>.

MaizeFinderPC has functionality similar to MaizeFinderWeb, but is a PC version, which is installed individually on each users PC, without any requirement of access to the Internet. The use of MaizeFinderPC is described in detail in the present documentation.

Head-to-head analysis

Introduction to head-to-head analysis

Plant breeding is essentially a process where superior germplasm is produced and identified. One of the fundamental problems in this process is the interaction between the environment and the genotypes (GE), which makes it difficult to determine which germplasm is superior. One approach to overcome this problem is the use of multi environmental trials (MET), which permits breeders to compare germplasm in different environments. The data from the MET can then be analyzed e.g. by comparing the results of varieties within an experiment or across a trial series.

Head-to-Head analysis of hybrids uses the data from the MET, but instead of confining the analysis to an individual experiment or trial series, data are used from all experiment where the two hybrids has been grown together, regardless of experimental design, number of replications, and other hybrids evaluated (See below for an example of head-to-head analysis). The advantages of head-to-head analysis include (Trifunovic 2002):

Head-to-head of hybrids:

- Possible to compare hybrids across experiments
- Identify advantages and weaknesses of each hybrid
- Use of data for stability analysis

Head-to-head of lines

- Provides specific information on what the line brings into hybrids
- Information on advantages and disadvantages of each line
- Identification of characteristics that needs improvement in line
- Basis for the planning of crosses for pedigree selection (F2 populations)

The analysis is carried out for each trait using a paired t-test, for the comparison. That they are 'paired' means that the observations for a given trait is taken from the first hybrid and paired with the observation in the same experiment of the same trait of the second hybrid. This means that a systematic comparison is undertaken of the two when grown under similar conditions. A paired t-test is used to determine if the difference is significant. Head-to-head analysis of lines is similar to the comparison of the hybrids, but with the added criteria that the two lines must be crossed with the same line within the experiment.

Only two hybrids are considered at a time in order to obtain the maximum number of locations and experiments used in the comparison (Troyer and Rosenbrook, 1983). As not all traits are recorded in all experiments the number of data points for each comparison and trait may vary. The greater the number of data points, the more reliable the results (Trifunovic et. al, 2003).

Details on head-to-head functionality in MaizeFinder

The Head-to-head analysis is carried out in three steps:

1. If you are comparing two hybrids enter the four parent lines, or alternatively if you are comparing two lines enter the pedigrees of each.
2. Select one of the four outputs: GIS, analysis, data to table, or data to CSV file.
3. Run the head-to-head analysis.

Below is a detailed description of each one of these steps.

Step 1: Enter the line names:

- 1.1 If needed click the New button  on the toolbar to clear the parents entered in the previous analysis.



- 1.2 Enter the first letters of Parent 1.1 e.g. CML and click the load pedigrees button  to the right. If no letter is entered all the hybrid parents in MaizeFinderData will be loaded. Wait for the pedigrees to load, and select the desired pedigree from the dropdown list.
- 1.3 Enter the first letters of Parent 1.2 and click the load pedigrees button  to the right. If no letters has been entered all the line parents in MaizeFinderData that has been crossed with Parent 1.1 will be loaded. Entering the first letters will reduce the list to the materials starting with these letters. Wait for the pedigrees to load, and select the desired pedigree from the dropdown list.
- 1.4 Similarly to the previous step enter the Parents 2.1 and 2.2.

The load pedigrees buttons  considers the pedigrees entered above, but not below, which means you should enter the parents starting from the top.

Step 2. Select the output from the head-to-head analysis

Select output

GIS

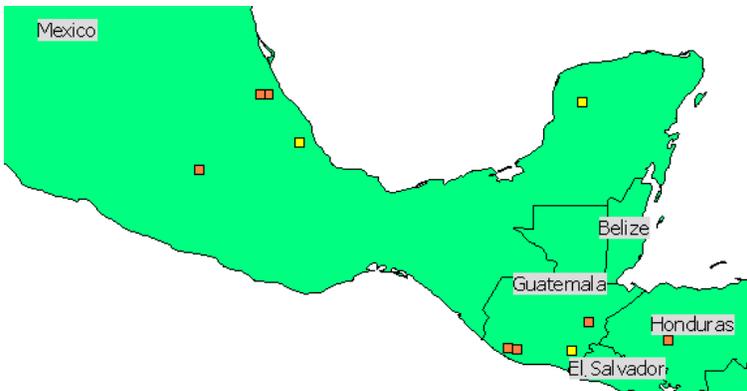
Analysis

Data to Table

Data to CSV

You can select four different outputs from the head-to-head analysis, GIS, Analysis, Data to table, and Data to CSV.

2.1 *GIS* output will on the map highlight in yellow the trial sites from which data is included in the head-to-head analysis i.e. the locations where the experiments were carried out. It should be noted that some traits might not have been scored at all sites.



The figure above illustrates an example of a GIS output, where the information from the trial sites has come from two sites in Mexico and one in Guatemala. These three sites are highlighted in yellow.

2.2 *Analysis* shows a report with the head-to-head analysis of the hybrids or lines:

Head to Head analysis

10/24/2003

Line 1 CML264

Line 2 CL-02305



	Line 1	Line 2	No. Observations	Difference	
Anthesis date	54.73	54.66	6	0.07	n.s.
ASI, days	0.17	0.83	6	-0.67	n.s.
Ears per plant, #	1.01	0.99	6	0.02	n.s.
Ear rot, %	6.97	6.78	6	0.19	n.s.
Exers. Turec., 1-5			0		
GLS, 1-5			0		
Grain moisture, %	18.22	17.42	6	0.80	n.s.
Grain texture, 1-5			0		
Grain yield, kg/ha		6,491.63	6	156.28	n.s.
Bad husk cover	4.03	1.75	6	2.28	*
Leaf rolling, 1-5			0		
Leaf senesc, 0-10			0		
Plant height, cm	240.85	246.77	6	-5.92	*
Pucc. Sorghi, 1-5			0		
Root lodging, %	21.20	38.95	4	-17.76	n.s.
Stem Lodging, %	0.23	0.33	6	-0.10	n.s.

The first section of the report shows the two lines that are being compared or in the case of hybrids the four parents of the two hybrids. The next section has the 16 traits that are being analyzed in the first column. The second and third column indicates the average values of the two lines or hybrids that are being compared. The fourth column indicates the number of observations for each trait. The fifth column indicates the difference between the average values and it also indicates if the difference is significant. N.s. means the difference is not significant, * is significant at 0.05 level, ** at 0.01 level and *** at 0.001 level.

In the example above the line CL-02305 has a significantly better (0.05 level) husk cover than the line CML264. Also a significant difference for plant height was observed.

2.3 *Output to Table* shows the data included in the analysis in a table.

Country	Location	AnthesisDate1	AnthesisDate2	AnthesisDifference
Mexico	COTAXTLA (INIFAP)	55.50	53.00	2.50
Guatemala	CUYUTA	54.48	53.99	0.49
Mexico	UXMAL	55.66	57.52	-1.86
Guatemala	CUYUTA	54.09	54.52	-0.43
Mexico	UXMAL	54.62	55.43	-0.81
Mexico	COTAXTLA (INIFAP)	54.00	53.50	0.50

The example above shows a small section of the table. The column AnthesisDate1 gives the values for the first line or hybrid and AnthesisDate2 gives the values for the second

line or hybrid. AnthesisDifference provides the difference between the two values. The other 15 traits are presented in the same manner. The table also includes additional information such as names of locations and countries where the trials were carried out.

If you click a column heading the table is sorted according to that column. If you click it again the sort order is reversed.

2.4 *Output to CSV* (Comma Separated Value file) exports the same data as is shown in a table to a CSV file.

F	G	H	I	J
Country	Location	AnthesisDate1	AnthesisDate2	AnthesisDifference
Mexico	COTAXTLA (INIFAP)	55.5	53	2.5
Guatemala	CUYUTA	54.4801	53.9867	0.4934
Mexico	UXMAL	55.6616	57.518	-1.8564
Guatemala	CUYUTA	54.0863	54.5199	-0.4336
Mexico	UXMAL	54.623	55.4282	-0.8052
Mexico	COTAXTLA (INIFAP)	54	53.5	0.5

The example above shows the exported data after it has been imported into Excel. To import the file into Excel:

1. In file menu Click Open and in file type select Text file.
2. Select the file you exported and click Open
3. In the window that opens click Finish to import the data.

Though, it is called a *comma* separated value file, MaizeFinder uses the character ‘tab’ to separate the values.

Step 3. Run the head-to-head analysis

To run the analyses click the funnel  in the toolbar or select Run Head-to-Head in the Head-to-Head menu.

Queries in MaizeFinder

Introduction to Queries in MaizeFinder

MaizeFinder permits you to define multiple criteria in order to obtain exactly the information you are looking for. You can freely combine different criteria related to the germplasm, the experiment or trial series, the characteristics of the location or the traits measured.

MaizeFinder uses the criteria you define to build a SQL query string which is then used to extract the data. All the criteria are combined with an AND which means that the result will only include the data that fulfills all the criteria. As the results in the database may not include measurements on all possible traits the answer may not include material, which could fulfill the criteria if the trait had been measured. This is particularly relevant if you include traits that are only measured infrequently. The query result will only include the experiments where the traits were measured, but omit the ones where it was not. If you expect that the tendency showed in the few cases where the trait was measured is valid across all experiments you may want to run additional queries e.g. looking for the complete data on a given germplasm, which has shown the desired trait in the measurements.

A query is a combination of both the criteria and the output, and you can vary both independently i.e. you can define a set of criteria and then run the query with different outputs or you can maintain a fixed output and vary the criteria. Your interpretation on the data may well vary both depending on the criteria and the chosen outputs, and you should use this flexibility to try to answer your question in different ways in order to obtain the best available answer.

The Queries in MaizeFinder provides very easy access to the data, and this will hopefully help you make better decisions. However, you should not base your decision on one single query but rather try to use different queries that may present different aspects of the data. One example is to use the outputs to reports and GIS to get a quick overview of the results, but this should then be complemented for example by head-to-head analysis (if relevant) and extraction of complete data sets on the material and additional analysis e.g. in Excel or SAS.

Specific Query functionality

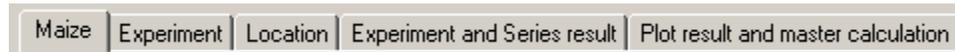
A query in MaizeFinder involves three steps:

1. Enter the criteria to be used for selecting the data
2. Select one of the four outputs: GIS, report, data to table, or data to CSV file.
3. Run the query

Furthermore, you can if desired save the query for future use.

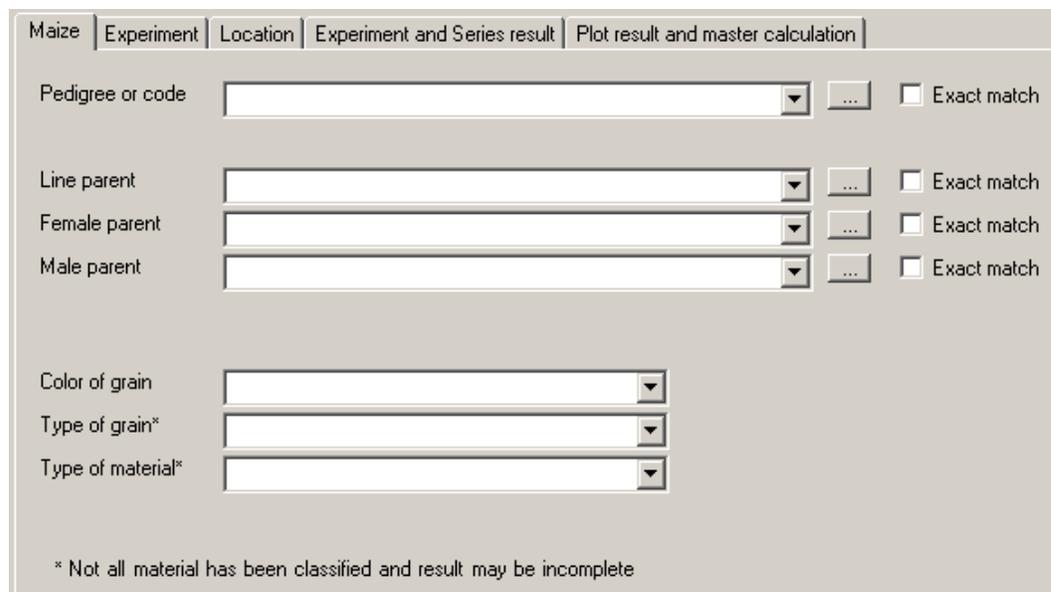
Step 1. Enter the criteria to be used for selecting the data.

The criteria are grouped thematically on five tabs:



- *Maize*, which has criteria related to the germplasm
- *Experiment* has criteria related to the experiment or trial series
- *Location* has geographical criteria related to the location site
- *Experiment and Series result* enables you to include criteria related to the traits measured.
- *Plot result and master calculation* has criteria on the trait in individual plots.

The query will return all the records that fulfill all the criteria you have entered.

A screenshot of a search criteria form. At the top, there are five tabs: 'Maize', 'Experiment', 'Location', 'Experiment and Series result', and 'Plot result and master calculation'. The 'Maize' tab is selected. Below the tabs, there are several input fields and checkboxes. The first field is 'Pedigree or code' with a dropdown menu, a '...' button, and an 'Exact match' checkbox. Below it are three more fields: 'Line parent', 'Female parent', and 'Male parent', each with a dropdown menu, a '...' button, and an 'Exact match' checkbox. At the bottom, there are three more dropdown menus: 'Color of grain', 'Type of grain*', and 'Type of material*'. A footnote at the bottom left reads: '* Not all material has been classified and result may be incomplete'.

The *Maize tab* contains criteria related to the germplasm. The first criterion is related to the pedigree or code of the material itself. The second to fourth criteria are related to the parents of the material, either line parent in the case of lines or female and male parent for hybrids. For all four you have the option to enter one or more characters of the pedigree or code and then use the load pedigree button  to load the matching pedigrees. If you tick the Exact match to the right of the pedigree the search will only return the materials, which are equal to the criterion. If it is not ticked it will return all materials that has the entered material at the beginning of the pedigree. The latter can be used e.g. for searching for several generations of a line by entering the shortest pedigree.

If you do not use exact match it is advisable to carefully check that your query only includes the material desired.

Grain color and type along with type of material can be used to narrow the output e.g to yellow single cross hybrid material. The table below presents all the options used in MaizeFinder for these traits; however, the drop down lists will only include the options, which are currently used in the database. This means that for example Orange may not occur in the drop down menu if no orange maize has been entered.

Grain color	Type of material	Grain type
White	Inbred line	Floury
Yellow	Hybrid (F1) SC hybrid (single cross)	Semi-floury
Purple	Hybrid (F1) TWC hybrid (three-way cross)	Dent
Variiegated	Hybrid (F1) DC hybrid (double cross)	Semi-dent
Brown	Hybrid (F1) Non-conventional hybrids	Semi-flint
Orange	Composite/OPV's	Flint
Mottled		Pop
White cap		Sweet
Red		Opaque
Blue		Tunicate
Black		Waxy

It should be noted that not all material in the database contain information on the type of material and grain type, which means that the answers returned from the query might be incomplete in the sense that they only include the ones which had information.

The screenshot shows the 'Experiment' tab in the MaizeFinder interface. It features several search criteria:

- Breeding program:** A dropdown menu.
- Trial series:** A dropdown menu.
- Year:** A range filter with input boxes for 'Year' and 'and <='.
- Management:** A dropdown menu.
- Collaborator:** A dropdown menu.

The *experiment tab* contains criteria related to experiments and trial series. It is useful for breeding programs or collaborators that are looking for data on their own work.

Management refers to the type of management applied in the experiment. The most common categories include:

Optimal, which are well fertilized rain-fed conditions using optimal site-specific agronomic practices

Low N: Trials grown in fields that have been depleted of nitrogen by growing unfertilized, non-leguminous crops for several seasons and removing the crop biomass after each season. Managed nitrogen stress maize trials are designed so that yields under managed N stress average 20 - 30 % of the yield of a well-fertilized maize crop at a given site.

Drought: Trials grown during a rain-free period, with irrigation applied at the beginning of the season to establish a good plant stand. Afterwards, irrigation is withheld so that the crop suffers drought stress during flowering and grain-filling, resulting in average yields of about 1 - 3 t/ha.

In some cases it may include factors, which are not strictly management but rather conditions particular to the location such as low soil ph.

The dropdown lists includes all the available options currently in the MaizeFinder database.

Maize	Experiment	Location	Experiment and Series result	Plot result and master calculation
Region		<input type="text"/>		
Country		<input type="text"/>		
Location		<input type="text"/>		
Mega environment		<input type="text"/>		
Latitude	>=	<input type="text"/>	and <=	<input type="text"/>
Longitude	>=	<input type="text"/>	and <=	<input type="text"/>
Maximum temperature	>=	<input type="text"/>	and <=	<input type="text"/>
Minimum temperature	>=	<input type="text"/>	and <=	<input type="text"/>
Mean temperature	>=	<input type="text"/>	and <=	<input type="text"/>
Precipitation	>=	<input type="text"/>	and <=	<input type="text"/>
Potential evapotranspiration	>=	<input type="text"/>	and <=	<input type="text"/>
Elevation (m)	>=	<input type="text"/>	and <=	<input type="text"/>
Day length	>=	<input type="text"/>	and <=	<input type="text"/>
Start month	>=	<input type="text"/>	and <=	<input type="text"/>

The *Location tab* includes criteria related to the location where the experiments were carried out. The first three criteria permit you to narrow in on the region, country or specific location. You can also use the criteria on latitude, longitude and elevation to define geographic areas. Latitude and longitude are in decimal degrees and elevation in meters above sea level.

All climatic criteria refer to the five-month optimum period, which is defined as the five consecutive months of maximum precipitation to potential evapotranspiration (P/PE) based on long-term normal monthly climate data. Initial month indicates the first month of the five-month optimum period. Maximum, minimum and mean temperature are in

degrees Celsius. Precipitation and potential evapotranspiration are the total in millimeters for the five-month optimum period.

Maize | Experiment | Location | Experiment and Series result | Plot result and master calculation

Experiment result
 Series result
 Experiment and series result

Trait 1: Bad husk cover <= 2

Trait 2: <=

Trait 3: <=

Trait 4: <=

Trait 5: <=

Trait 6: <=

The *Experiment and Series result tab* allows you to enter criteria related to the different traits measured in the experiments or calculated across the trial series. The default is to search on the experiment results, but alternatively you can search on the trial series results alone or search on both types of results at the same time. Most data in MaizeFinder is at the Experiment level and this is the recommended option.

The first three drop-down lists include only the 20 most common traits, where as the three last drop-down lists includes all traits currently in the MaizeFinder database. You define a criterion by first selecting the trait from a drop-down list. Then you define the operator to less than or equal to (<=) or more than or equal to (>=). And finally you define the value to be compared to.

Many traits are measured on a scale of 1-5 where one is the best and five the worst. Annex one includes a complete list of all traits and their definition.

Maize | Experiment | Location | Experiment and Series result | Plot result and master calculation

Plot Result
 Master calculation
 Plot and master calculation

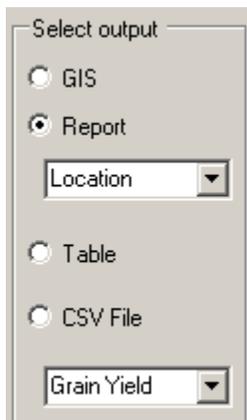
Trait: <=

The plot result and master calculation tab permits you to query data at the plot result and master calculation level. The plot results are the measurements obtained directly in the field without any calculation or adjustment. The master calculations include a series of calculations of the data at the plot level and these are the basis for calculating the experiment results in Fieldbook.

The drop-down list of traits includes all the traits both of plot results and master calculations, although a plot result will not contain the information if it is a calculated value.

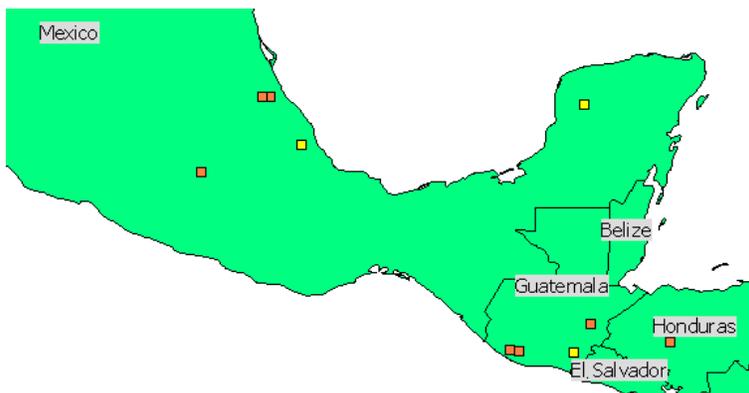
In a plot a trait may be measured repeatedly and the trait may occur with different names in the trait list, i.e. the same trait may have more than one name in the list. This makes it difficult to obtain complete answers from a query. Furthermore, MaizeFinder currently include mainly experiment results and only few of the plot results. In general it is therefore preferable to query the experiment results where the traits are standardized unless you are specifically interested in the individual plot results.

Step 2. Select the output from the query



You can select four different outputs for a query: GIS, Report, Table, and CSV file.

2.1 *GIS* output will on the map highlight in yellow the trial sites from which data is included in the query output i.e. the locations where the experiments were carried out.



The figure above illustrates an example of a GIS output, where the information from the trial sites has come from two sites in Mexico and one in Guatemala. These three sites are colored yellow. For more information see the section on GIS in MaizeFinder.

2.2 *Report* presents a standard selection of fields.

MaizeFinder includes the following reports:

- *Trial Series* gives an overview of the trial series that are included in the query.
- *Location* provides information on the locations in the query.
- *Grain yield* gives average values for grain yield and other common traits. The averages are calculated based on the data included in the query.
- *Pests* provides average values from the query on the most common pests.
- *Disease_1* and *Disease_2* provide average scores on the most common diseases. They have been divided into two reports due to the large number of diseases.

Different tools allow you to navigate the reports e.g. by using a group tree, by searching for specific text strings or simply by moving from page to page. In some cases the reports also allow you to click on a link to obtain more detailed information. Finally, you can print the report or export it as a pdf-file or to Word or Excel. For more information on this see the section on Reports in MaizeFinder.

2.3 *Table* shows the output in a table. You must select the specific output from the dropdown list below the CSV file option.

Standard Pedigree	Anthesis-silking interv ▾	Bad husk cover	Days to silk
CL-SPLW04/CL-04368	8.50		107.16
CL-SPLW04/CL-04368	0.50		92.96
CL-SPLW04/CL-04368	0.10	9.58	54.50
CL-SPLW04/CL-04368	0.09	13.54	60.71

If you click a column heading the table is sorted according to that column. If you click it again the sort order is reversed e.g. from ascending to descending. You can create new outputs or edit the existing ones. For more information on this see the section on Tables and CSV file.

2.4 *CSV file* exports the same data as is shown in a table to a CSV file. You do not necessarily have to load it into the table first, but this is often useful precaution to ensure that the query returned the desired result.

B	C	D	E
Standard Pedigree	Anthesis-silking interval.	Bad husk cover	Days to silk
CL-SPLW04/CL-04368	0	0.893	57.2952
CL-SPLW04/CL-04368	0.1	9.582	54.5
CL-SPLW04/CL-04368	8.5		107.1594204
CL-SPLW04/CL-04368	0.5		92.96094199

The example above shows the exported data after it has been imported into Excel. To import the file into Excel:

1. In file menu Click Open and in file type select Text file.
2. Select the file you exported and click Open

3. In the window that opens click Finish to import the data.

Though, it is called a *comma* separated value file, MaizeFinder uses the character 'tab' to separate the values. For additional information see the section on Tables and CSV files.

Step 3. Run the query

To run the query click on the funnel  in the toolbar or select Run query in the Query menu.

Save, load and delete query criteria

In MaizeFinder a query can be made up of many different criteria, and you may want to run a query several times with minor variations in criteria and output definition. To facilitate this you can store the query criteria.

1. Enter your criteria in the query form.
2. Click the Save Query As button  or select Save Query As in the File menu.
3. Enter the name of the Query in the dialog that appears.
4. Click OK, and your query criteria are now stored.

To load the query again at a latter stage simply select the query in the Open Query dropdown list .

To delete a query from the list first select the query in the Open Query dropdown list  and then push Delete Query .

GIS in MaizeFinder

Background on GIS

GIS means geographical information system, and refers generally to information systems that can handle georeferenced data. Georeferenced means that you either directly know the coordinates that indicates the features location on the earth or indirectly which e.g. is the case when you know the name of a town or municipality which in turn can be used to find the coordinates of the location.

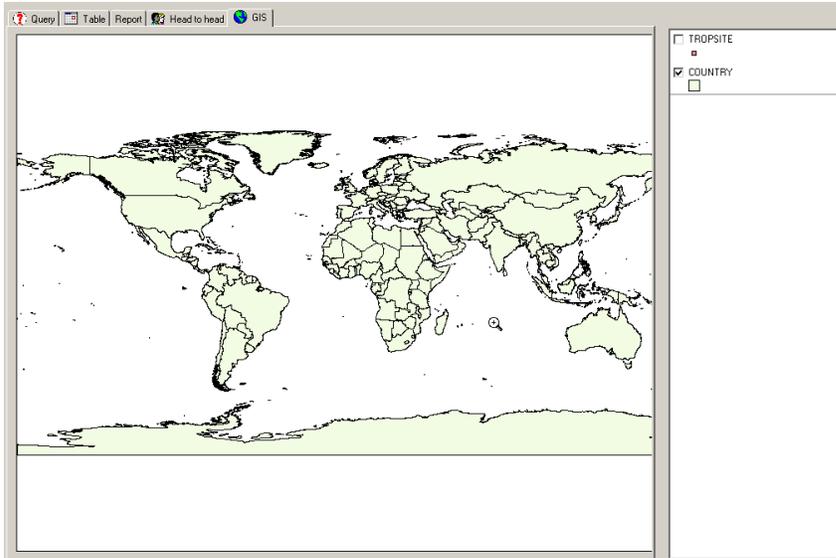
The backbone of MaizeFinder is a relational database called MaizeFinderData, which includes a table with the characteristics of the trial locations including latitude and longitude. This information enables us to show the trial sites on a map and combine the information with other map layers e.g. on maize mega environments. The MaizeFinder database includes information on all sites that recurrently are used for trials. For example with on farm trials we may not always have the location coordinates. The location coordinates are used to obtain climatic data; thus if the coordinates are missing the climatic data will also be missing.

The GIS in MaizeFinder is very simple and is termed a ‘map viewer’, which is a system that mainly can be used for viewing maps. MaizeFinder only includes very few map layers, but you can use African, Asian or Latin American Maize Research Atlases previously published by CIMMYT to obtain many more layers. These also include a GIS viewer, MapExplorer, and many GIS projects, which may help you, render the layers in a useful manner. Additional freely available GIS software is DIVA, which you can download from www.diva-gis.org. We suggest you complement the limited GIS in MaizeFinder with these resources.

The reason why the GIS in MaizeFinder is simple is that we wanted to be able to distribute the system freely to users and allow them to distribute it to other users as needed. Other more sophisticated systems such as ArcGIS give many more analytical capabilities especially in the integration of GIS and databases; however, to access these capabilities the users have to pay a license fee. At a later stage we hope to import MaizeFinder into ArcSDE, a geodatabase system that will enable users of ArcGIS to access all the data in MaizeFinder directly from the GIS-system.

Using GIS in MaizeFinder

The GIS interface is made up of two parts, the map and the legend. A map is a collection of one or more layers that are drawn on top of each other. The layers in the map are listed in the legend, which shows the name of the layer and the symbol used to draw the layer. The check box next to the name of the layer indicates whether the layer is visible or not. If it is checked it is visible i.e. it is currently drawn on the map.



The legend

The order of the map layers is important as the top layers may cover up information on the layers further down. Generally, it is most useful to have point layers such as trial sites on top, followed by line layers such as roads, and the polygons (e.g. countries) at the bottom. The drawing order on the map is also reflected in the legend. For example, the layers at the top are drawn on top of those below it.

The legend can be used to perform four different operations

1. *Toggle visibility of layer:* To toggle the visibility of a layer click the check box to the left of the layer name in the legend. If the box is checked the layer is visible, clicking it will make it invisible. Or conversely if it is not checked the layer will become visible.
2. *Change drawing order of layers:* To change the drawing order of layers on the map, you can drag a layer up or down in the legend. The change in the order on the legend will then be reflected on the map.
3. *Hide/show classification of layer:* To hide or show the classification of a layer, double-click on the name of the layer in the legend. If the classification includes a lot of classes the drawing may be slow and it may be preferable to hide it in order to speed up the process of redrawing the legend or maintaining the other layer in the visible part of the legend.
4. *Select active layer:* You make a layer active by clicking on it in the legend. It will then be underlined to mark it as the active layer. Setting the active layer is necessary in operations that are performed on a specific layer. For example, changing the rendering of the layer or using the information tool to extract information.



The legend above has two layers TROPSITE and COUNTRY. The TROPSITE layer is rendered with an orange square on the map, but as the check box is unchecked it is not drawn on the map. The COUNTRY layer is a polygon layer drawn in a light green color, and it is visible on the map as the check box is checked. The COUNTRY layer is also underlined, which means it is the active layer.

Navigating the map

In order to be able to move around on the map MaizeFinder provide a series of standard navigation tools: . If you have already used a GIS these will be familiar, but if not here is a brief description of each navigation tool:

 *Zoom to full extent* button will zoom to the full extent of all layers on the map. This can be used e.g. if you have zoomed in a couple of times and want to move all the way out again.

 *Zoom to active layer* is similar to Zoom all out, but will only zoom to the extent of the layer marked as active in the legend.

 The *Zoom in* tool is used for zooming to a particular area on the map. After clicking the tool, drag a box on the map and it will zoom to that area.

 The *Zoom out* tool is used for zooming out and centering the map at a new location at the same time. To do this, select the zoom out tool and click on the map. The map will now zoom out and center on the place you clicked.

 The *Pan* tool is used for moving around on the map without zooming in or out. After having selected the tool, keep the left mouse button pressed while dragging the mouse. This will move the map in the direction you drag. Release the mouse button and wait for the map to redraw.

Adding a standard layer to the map

MaizeFinder includes 13 standard map layers:

Trial sites
Country
Administrative boundaries
Cities
Rivers
Lakes
Precipitation
Potential evapotranspiration
Maximum temperature
Minimum temperature
Elevation
Soils
Maize mega environments

The first 6 layers are global whereas the last 7 are regional. The reason for making them regional is the size of the files, which for global layers would be too big for most computers. The four climatic layers cover the five-month optimum period.

To add a standard layer simply select it on the Map layers menu, and the layer is automatically added to the map and rendered in a default manner. Once a standard layer is loaded it is like any other layer, and you can for example change the rendering or move it up or down in the legend.

If the regional maps are not from your region of interest you may have to go to Options in the Windows menu, and change the selection of regional layers. You simply click on the load button  next to the regional layer you want to change, and in the dialog that opens navigate to the equivalent layer in the folder with map layers from the region of interest. See the description for the Setup of MaizeFinder for a more detailed description.

Adding and removing layers

In MaizeFinder GIS you can add your own layers to the map. Although, many different formats for GIS files exist, MaizeFinder GIS is limited to shape files. These are the standard files used in ArcView, and the most common GIS file type and most GIS systems can both read and write this file type.

To add a layer:

1. Click the Add layer button  and the add layer dialog opens.
2. Navigate to the directory that contains the layer you wish to add.
3. Click on the layer you wish to add and click open. The layer is now added as the top layer on the map.

To remove a layer:

1. Click on the layer you wish to remove in the legend in order to make it the active layer.
2. Click on the Remove layer button  on the toolbar to remove the layer.

This will remove the layer permanently. If you wish to use it later again it may be better to use the check box of the layer to make it invisible.

Symbolizing data

Maps are valuable because they can communicate very complex information much more effectively than say tables or lists. The reason is that maps take advantage of our natural abilities to distinguish and interpret colors, patterns, and spatial relationships. When you display your data on a map you may identify distributions, relationships and trends you could not previously see. You can also use maps to effectively communicate information and results to others (ESRI 1998).

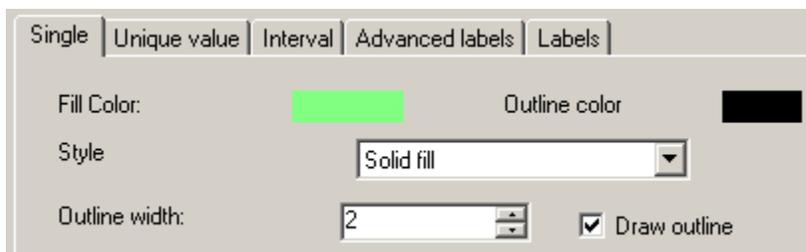
You can symbolize your data in five different ways. The first three ways Single, Unique value and Interval changes how the data is actually drawn where as the two last ways assigns a label to each feature e.g. the names of all the countries. The specific options for symbolizing the data depends on the type of data i.e. if it is points, lines or polygons.



Symbolizing your data has a major impact on how your data is perceived, and you should experiment with different ways of symbolizing a particular set of data in order to identify how it can be best presented. Symbolizing involves deciding on color schemes and symbols to represent features. It may also involve grouping or classifying the features according to their characteristics (ESRI 1998).

Single symbol

Single symbol renders all the symbols in a layer with the same symbol. It can be combined with the labeling options in the way that you first assign the single symbol and then the labels.



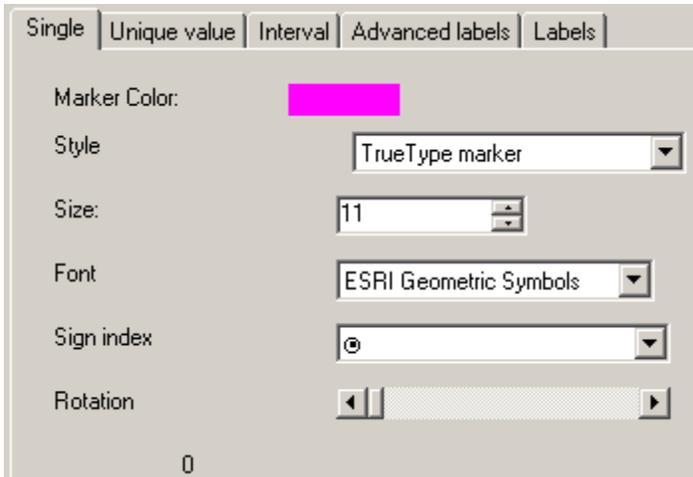
The figure above illustrates the options for a single symbol for a polygon layer. If you click on either the Fill color or the Outline color, the standard windows color dialog opens and you can choose the color you want.



You can also choose between different fill patterns as shown in the figure below.



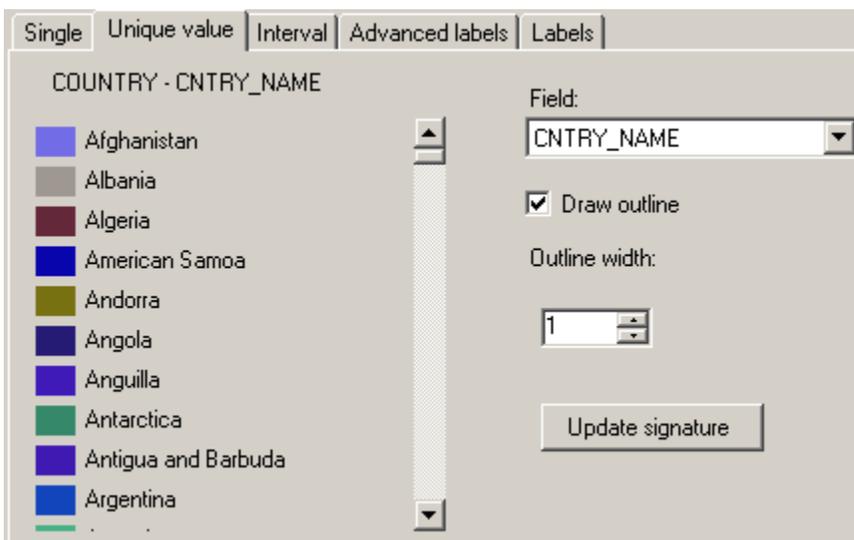
Finally, you can indicate the width of the outline, or choose not to draw the outline i.e. it will have the same type as the fill.



For a point layer you can also choose the color of the symbol by clicking on it. Furthermore, you can choose the style or marker type to be used and the size of it. The most common options are square, circle, triangle, or cross markers, but you can use any true type font if desired. If the marker is a True type font then the Font drop down will have a complete list of all the markers available from that font. For the true type fonts you also have the option to rotate the chosen symbol.

Unique value

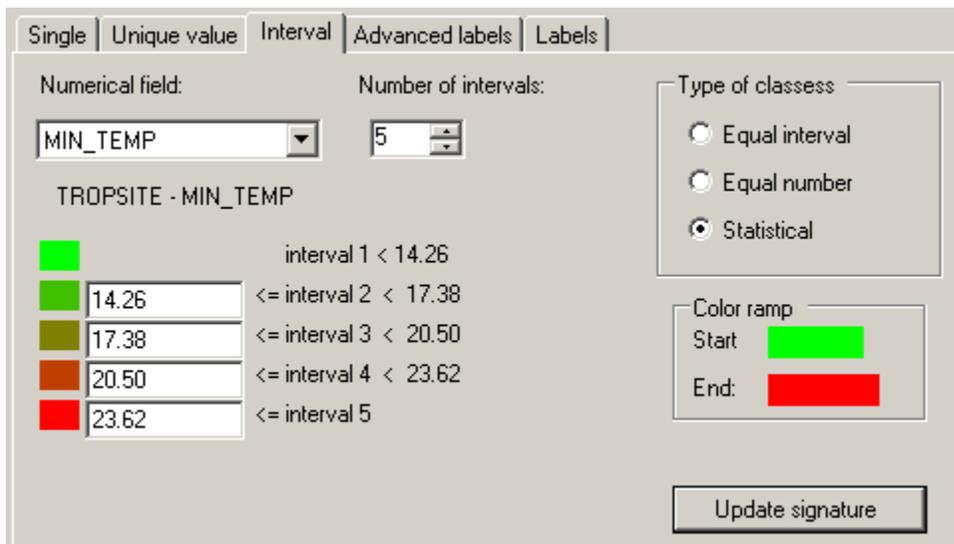
The unique value renderer assigns a specific symbol to each unique value in a specific field. To apply a unique value renderer you must first select the field in the dropdown list. Then if desired select if the outline should be drawn and the width of this. Then you must press the button Update signature. This generates a list of unique values and assigns a random color to each one. If you do not like the colors assigned you can press Update signature, and it will generate a new random color scheme. Alternatively, you can click on one of the individual colors in the list and the select color dialog opens; from this you can freely choose a color.



If a layer with a unique value renderer has many different values such as a list of all the countries in the world, the drawing of the legend will be slow and it will be very difficult to pick out a specific color on a list of more than 100. In that case it might be preferable to double-click on the layer in the legend so that the symbols are not redrawn every time the map is refreshed.

Interval Renderer

Interval or class break renderer applies symbols to a set of discrete values. It only works on numerical fields where it is possible to define a range of values.



To assign an interval renderer you must first generate a draft renderer, which you can then customize further.

To generate the draft renderer:

1. Select a numerical field and the number of intervals desired.
2. Select the type of classes i.e. the system to apply in the classification
 - *Equal interval* creates ranges of equal size.
 - *Equal number* attempts to create intervals with the same number of observations in each interval. This may not work well if you have many repeated values, which makes it difficult to create groups of equal size.
 - *Statistical* uses the standard deviation and the mean of the values to create the ranges. The ranges are defined as the mean plus or minus one, two, or three times the standard deviation. A range is only created if it contains values.
3. If desired click on the colors in the color ramp to define the start and end colors of the ramp.

4. Click on Update signature to create the draft interval renderer.

Customize the draft renderer

- Click on the individual colors to change them such as in order to make a specific range stand out.
- Change the values that define the limits of the ranges such as to use round numbers. When changing ranges make sure that the values are increasing from top to bottom and that they are non-overlapping.

Labels

MaizeFinder has two options for adding text to the map: Labels and Advanced labels. Both take information from a field related to a map layer and add it to the map. Labels is the simplest to use and is recommended for most tasks. Advanced labels provides more options, but has the limitation that it will allow different labels to overlap, which makes it difficult or impossible to read them. Advanced labels are mainly useful if you are zoomed to far in that the overlapping labels are not a problem. The Labels option only adds the labels that will fit on the map: As you zoom in it dynamically adds more labels as space become available. This is generally preferable and you should use the Labels option for most tasks.

Tip: You can first define a single renderer and then add labels, which will maintain the labels. If you want to use a unique value renderer or an interval renderer you cannot add labels to the layer without losing the rendering. To solve this problem you can add the layer twice. On the bottom copy apply the unique value or interval renderer and on the top layer apply the labels, but make sure to deselect the option Draw object, which means only the labels are drawn.

Single Unique value Interval Advanced labels Labels

Text field
CENTRY_NAME

Font Tahoma f

Mask color

Label placement

- Above - Right
- Above - Center
- On top of object

Draw object
 Mask labels

To add labels:

- Select the field that contains the text to be added.
- If desired change the font by clicking on the f –button and use the standard change font dialog.
- If you leave Draw objects checked, the map layer will be drawn; if it is unchecked only the labels are drawn.
- If you check Mask labels a back color is added to the label, which may make it easier to read. By clicking on the colored square you can change the color of the mask.
- For point layers you can choose to place the label above to the right, centered or on top of the object.

Advanced Labels

The Advanced labels option provides very detailed control over the placement of the individual labels, but it has the draw back of overlapping the labels if space is limited. For most purposes, the Labels option is recommended.

To add advanced labels:

- Select the field that contains the text to be added.
- If desired change the font by clicking on the New font button and use the standard change font dialog.
- If you leave Draw objects checked, the map layer will be drawn and if it is unchecked only the labels are drawn.

- The permit repetitions can be used to eliminate labels that repeat the same information.
- The horizontal and vertical options can be used to place the label in relation to the object.
- The rotation can be used to put the labels at an angle e.g. setting it to 90 will draw the labels vertically.
- Y-Offset and X-Offset can be used to control each label individually by having two columns in the shape file that defines the vertical and horizontal offset.

Reports in MaizeFinder

Background on reports

A report provides a standard output that is always available to you. Different tools allow you to navigate the reports e.g. by using a group tree, by searching for specific text strings or simply by moving from page to page. In some cases the reports also allow you to click on a link to obtain more detailed information. Finally, you can print the report or export it as a pdf-file or to Word or Excel.

Nevertheless, reports have three limitations.

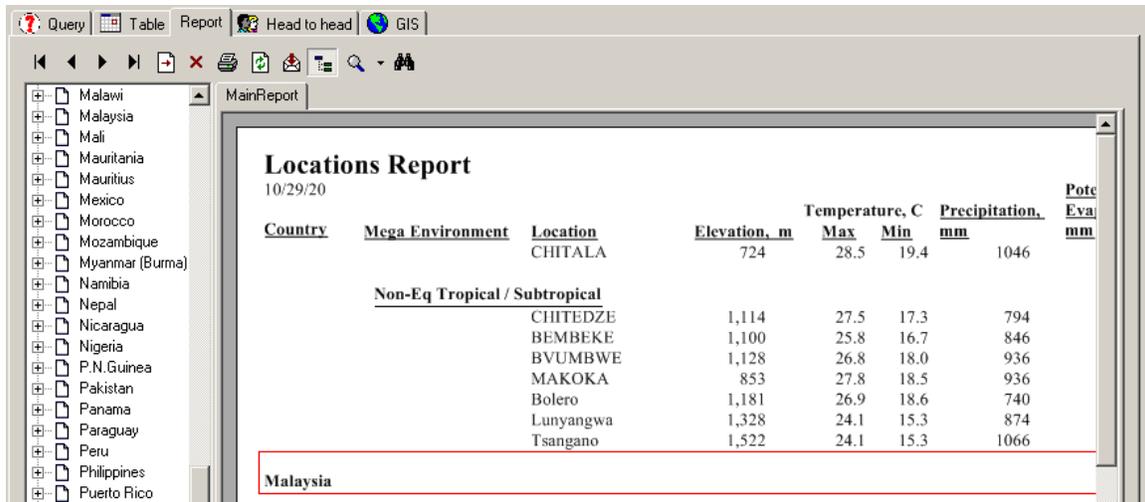
1. The layout is limited to the size of paper, which limits the amount of information.
2. The information included in the report is fixed and you cannot add or remove fields.
3. The information is sorted according to default criteria and you cannot change that.

An alternative output is the table where you are not limited by paper size and you can choose which fields to include, and sort them dynamically according to different fields. See below for more details on tables.

MaizeFinder includes the following reports

- *Trial Series* gives an overview of the trial series that are included in the query.
- *Location* provides information on the locations in the query
- *Grain yield* gives average values for grain yield and other common traits. The averages are calculated based on the data included in the query.
- *Pests* provides average values from the query on the most common pests.
- *Disease_1* and *Disease_2* provide average scores on the most common diseases. They have been divided into two reports due to the large number of diseases.

Specific report functionality



The figure above illustrates the main elements in the report page. In the upper left corner you have a toolbar, which is described in detail below. To the left is a group tree, and to the right a page of the report itself.

The group tree.

The group tree can be used to get an overview of the data available, but also for navigation of the report. In the example Malaysia was clicked in the group tree and the report automatically scrolls to the Malaysia group and highlights it with a red square. By clicking the Toggle group tree button in the toolbar you can hide or display the group tree.

The Report toolbar



The report toolbar includes the following tools

◀ Go to first page

◀ Go to previous page

▶ Go to next page

▶ Go to last page

➦ Go to page, which enables you to go to a specific page number

✖ Close current view, which is used to close views that open with additional details.

🖨 Print report opens the standard print dialog for printing the report or part of it.

-  Refresh is used to reload the data from the database, but this is rarely relevant in MaizeFinder.
-  Export Report opens the standard save file dialog. You can choose to save the report in pdf, Excel, Word or rich text format.
-  Toggle group tree hides or shows the group tree.
-  Zoom allows you to zoom to a certain percentage page size or select Page width, which will fit the report to the size of the report window.
-  Search text allows you to search for a specific text string that you enter in the dialog that opens when you click the button.

Tables and CSV files

Introduction to the data table and CSV files

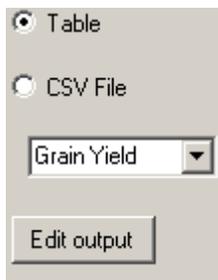
One of the fundamental ideas in MaizeFinder is to provide easy access to the data you need, and that you should be able to select exactly the data you are interested in. The selection of data is achieved by entering different criteria in a query. If you want to be able to select the fields included in the output you must choose table or CSV file as the output.

If you choose table the result of the query is loaded into a table where you can look at the data and sort it according to the different fields. If you choose CSV the data is saved to a CSV file, which can subsequently be imported into other applications such as Excel or SAS.

CSV means Comma Separated Value file and is one of the most commonly used standards for data exchange among disparate applications. Recently XML has entered as a new data exchange standard. MaizeFinder actually uses XML as the internal data standard. But XML is too new to be widely used and the query result is therefore converted to CSV when it is exported. An additional benefit is that CSV files generally are smaller than XML files.

Specific table and CSV functionality

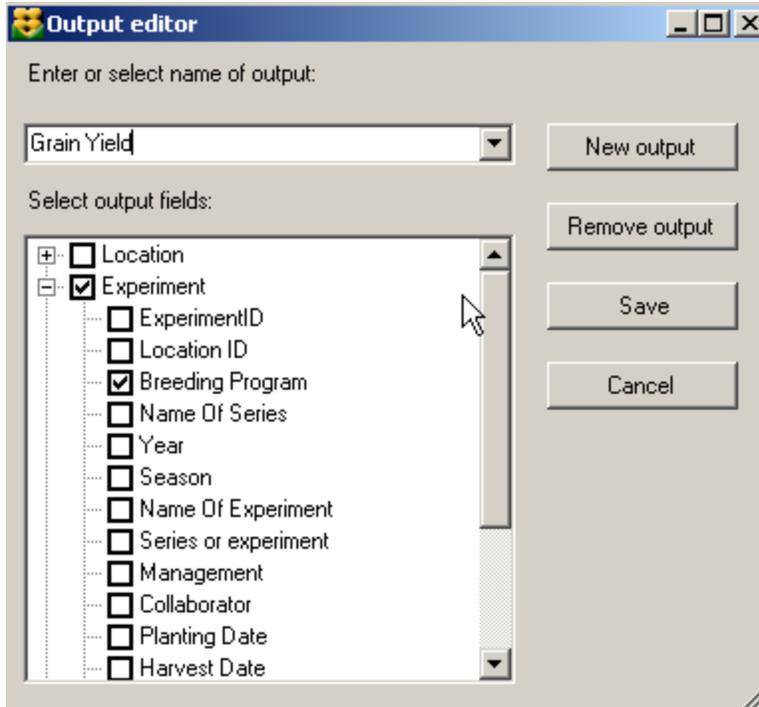
How to select the output for tables and CSV files?



The image shows a user interface for selecting output format. It features two radio buttons: 'Table' (selected) and 'CSV File'. Below the radio buttons is a dropdown menu with 'Grain Yield' selected. At the bottom is an 'Edit output' button.

In the Output panel for the Queries select output as table or CSV file, and then select the desired output from the dropdown list. The choice in the dropdown list applies both to table and CSV file output. You can edit or create new outputs as described in the next item.

How to define the fields in the output for tables and CSV files?



1. Click the Edit output button in the Output panel to the right.
2. If you want to define a new output click New Output and enter the name for a new output. If you want to edit an existing output select it from the dropdown list and the output will then be loaded into the node tree.
3. To open a node click on the + to the right of it. You can now include fields by marking the checkboxes or alternatively remove fields by removing the checkmark.
4. If you tick the check box at a node this will include all fields of that node in the output, and if you uncheck it all the fields are removed from the output. When you load an output and the node is checked this only indicates that some but not necessarily all fields in a node are included in the output.
5. Once you are satisfied with the content of the output, click Save and this will close the output editor. The output will now appear in the dropdown list with outputs.

Functionality of the table

Standard Pedigree	Anthesis-silking interv ▾	Bad husk cover	Days to silk
CL-SPLW04/CL-04368	8.50		107.16
CL-SPLW04/CL-04368	0.50		92.96
CL-SPLW04/CL-04368	0.10	9.58	54.50
CL-SPLW04/CL-04368	0.09	13.54	60.71

A MaizeFinder table is very simple and principally serves to allow one to take a quick look at the data. Functionality is limited to being able to sort the data according to a column. Clicking in the column heading does this. If it is clicked again the sort order is reversed. In the example above the data has been sorted according to Anthesis-silking interval. The white arrow pointing downwards in the heading indicates that the sort order is descending.

The width of the columns in the table can be changed by placing the mouse pointer between two column and waiting for it to change to a double arrow.

 The mouse can then be dragged in order to change the column width.

Export to CSV file and import into Excel.

If you choose output as CSV file the same data as is shown in a table is exported to a CSV file. A dialog will open requesting you to enter a name for the file.

B	C	D	E
Standard Pedigree	Anthesis-silking interval.	Bad husk cover	Days to silk
CL-SPLW04/CL-04368	0	0.893	57.2952
CL-SPLW04/CL-04368	0.1	9.582	54.5
CL-SPLW04/CL-04368	8.5		107.1594204
CL-SPLW04/CL-04368	0.5		92.96094199

The example above shows the exported data after it has been imported into Excel. To import the file in Excel:

1. In file menu Click Open and in file type select Text file.
2. Select the file you exported and click Open
3. In the window that opens click Finish to import the data.

Though, it is called a *comma* separated value file, MaizeFinder uses the character ‘tab’ to separate the values.