SMIS End User Manual

Running SMIS to inform soil management guidance

Date: November 2018

Table of contents

List of figures

List of tables

1. Introduction

The purpose of this document is to aid the end user in using SMIS and interpreting the outputs from SMIS. Ultimately, the outputs from SMIS aim to give better insights and guidance on soil management practices in horticulture.

There are near infinite combinations of queries that can be asked of SMIS, given the number of crops / crop varieties / previous crop / soil types / year combinations held in the SMIS database. This Manual includes some illustrative case studies (example scenarios) of how the 3 components of SMIS (Grower database, Rule Bases and Established Queries) can be run both individually and / or in combination.

It should be noted that not all queries will generate output (in the form of pie charts, histograms, links to articles etc.). This is because at present there is not sufficient data in the system to generate the statistical relationships needed to generate these displays for all horticultural crops listed (or their combinations). As more data is uploaded into SMIS, more scenarios that generate statistically significant relationships can be displayed. Also, the performance of individual crop varieties identified by SMIS should not be taken as any kind of official endorsement or embargo.

Since this End User Manual was written, new data has been added to SMIS. As a result, some of the graphics and results of the scenario testing examples below may be slightly different from the current version of SMIS. However, this updating of the database has not affected the principles or functionality of the system, or the end user experience.

1.1. Navigating SMIS

SMIS can be accessed by AHDB staff at the following address: smis.ahdbdigital.org.uk. The home page is shown in Figure 1.

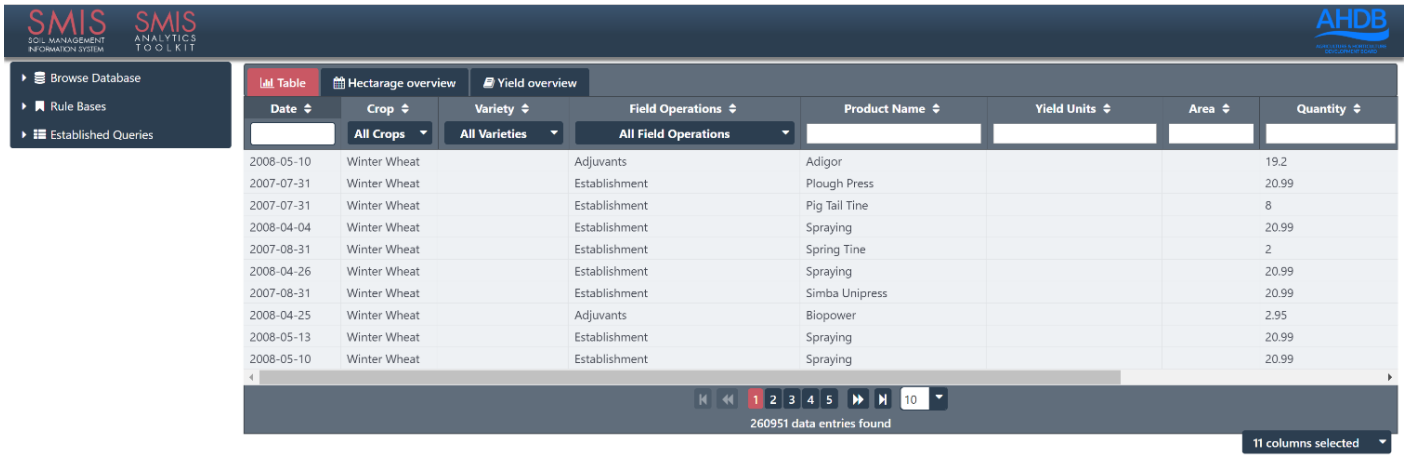

Figure 1. Opening SMIS home page

→ Clicking on the ‘Browse Database’ tab on the left hand side reveals a drop down menu of ‘Grower data’, ‘Experimental data’ and ‘Literature data’.

→ Clicking on ‘Rule Bases’ tab reveals ‘Browse Rule Bases’

→ Clicking on ‘Established Queries’ reveals a drop down menu of ‘Factors affecting Yield’, ‘Compaction

- Risk’, ‘Foot rot index’, ‘PCN level’ and ‘Cavity spot’.

The data in the spreadsheet shown on the SMIS home page is the grower data (see 2.1 below for a more detailed explanation).

2. Browse database

2.1. Grower data

The Grower database contains over 325,000 data items related to soil management in horticulture. The data originates primarily from Gatekeeper data (https://farmplan.co.uk/support/gatekeeper/) and farm records supplied by participating farmers and growers. It contains information on tillage, harvesting, pesticide applications (to crop and soil), fertiliser applications, soil analyses, yield data, etc. The data held can be found by clicking the ‘All field Operations’ tab on the Grower data window (Figure 2).

→ Clicking on the ‘Browse Database’ tab on the left hand side reveals the ‘Grower data’ tab.

The rows in the spreadsheet represent individual data entries from the grower database (Figure 2). The user can expand or decrease the number of rows shown (10, 20 or 30) using the drop down menu at the bottom of the spreadsheet. Some crop/soil/management combinations may have less than 10 entries; others may have considerably more. The user can scroll through the database using the arrow buttons, also at the bottom of the spreadsheet.

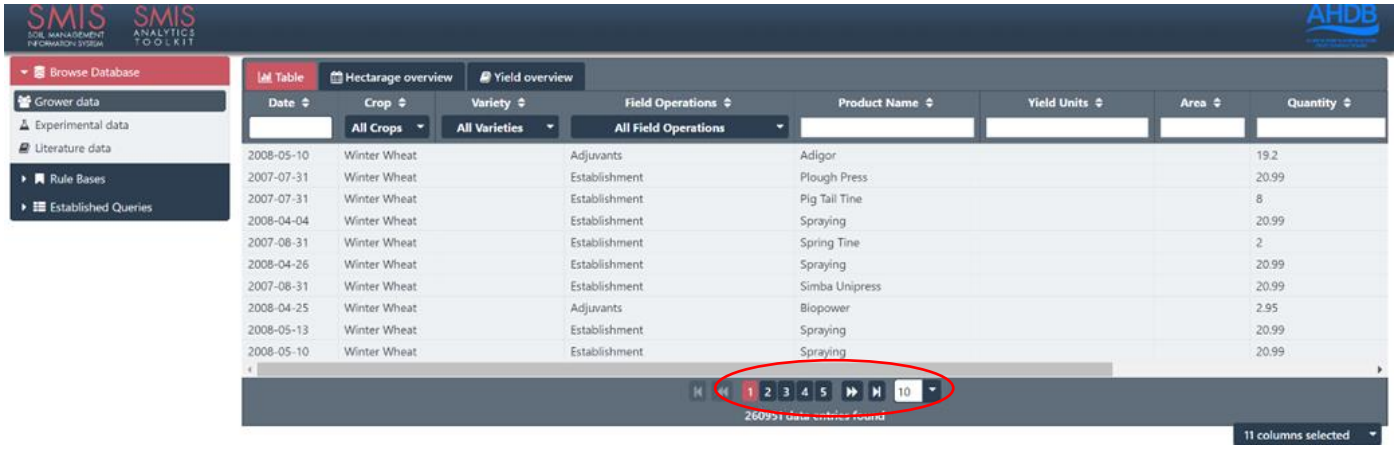

Figure 2. Grower data spreadsheet

The spreadsheet columns represent factors pertinent to soil management that are captured in the Gatekeeper records, the primary source of grower data in SMIS (
Table 1). The categories of data held can be found by clicking the ‘All Field Operations’ tab e.g. Establishment, Fungicides, Herbicides, Lime, Harvest, Organic Manure, Trace Elements, Fertiliser etc.

(Table 1). The user can then select a category of interest e.g. ‘Establishment’ by making a selection using the drop down list (Figure 3).

The user can select the number of columns (factors) of interest that are then shown in the spreadsheet, by using the tab at the bottom right of the spreadsheet (Figure 4).

Factors can be typed in manually via the search bar or by ticking the box next to the factor of interest. The tab then displays the numbers of columns selected (e.g. 11 columns in Figure 2 and Figure 4).

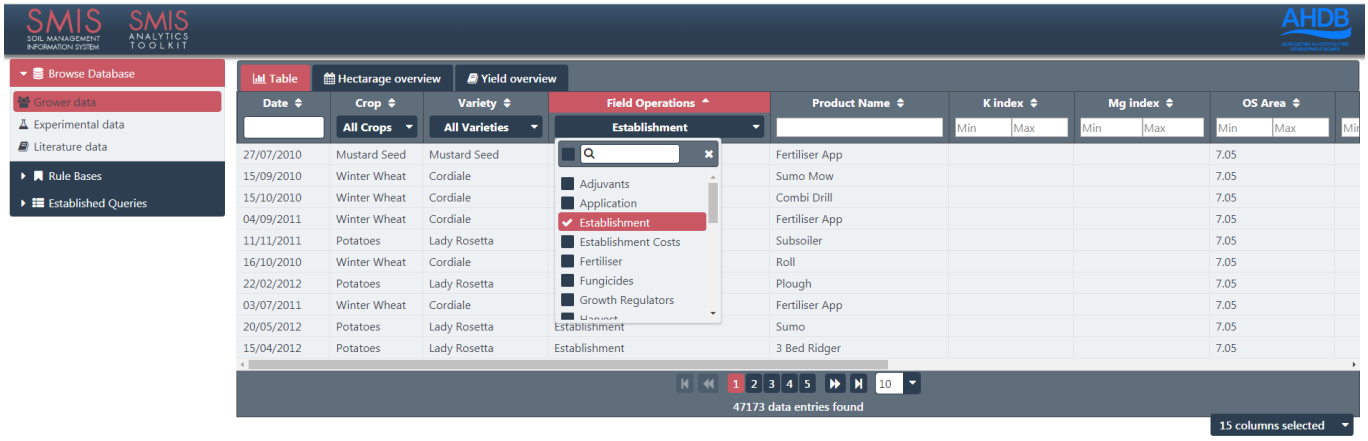

Figure 3. Selecting a category of interest e.g. ‘Establishment’ by making a selection using the drop down list

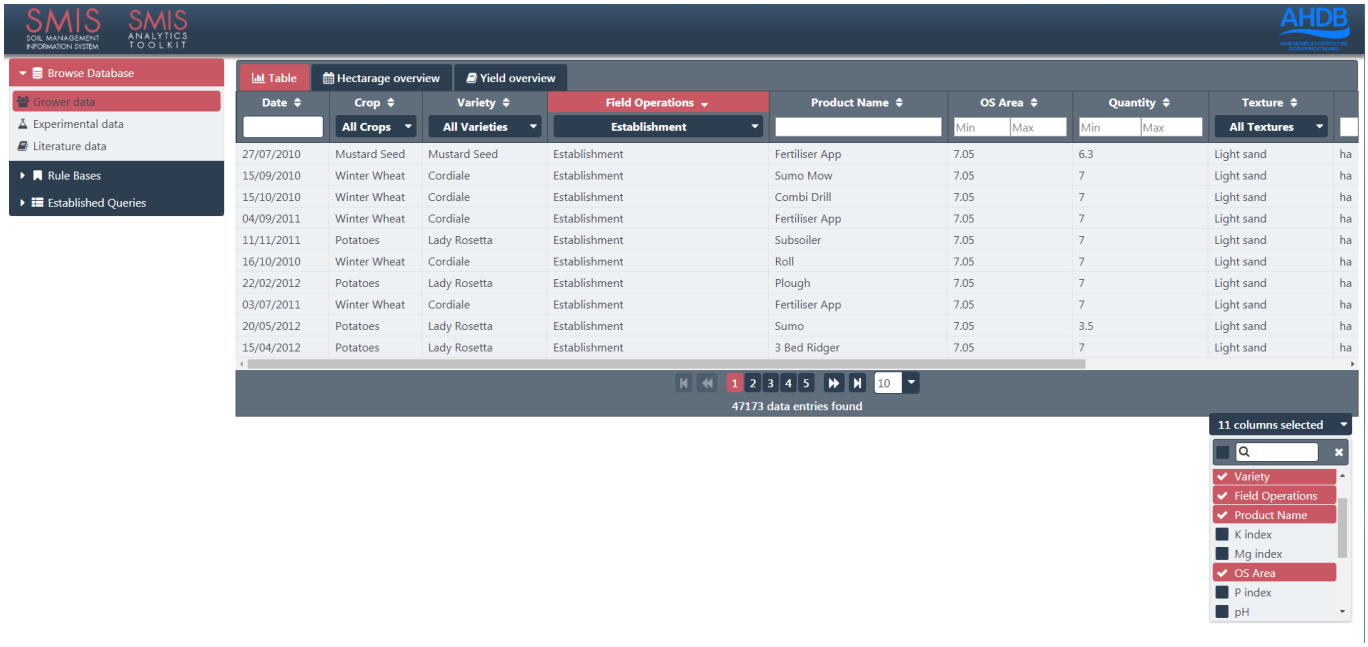

Figure 4. Selecting the number of columns of interest

Table 1. Factors relating to soil management practices

Date	Variety (Crop specific with >200 listed)	Crops >70 (See Table 2)
Field Operations including Adjuvants Applications Establishment Fertiliser Fungicides Growth regulators Harvest Herbicides Insecticides Lime	K index	Product Name*
Mg index	P index	OS Area [Field area (ha) derived from Ordnance Survey and extracted from Gatekeeper
pH	SubfieldId	Quantity
Texture [Deep Clay, Deep Silt, Light Sand, Medium].	Yield	Units
Yield Uni		

N.B. K, Mg and P Indices obtained from Gatekeeper farm data following RB209. pH derived from Gatekeeper farm data .

**Product name refers to 'Product' as listed in Gatekeeper. This includes tillage, spray and harvesting machinery and specific pesticides, adjuvants, fertilisers.*

The end user can then run queries on the database, as shown in the examples below. These scenarios have been created to illustrate the functionality of SMIS: they cannot cover every scenario held within SMIS, so have been selected at random and were presented at the Stakeholder Workshops in June and July 2018.

The user can also select the dates of interest over which the queries are to be run, by using the calendar in column 1 (Figure 5). This allows the user to investigate the impact of 'extreme' weather (e.g. the wet summer of 2010) on soil management decisions and crop yields. The default is to include all years.

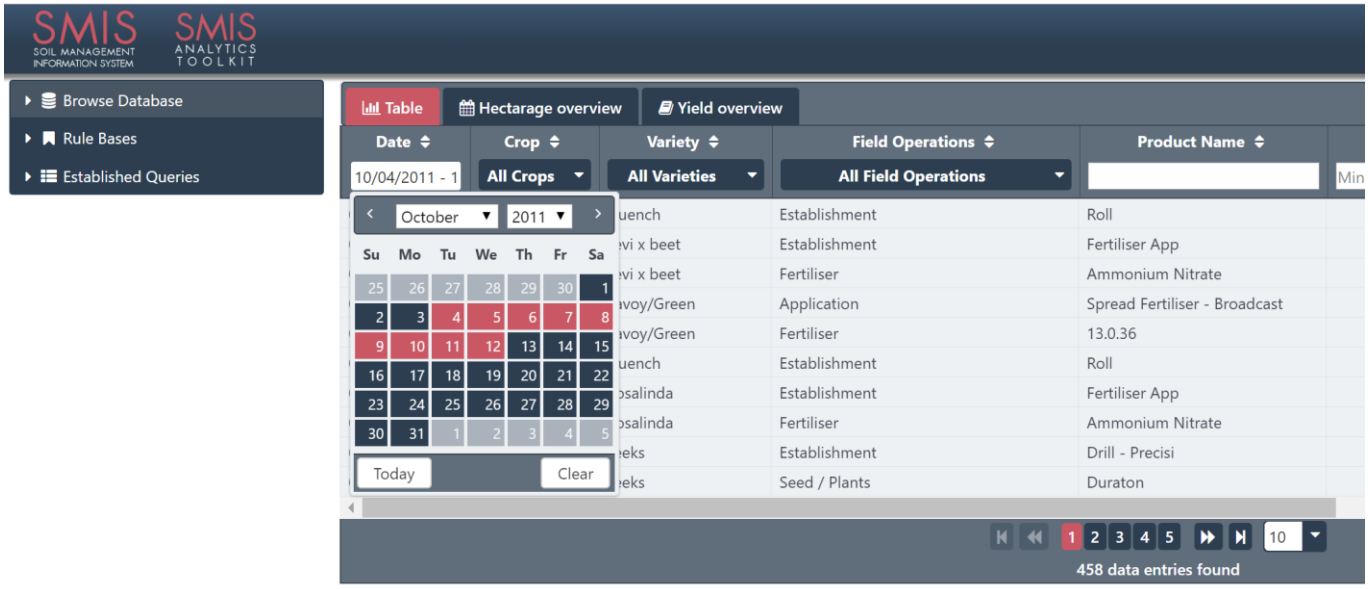

Figure 5. Selecting the dates of interest to run queries.

The user can find how many entries are given for each crop in the SMIS database. The 2nd column of the spreadsheet has a drop down menu of all the crops in SMIS. Many are horticultural crops, but there are also arable, grassland crops and other land uses – reflecting the unique cross rotational character of the SMIS database. This also allows the effect of the 'Previous Crop' (and associated field operations) to be included as a factor influencing horticultural crop yields, as well as soil management decisions. Crops can also be typed in manually via the search bar or by ticking the box next to the crop(s) of interest. Single or multiple crops can be selected.

When new queries are to be run, it is important to 'untick' the crop used in the previous query (unless it is to be included as well as the 'new' crop).

Table 2. Crops and land covers included in the SMIS grower database.

Horticultural Crops			Other non-horticultural crops and land uses	
All Carrots	All Vining Peas	Artichokes Jeru	Spring Wheat	All Potatoes
Asparagus	Beans French	Savoy/Green	Winter OSR	Beans Dried Spring
Broccoli	Carrots	Bulbs	Winter Wheat	Fallow
Celeriac	Collards	Chard	Winter Barley	Grass Ley
Daffodils	Fennel	Drilled Onions	All Cereals	Spring Barley Malting
Kale	Green Cover Crop	Beetroot	Salad Potatoes	Sugar Beet
Cabbage	Round Carrots	Kohlrabi	Potatoes	Millet
Chicory	Organic Carrots	Spring Green	Non Production	Winter Rye
Peas Dried	White Cabbage	Curly Kale	Seed Potatoes	Mustard Seed
Red Pointed Cabbage	Round Head Cabbage	Rhubarb	Grass	Forestry
Salads	Winter Savoy / Green Cabbage	Savoy Cabbage	Maize Forage	Woodland
Spinach	Spring Beans	Winter Greens	Spring Barley	Trees
Haricot Beans	Tender Stem	Squash	Barley Winter Malting	Set Aside

Vining Peas	Leeks	Vining Peas (petits pois)	Game Cover	
Parsnips	Onions	Cauliflowers		
Red Cabbage	Pumpkins			

Having selected the crop(s) of interest, a number of queries can be run.

For example:

- Number of data entries for that crop (displayed at the bottom of the spreadsheet)

- Area grown of the chosen crop(s) for each year of entry in the database (clicking on the 'Hectarage overview' tab above the spreadsheet reveals the area (ha) grown; see Figure 6 and Figure 7). Where varieties of the crop are known (e.g. Celeriac), then these are displayed for each year. A particular crop or variety can be excluded from the display by clicking on it in the legend (on right hand side of the graph (Figure 7). This causes that line(s) of data to be removed from the display.

- Yields for each crop for each year (by clicking on the 'Yield overview' tab at the top of the spreadsheet). Where varieties of the crop are known (e.g. different varieties of Celeriac), then these are displayed for each year (Figure 8).

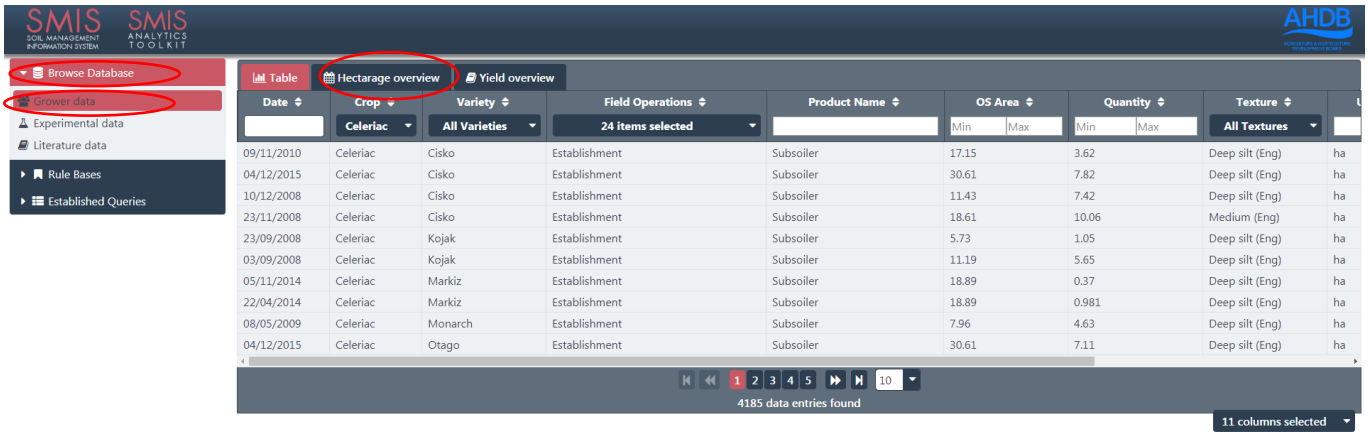

Figure 6. Selecting 'Hectarage overview' to discover area of crop(s) grown

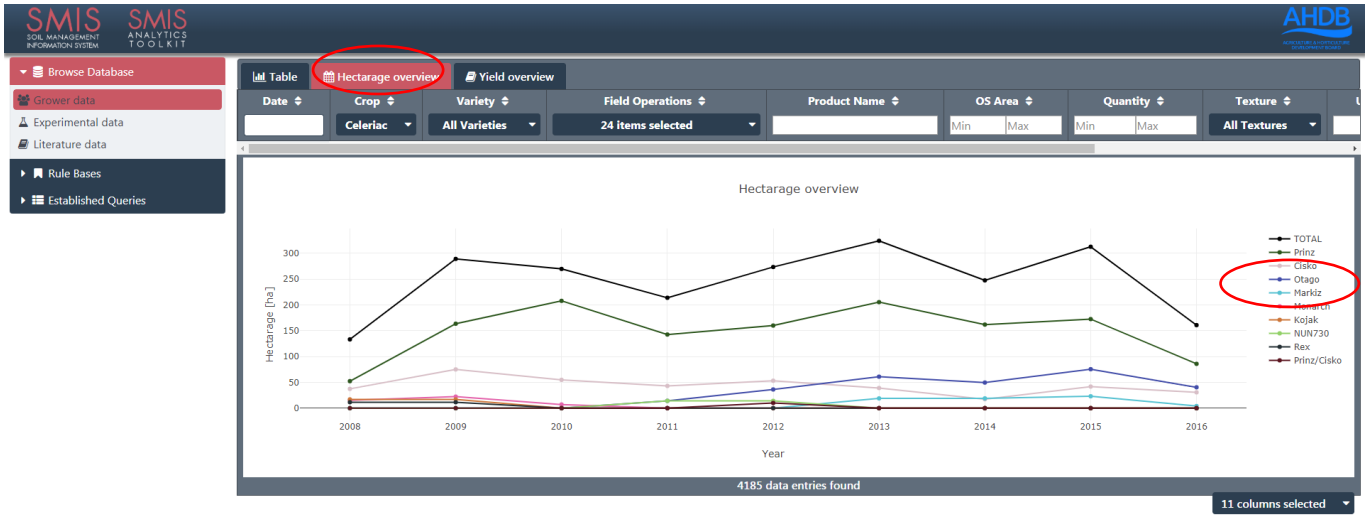

Figure 7. Graphical representation on area of crop(s) by year.

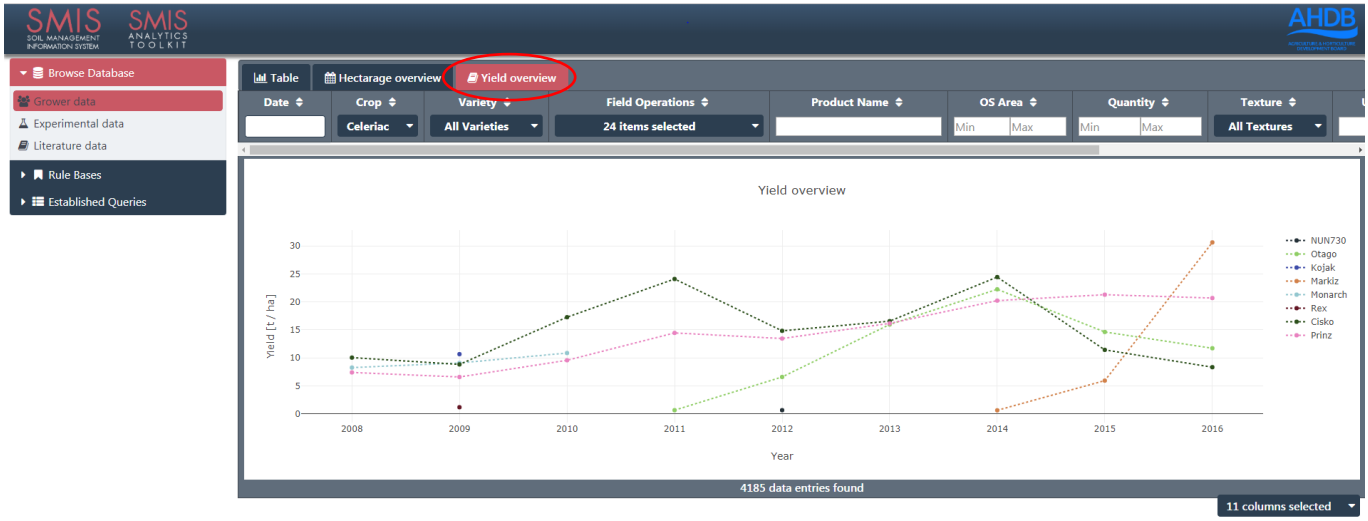

Figure 8. Graphical representation of yield by crop(s) by selecting 'Yield overview' tab

All outputs can be filtered by selecting options in the drop down boxes on column headers allowing for selection of:

- specific crop type(s) (Figure 9),
- varieties of the selected crop(s) (Figure 10),
- field operations (Figure 11) (e.g. fertiliser, fungicides, growth regulators, herbicides, insecticides, lime, molluscicides, organic manure, seed / plants, seed dressings, trace elements); and
- soil texture (deep clay, deep silt, light sand and medium) (Figure 12).

Displays are only possible where sufficient data is available within the SMIS database. The expectation is that more data will be imported in the future, so increasing the number of scenarios that can be displayed.

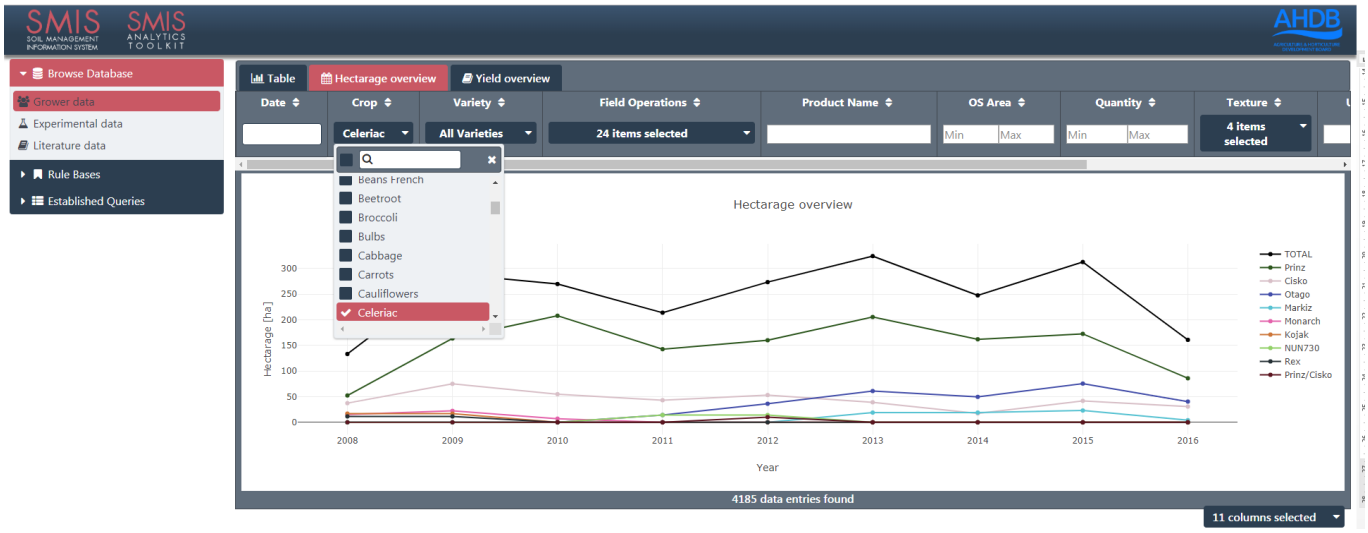

Figure 9. Output filtering using Crop Type.

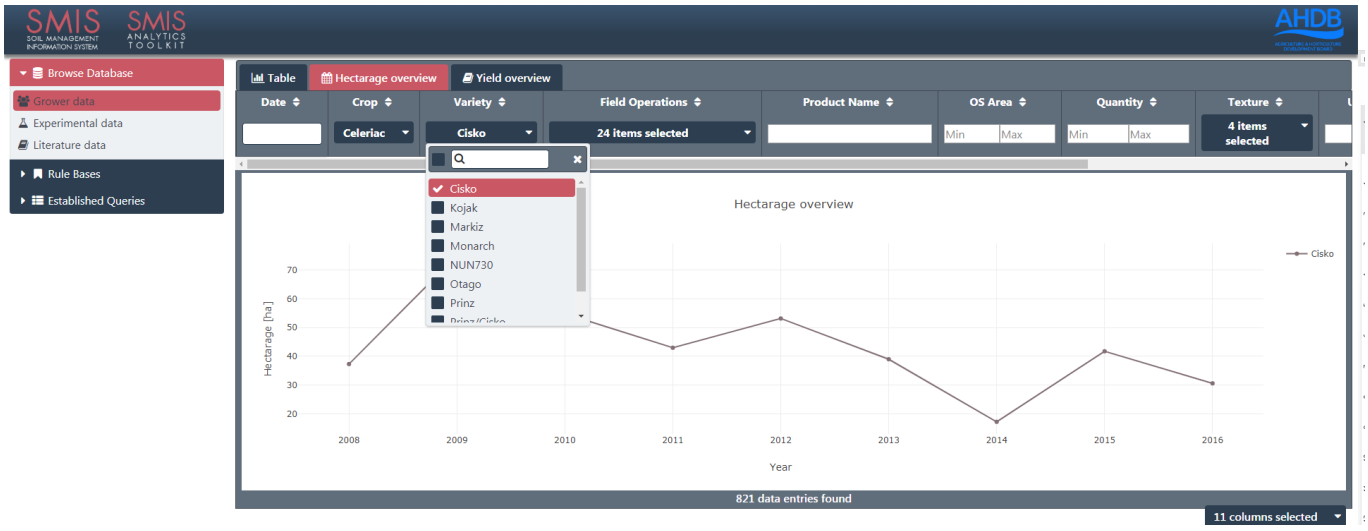

Figure 10. Output filtering using Variety

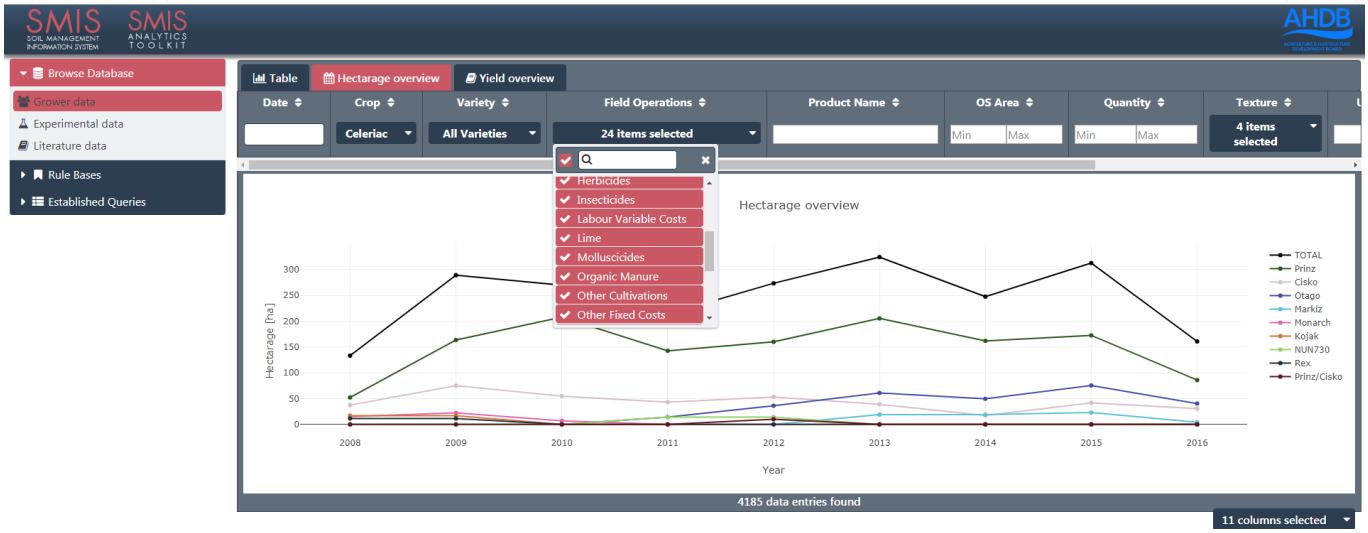

Figure 11. Output filtering using Field Operations

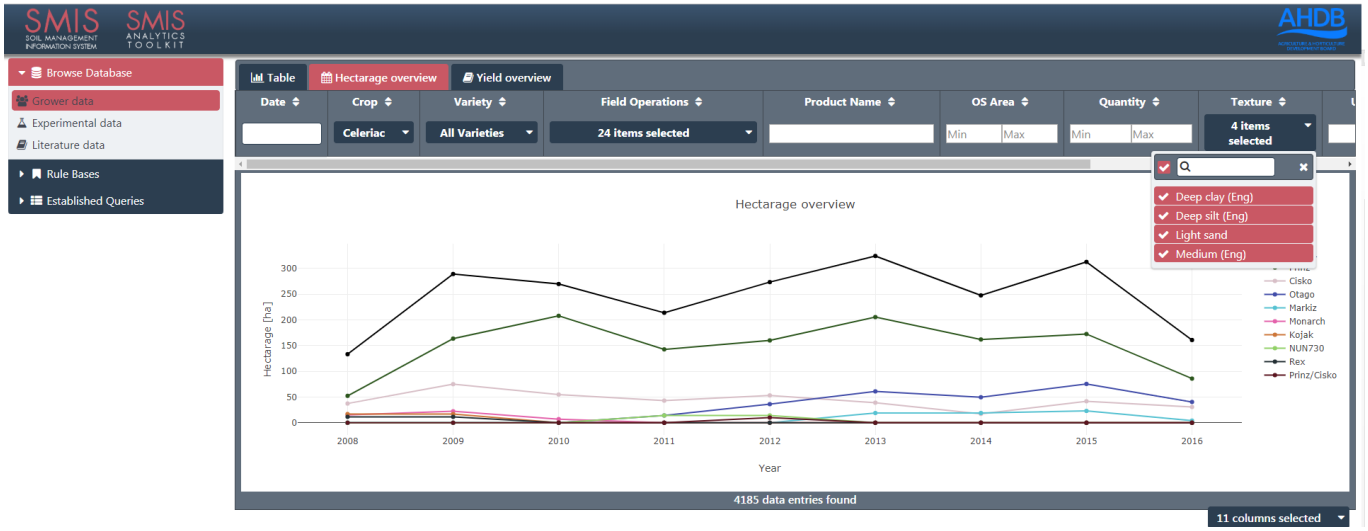

Figure 12. Output filtering using Soil Texture

When new queries are to be run, it is important to ‘untick’ the option used in the previous query (unless it is to be included in the ‘new’ query).

Table 3 and Table 4 show the results of some example queries for selected crops.

Table 3. Results of example queries for bulbs and celeriac

Query run in SMIS Grower database		Selected crops (as examples)	
		Bulbs	Celeriac
Number of data entries?		6820	8459
Hectarage Overview	Year of maximum hectarage?	2014	2013

	Dominant variety in 2015?	Tamsyn	Prinz
Yield Overview	Highest yielding variety?	California	Cisko
	Least consistent yielding variety?	Delian	Markiz

Table 4. Results of example queries for vining peas and sugar beet

Query run in SMIS Grower database	Selected crops (as examples)	
	Vining peas	Sugar beet
Soils with highest hectarage in 2015?	• None on light soil • 80 on medium • 118 on deep clay • 280 ha on deep silt	• Light 250 • 255 medium • 478 on deep clay • 197 on deep silt
Highest yielding varieties in 2015?	Trophee	Springbok

Figure 13 shows how the user can select particular field operations. Table 5 gives examples of outputs for selected crops, potatoes (all varieties) and winter wheat (Duxford variety only). The last query “yields in 2011 when growth regulators were used” produces no displays as there is currently insufficient data in SMIS to generate the plots.

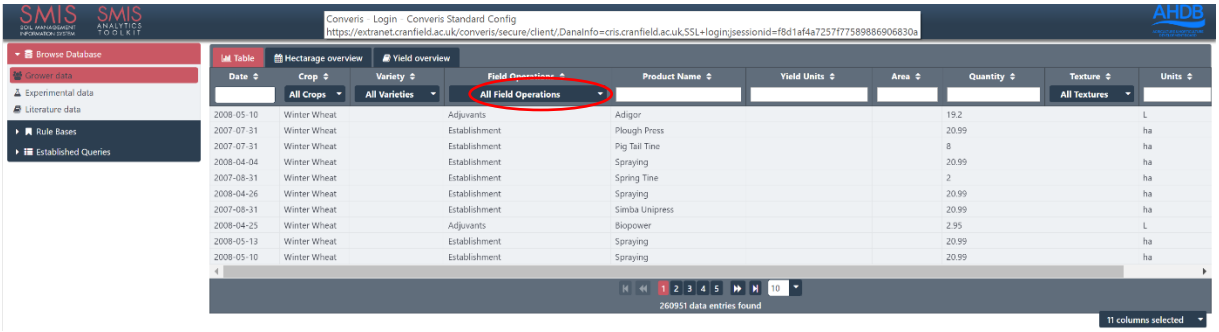

Figure 13. Selection of field operations in SMIS Grower database

Table 5. Results of different field operations

Query run in SMIS Grower database	Selected crops (as examples)	
	Potatoes (all varieties)	Winter wheat (Duxford variety)

Hectarage with organic manures in 2012	557 ha	31.6 ha
Yields where herbicides have been used in 2013	Add the yields of all varieties under 'Yield overview' tab) e.g. Maincrop = 42.97 t/ha	9.32 t/ha
Yields in 2011 when growth regulators have been used	Yield plot could not be generated for the selected dataset.	Insufficient data – no display

2.2. Experimental database

The SMIS experimental database is accessed by clicking on the 'Browse database' and Experimental data' tabs on the SMIS home page (Figure 1) and is shown in Figure 14. The SMIS experimental database (supplied by ADAS, AHDB Project CP107C). There are currently 369 items of field experimental data in SMIS.

The rows in the spreadsheet represent individual sites / fields / locations where soil structural assessment measurements under CP107C have taken place. The columns represent different soil properties and site conditions, including Date, Crop, Soil texture, Soil erosion, Soil colour, Soil porosity, Moisture condition and Clod development. The data can be sorted according to these columns / properties. In the Experimental data, it is possible to run queries on the soil properties associated with a particular crop e.g. 'what is the mean level of Extractable P (mg/l) in cabbages on light soils' etc.

Experimental data can be connected to the grower database and literature database in the visualisation suite of the Rule Bases (see below). However, due to current limited availability of experimental data within SMIS, the linkages are not currently shown visually. This is because only 2 sites that also appear in the Grower database have provided experimental data, which is too limited for any statistical relationships to be run.

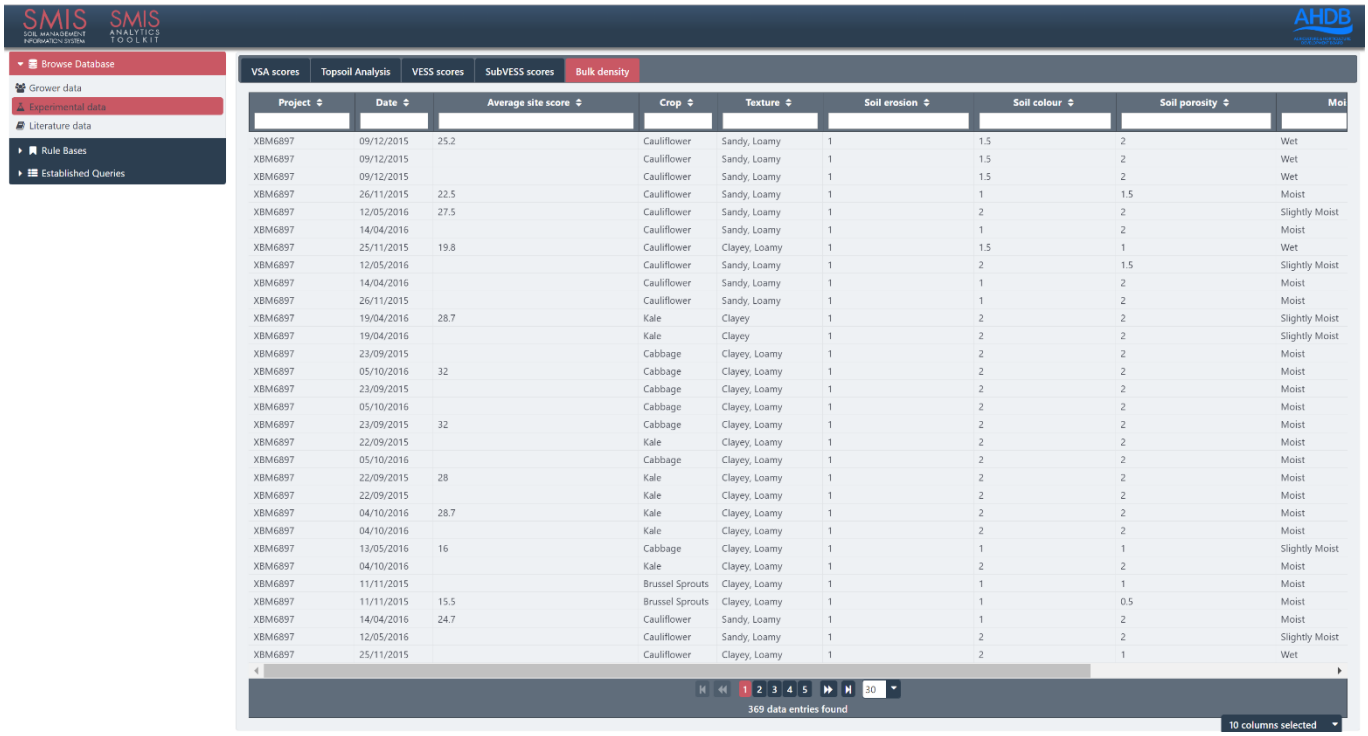

Figure 14. The SMIS experimental database (supplied by ADAS, AHDB Project CP107C)

2.3. Literature database

This database represents an unprecedented repository of literature related to soil management issues and solutions specific to horticultural crops. The database of compiled literature can be accessed by clicking on the 'Browse Database' tab so the 'Literature data' tab is revealed (Figure 15).

Figure 15. The home page of the literature database in SMIS.

The rows in this spreadsheet refer to each individual item of literature. This includes academic papers published in scientific, peer reviewed journals; Conference proceedings / papers; Research reports; and Grey literature (e.g. articles on websites and in trade magazines). The user can select whether 10, 20 or 30 rows are shown on the page using the drop down menu at the bottom of the spreadsheet (Figure 16).

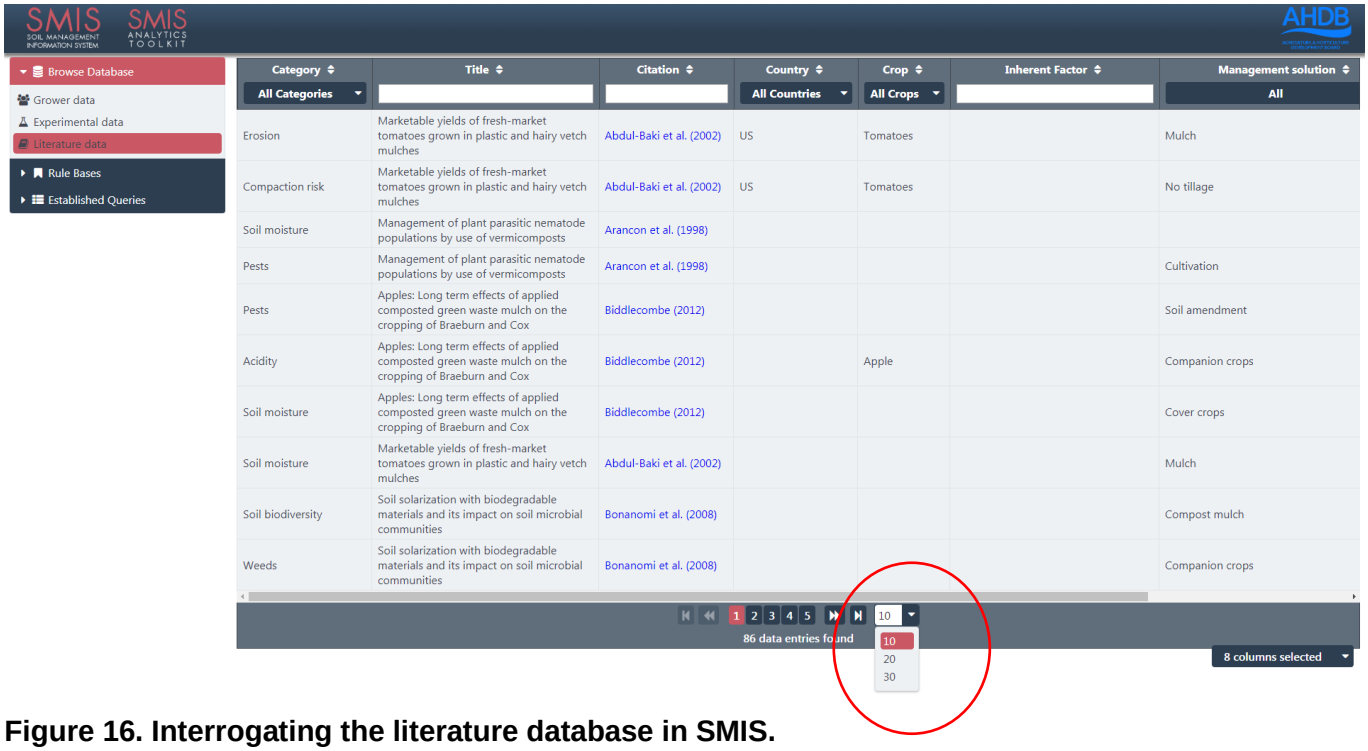

Figure 16. Interrogating the literature database in SMIS.

The columns categorise the items according to the headers listed in Table 6. Clicking on the column header sorts the entries into alphabetical order. For some headers, the user can filter the results to only show selected cases (e.g. under 'Category', the user can select items that only deal with 'erosion' or acidity' for example). Other options are given in Table 6.

Table 6. Columns in the literature database

Column header	Description
Category	The soil management issue addressed in the item. Categories are: Acidity; Compaction risk; Erosion; Nutrient supply; Pests; Soil biodiversity; Soil moisture; and Weeds. The issue(s) of interest can be typed in manually via the search bar or by ticking the box next to the issue of interest.
Citation	Authorship (in alphabetical order) and date of publication. Typing a name in the search box will identify items authored by that name.
Title	Typing in the search box identifies literature item titles with that key word.
Country	The country/countries of interest can be typed in manually via the search bar or by ticking the box next to the country/countries of interest.
Crop	The crop(s) of interest can be typed in manually via the search bar or by ticking the box next to the crop of interest.
Inherent Factor	This category includes factors such as 'periods of bare soil', 'competition' and 'monoculture crop'.
Management solution	The soil management solution(s) of interest can be typed in manually via the search bar or by ticking the box next to the management solution of interest. These include:

Additional N to grass amendment	Biocidal green manure	Biofumigant crops;
Companion crops	Compost	Compost mulch
Conservation tillage	Cover crops	Cultivars

	Cultivation	Fallow	Green manures
	Harvest	Irrigation	Management practice
	Manure	Mulch	Mulch/residues
	Mulch/tillage	Mulches/biosolids	N application rate
	No tillage	Organic	Organic amendments
	Organic mulch	Plant residue mulches	Residue
	Residue retention	Ridging	Rotation
	Soil amendment	Soil disinfestation	Soil solarisation
	Surface mulch	Tillage	
Reference type	Each item of literature is classified by knowledge type; 'quantitative' (based on empirical evidence from field work: laboratory studies were excluded due to the limitations of extrapolating practical, applied results from small spatial scales); 'qualitative' (based on observations during a field-based experiment); and 'anecdotal' (unreferenced statements). This classification was used to evaluate and quantify the confidence in outputs / findings from each item (i.e. the 'weight of evidence' within the SMIS database).		
Note	Concise notes about the item. Typing in the search box identifies notes that contain that key word.		

→ Clicking on the author name(s) in the Citation column takes the user to the original article (Figure 17). This might be to a webpage of an abstract (e.g. a link to ScienceDirect or journal publisher) or a pdf file of the complete article (if in the public domain). Not all sources are available due to copyright restrictions.

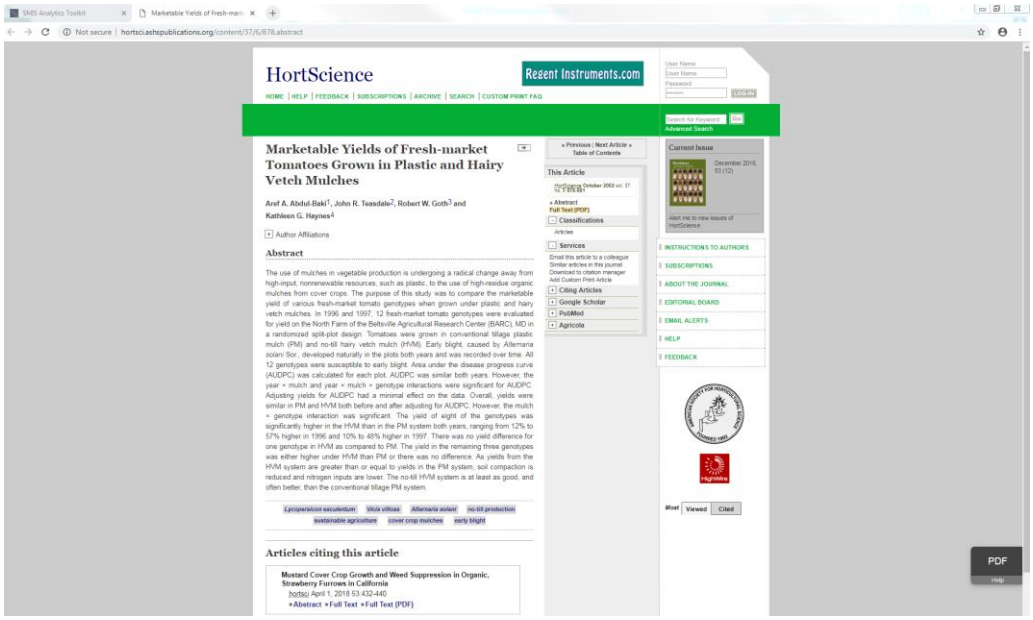

Figure 17. Example of the link to an original article

Table 7 shows some outputs from example queries run on the SMIS literature repository (as run in the Stakeholder Workshops in June and July 2018).

Table 7. Example queries using the literature database

Query raised	Result

How many items relate to soil management issues in asparagus?	2 (Compaction risk and Erosion)
What are the key findings of the papers dealing with soil erosion issues in apples?	Reported in a project report to AHDB by Edwards-Jones (2010). Clicking on the citation brings up the following AHDB report as a pdf from the following website: (http://www.hdc.org.uk/sites/default/files/research_papers/CP%20062_Report_Annual_%202010.pdf) **Project title:** Carbon storage in orchards and amenity plantings **Project number:** CP 62 **Project leader:** Professor Gareth Edwards-Jones, Bangor University **Report:** July 2010 **Previous report:** July 2009 **Key staff:** Professor Douglas Godbold, Professor Davey Jones **Location of project:** Bangor University **Industry Representative:** N/A **Date project commenced:** 29 September 2008 **Date project completed (or expected completion date):** December 2011
What is the most common method for managing weeds and how many items quantify this?	Mulching, with 4 papers that are quantitative: Neilson et al. (2003); Szewczuk and Gudarowska (2004) Stirling (2008); and Szewczuk and Gudarowska (2006)

3. Rule Bases

The 'Rule Bases of SMIS' is a visualisation suite that aims to link the three data components of SMIS: the grower data, the literature and the experimental data.

→ Clicking the Rule Bases tab on the left hand side of the SMIS home page reveals the 'Browse Rule Bases' tab (Figure 18).

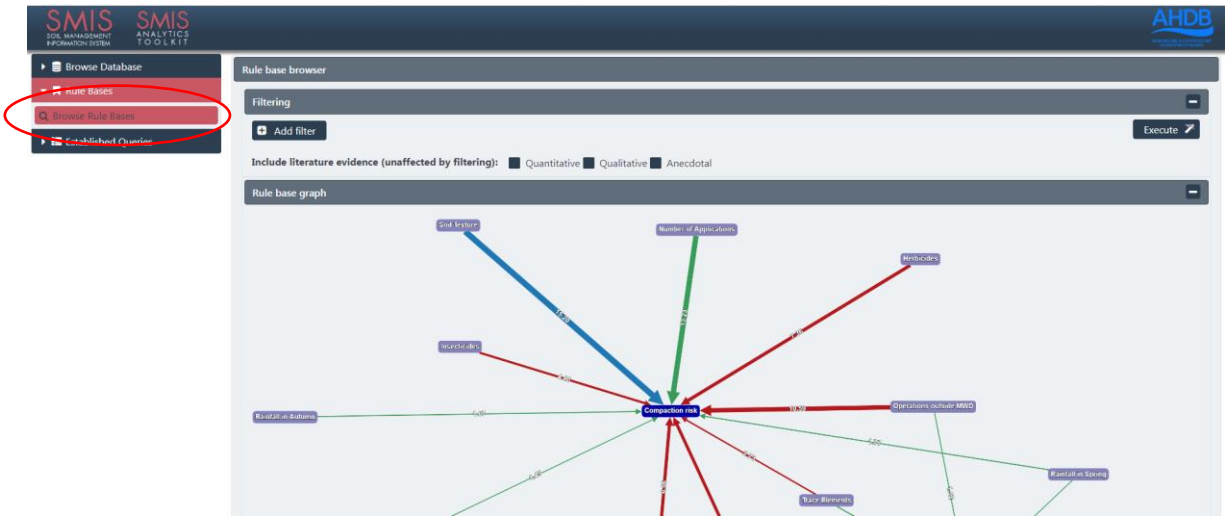

Figure 18. Browse Rule Bases tab

Clicking on the 'Add filter' box (Figure 19) allows the end user to select one or more of the following options (filters):

- Crop;
- Previous crop;
- Variety;
- Soil texture and/or
- Year.

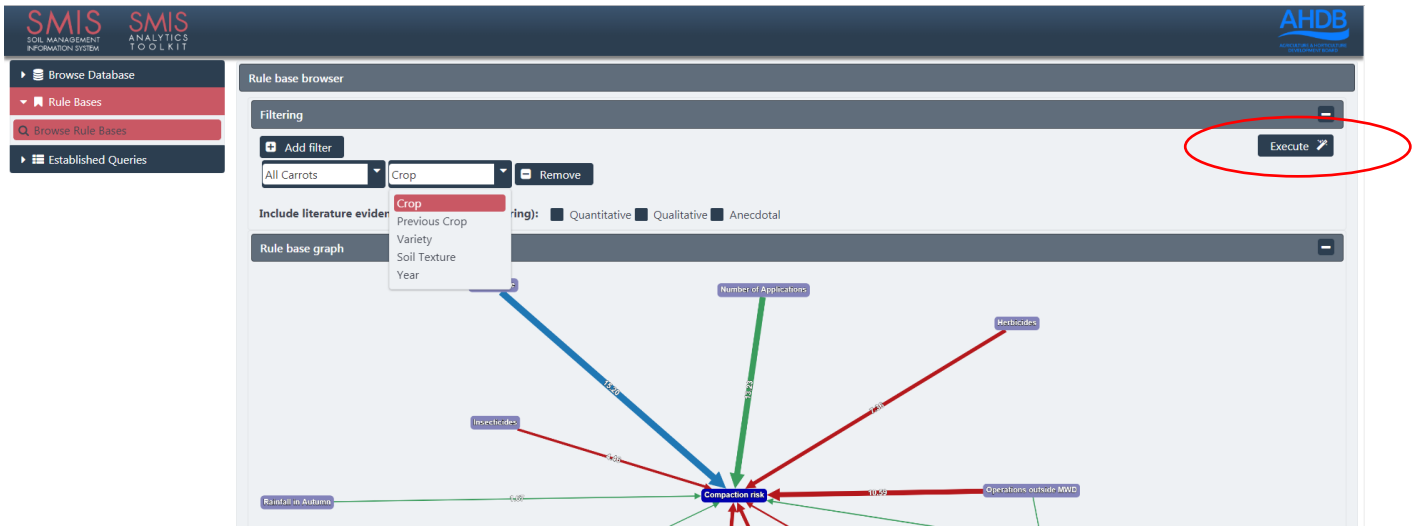

Figure 19. Filtering the Rule Bases

It is important that the 'Execute' button on the right hand side of the page is clicked (Figure 19) and refreshed when a new filter is added or removed.

3.1. Explaining the symbols used in the Rule Bases visualisation

- Blue nodes (or labels: Figure 20) show the identified soil management issues in horticulture.
- Purple nodes (or labels: Figure 20) are factors affecting these soil management issues. Clicking on these reveal the data behind the relationship (if any present) e.g. all crops (unfiltered), rainfall in autumn on compaction risk.
- The lines connect to nodes where the data hold a relationship. Clicking on these show where (and which) data are important in determining the relationship between cause and effect.
    - In the displays representing the grower data:
        - The thickness of the line reflects the strength of the relationship.

- Green lines reflect a positive effect i.e. increase the issue e.g. increased rainfall in autumn increases compaction risk.
- Red lines imply a negative effect i.e. reduce the issue e.g. Fewer operations outside 'mean workability days' causes a reduction in compaction risk.
- Blue lines represent categorical data, such as 'soil texture; previous crop', that cannot be statistically analysed in the same way as ordinal, interval or ratio data can be.
- Clicking on these lines will reveal the relationship between cause and effect (e.g. histogram of data collected), if it is available.
- The numbers on the lines reflect the relative importance of given variable in the model (scaled from 0 - 100).

- In the displays of the literature evidence:
  - Purple lines show evidence of cause and effect from the literature and whether this is 'quantitative', 'qualitative' or 'anecdotal'. These will only be visualised when the 'Include Literature Evidence' boxes are ticked (Figure 21)
  - The thickness of the line represents the number of articles / papers.
  - Clicking on the line will reveal the evidence (e.g. a citation of a paper from the literature. Clicking on the author [citation column] will take the user to the original source).

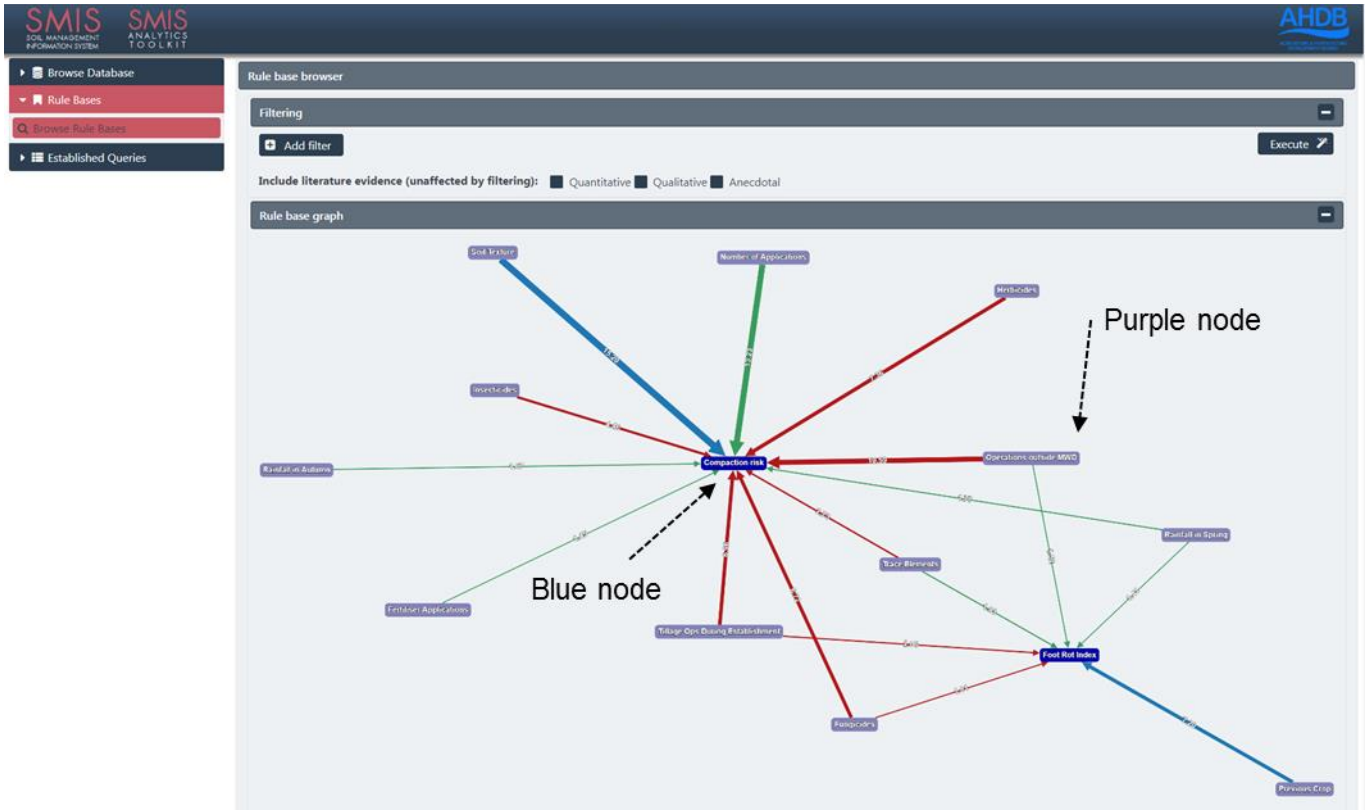

Figure 20. Illustration of Blue and Purple Nodes (labels) within the SMIS Visualisation Suite.

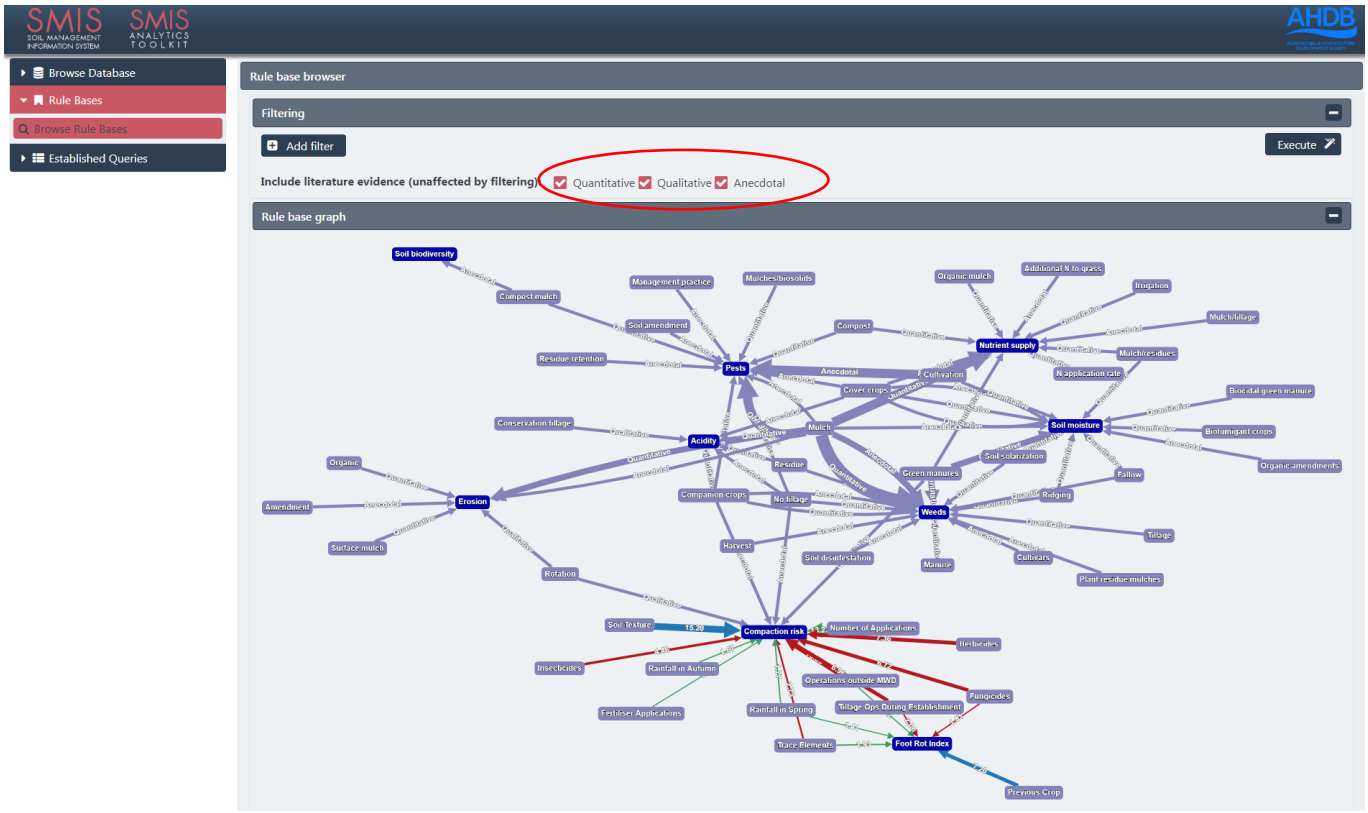

Figure 21. Visualisation of 'quantitative', 'qualitative' or 'anecdotal' evidence from the Literature Database

3.2. Interrogating SMIS Rule Bases

Figure 22 shows how to access the Rule Bases of SMIS, showing as an example, vining peas as the crop of interest. This shows that here, the only data available relating to soil management issues is 'Yield'.

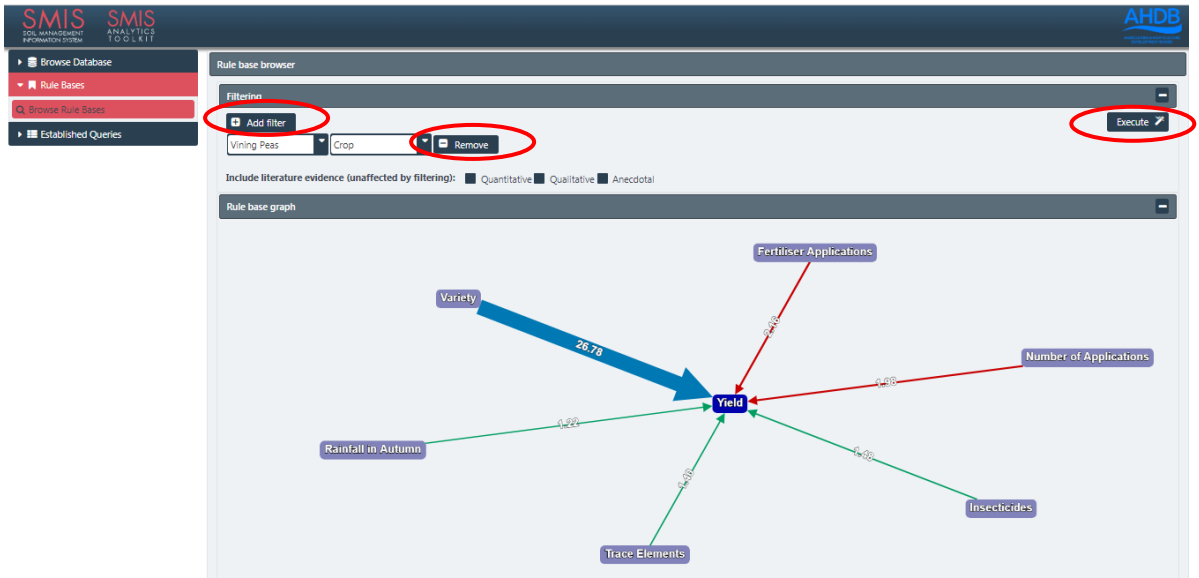

Figure 22. Browsing the Rule Bases: vining peas example

More filters can be added (or later removed) by clicking on the '+ Add filter' or '- Remove' boxes (Figure 22). (The more filters added, the narrower the database and fewer results will be displayed). Figure 23 shows the effect of adding 'soil type' (here, deep clay) to the filters – only instances of vining peas on clay soils will be displayed. Figure 24 shows the effect of changing soil type from deep clay to deep silt.

It is important that the 'Execute' button on the right hand side of the page is clicked and refreshed when a new filter is added or removed.

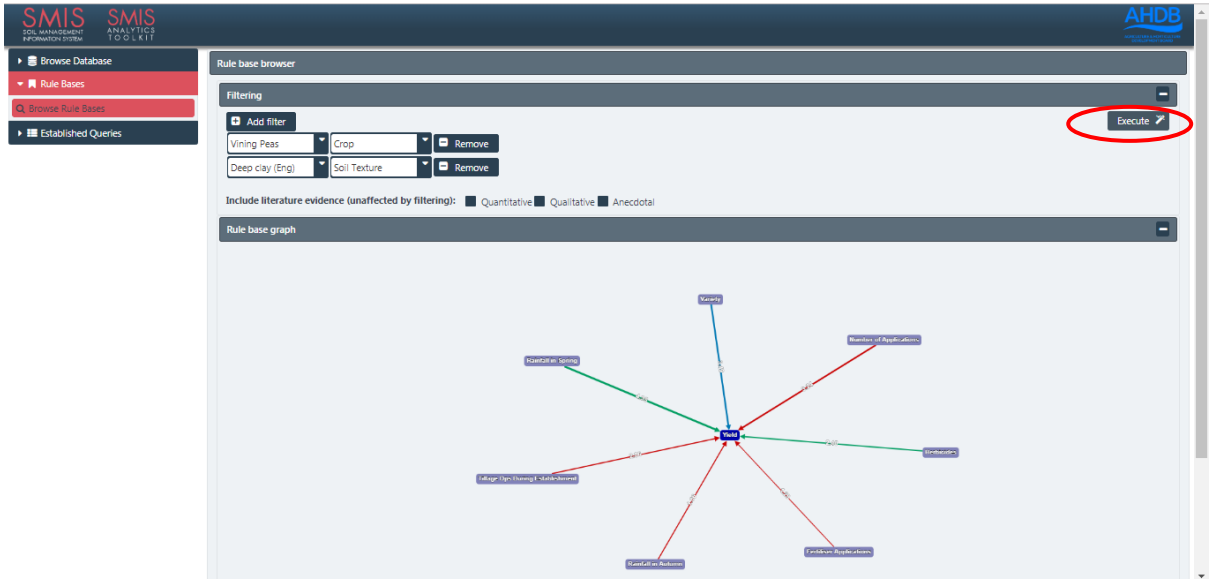

Figure 23. Rule Base for vining peas on deep clay soils.

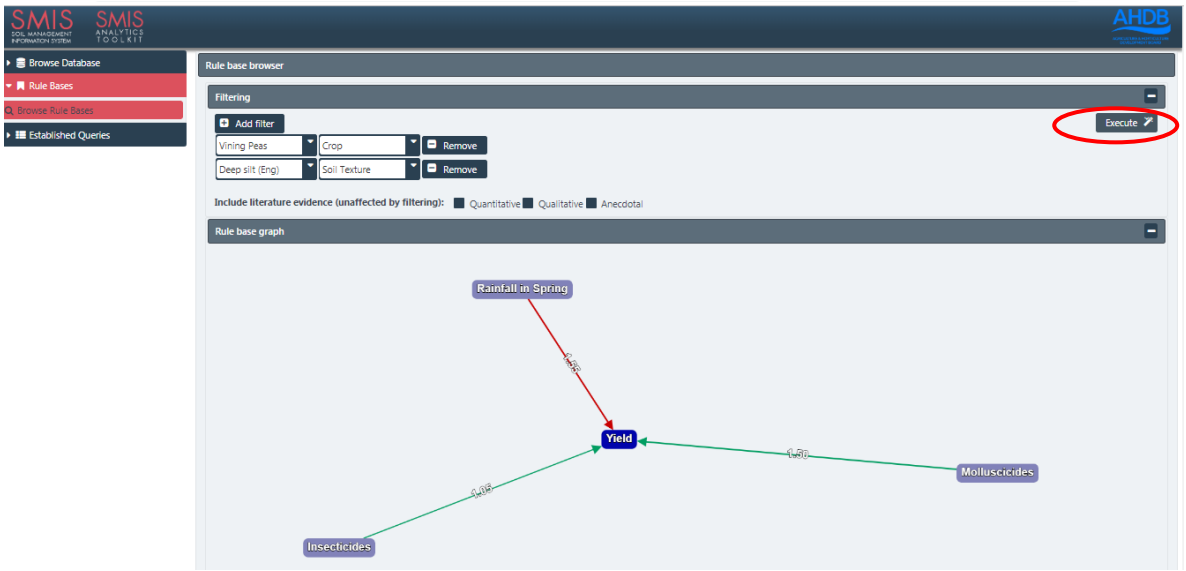

Figure 24. Rule Base for vining peas on deep silt soils.

Figure 25 shows the effect of selecting a crop with more extensive records in the database (here, Winter Wheat as an example). In this example, soil management issues of 'Yield' and 'Compaction Risk' are displayed, with factors affecting them shown as purple nodes. In some cases, these factors are common to both soil management issues. Clicking on any of the purple nodes (factors affecting the issue) reveals more information about that factor e.g. effect of Soil Texture (Figure 26) and effect of Previous Crop (Figure

27) on the incidence of soil compaction. Figure 28 shows the display of information on the soil management issue i.e. the effect of the number of operations outside Mean Workability Days (MWD) on the incidence of soil compaction in Winter Wheat.

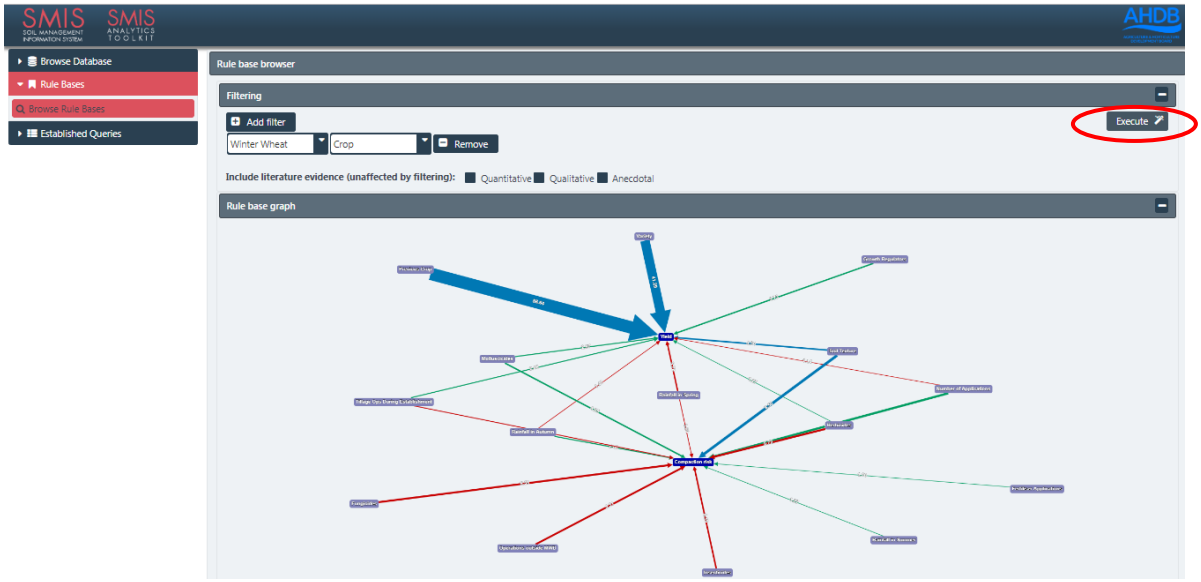

Figure 25. Rule Base for Winter Wheat

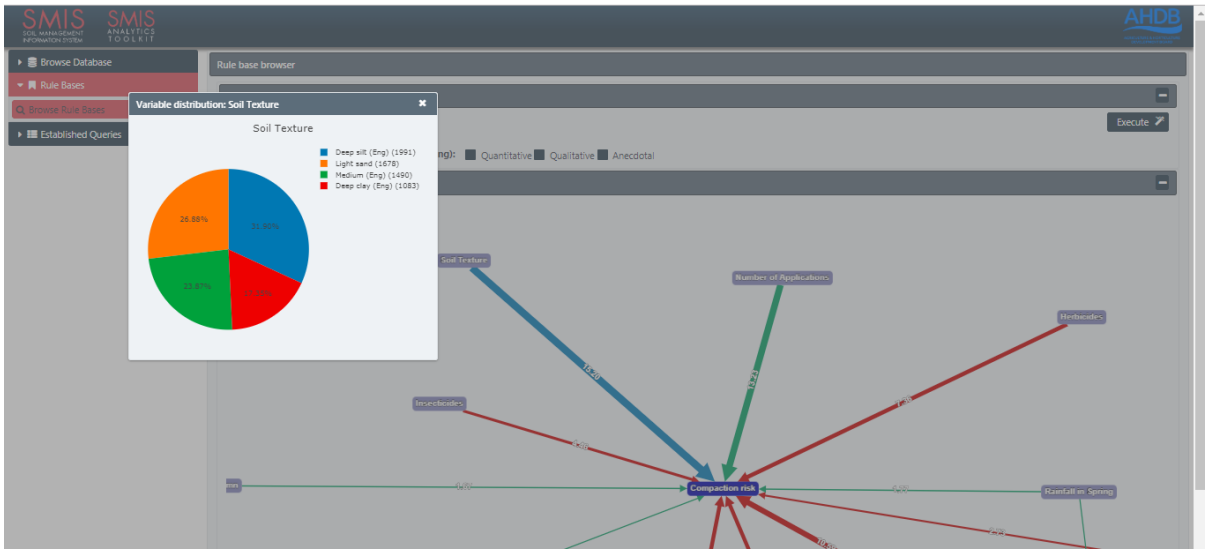

Figure 26. Revealing more information on the soil management issue: Effect of soil texture on incidence of soil compaction.

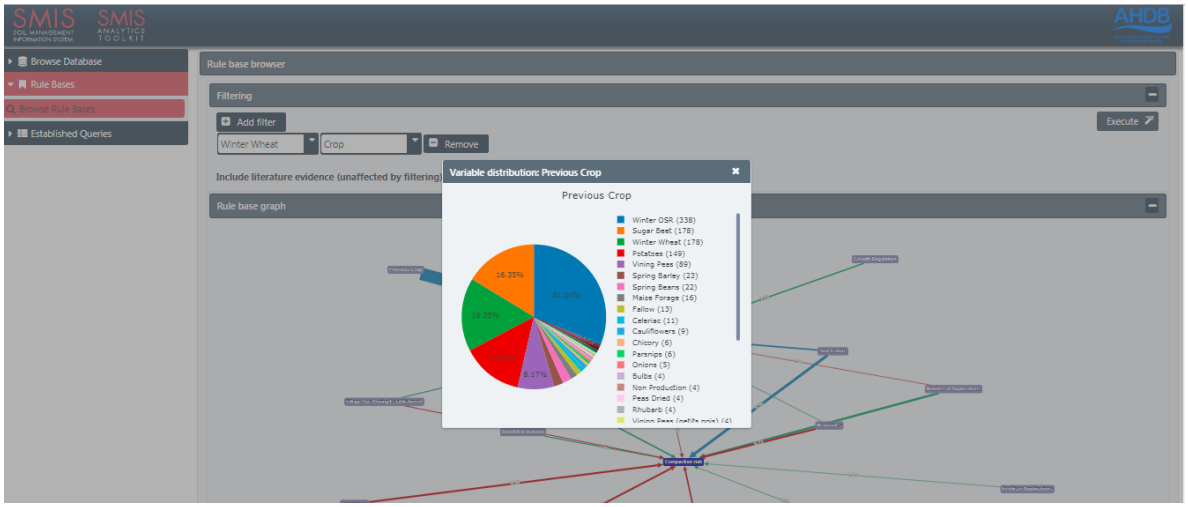

Figure 27. Revealing more information on the soil management issue: Effect of previous crop on incidence of soil compaction in Winter Wheat.

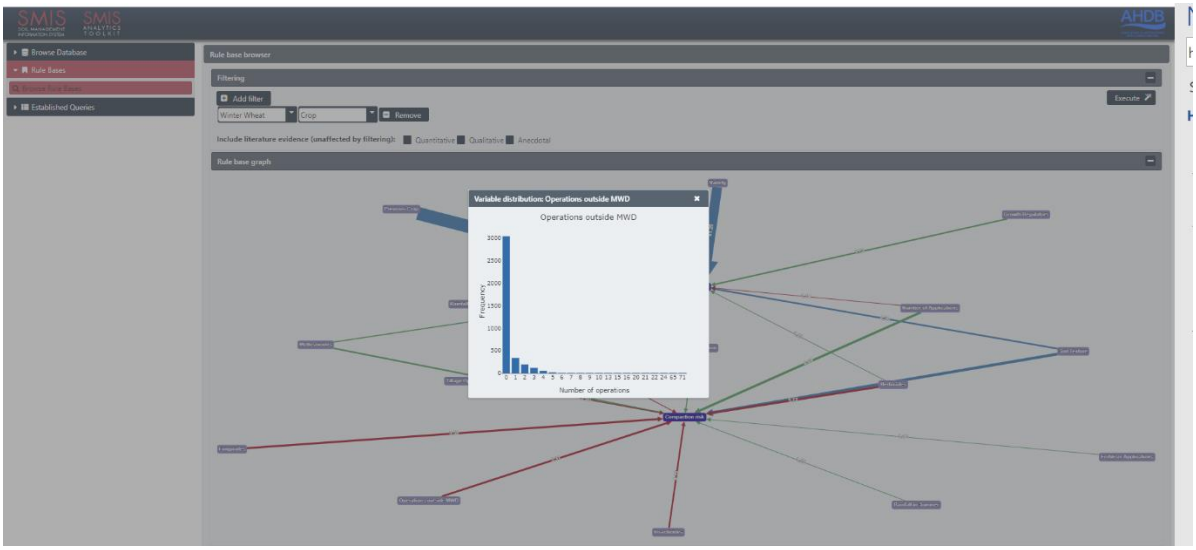

Figure 28. Revealing more information on the soil management issue: Effect of number of operations outside Mean Workability Days (MWD) on incidence of soil compaction in Winter Wheat.

All queries can be filtered by crop; previous crop; variety; soil texture and/or year. (For example, Figure 29. Rule Base for Winter Wheat on deep clay soils only: effect of crop variety). When changing options / filters, **it is important that the 'Execute' button on the right hand side of the page is clicked and refreshed when a new filter is added or removed.**

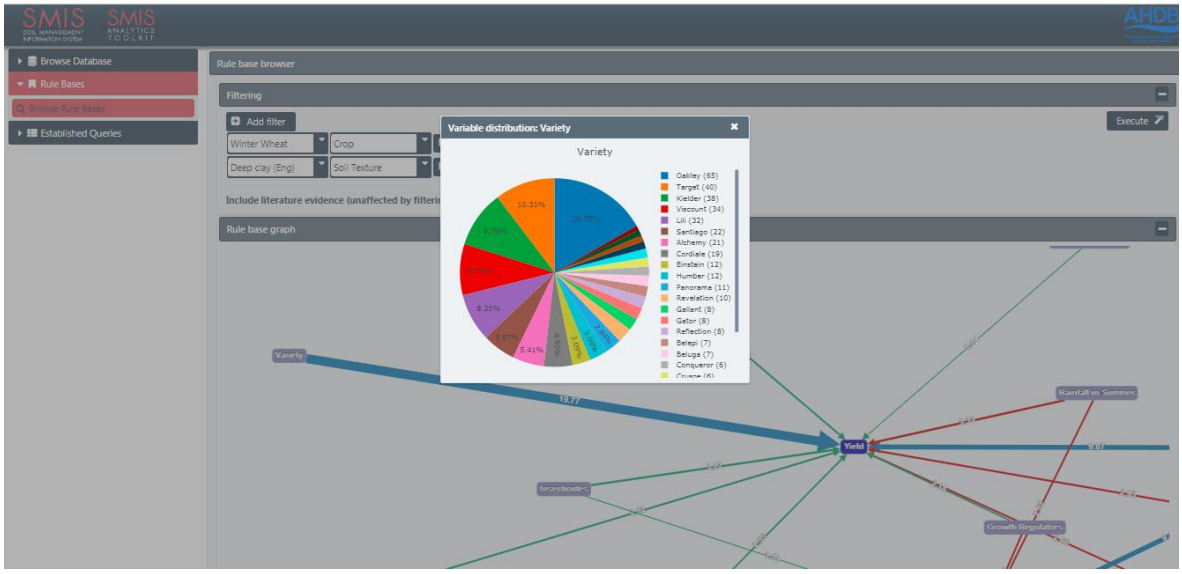

Figure 29. Rule Base for Winter Wheat on deep clay soils only: effect of crop variety on yield

3.2.1. Adding the literature and experimental data to the grower database

The user has the option of including all sources of information within SMIS: the grower data, literature and experimental evidence. Figure 30 shows the effect of adding the literature and experimental evidence to the grower database. Figure 30 shows the whole SMIS database; including literature evidence that is quantitative, qualitative and/ or anecdotal by checking the boxes next to “Include literature evidence”. The blue nodes show the soil management issues identified, including soil erosion, nutrient supply, foot rot index, compaction risk, soil moisture, pests, acidity, weeds and soil biodiversity. The green and red lines display the grower data and the purple lines display the evidence from the literature. The display can be filtered to focus on particular crops (e.g. Figure 31; potatoes); previous crops; variety; soil texture and year. An example of filtering can be seen in Figure 32, which shows the visualisation of the grower database, but including anecdotal literature only.

For some crops (e.g. carrots), adding the literature reveals considerably more relationships between cause and effect than the Grower data alone. As more data is added to SMIS, the relationships will be stronger (with more confidence).

It is important that the ‘Execute’ button on the right hand side of the page is clicked and refreshed when a new filter is added or removed.

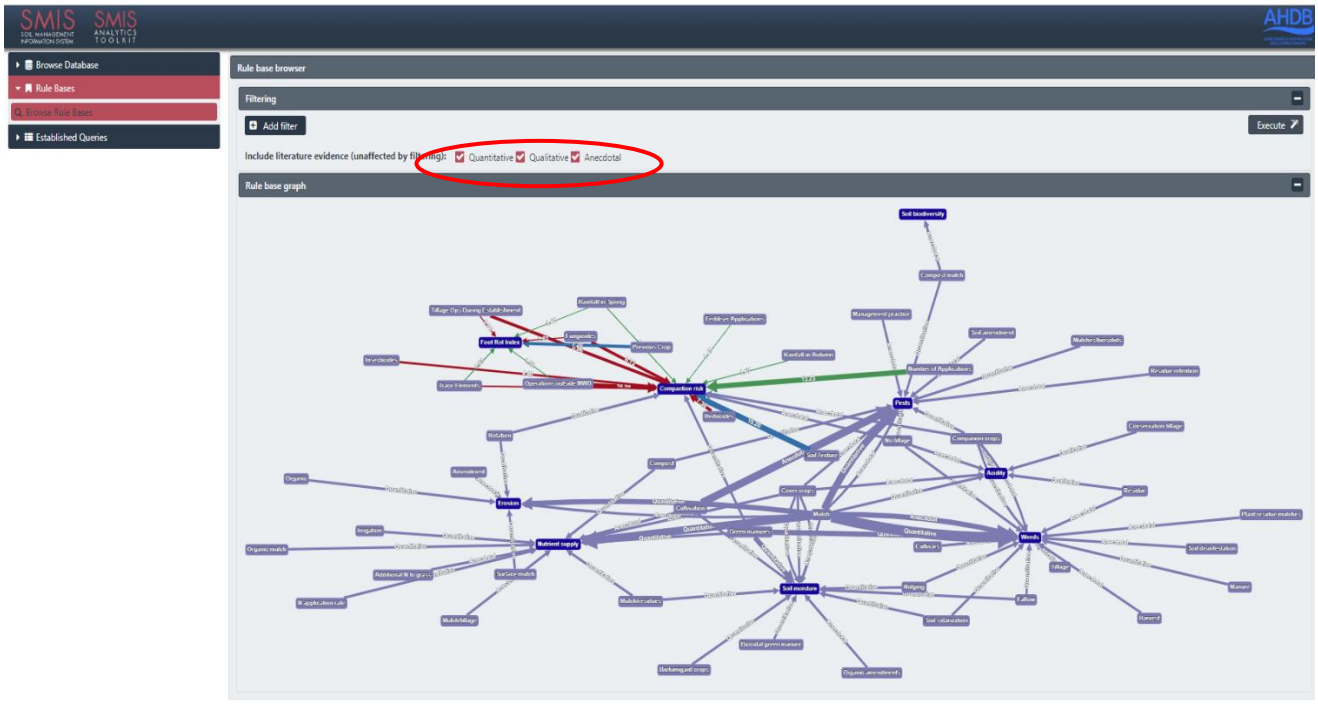

Figure 30. Visualising the SMIS database: grower data, literature evidence and experimental data

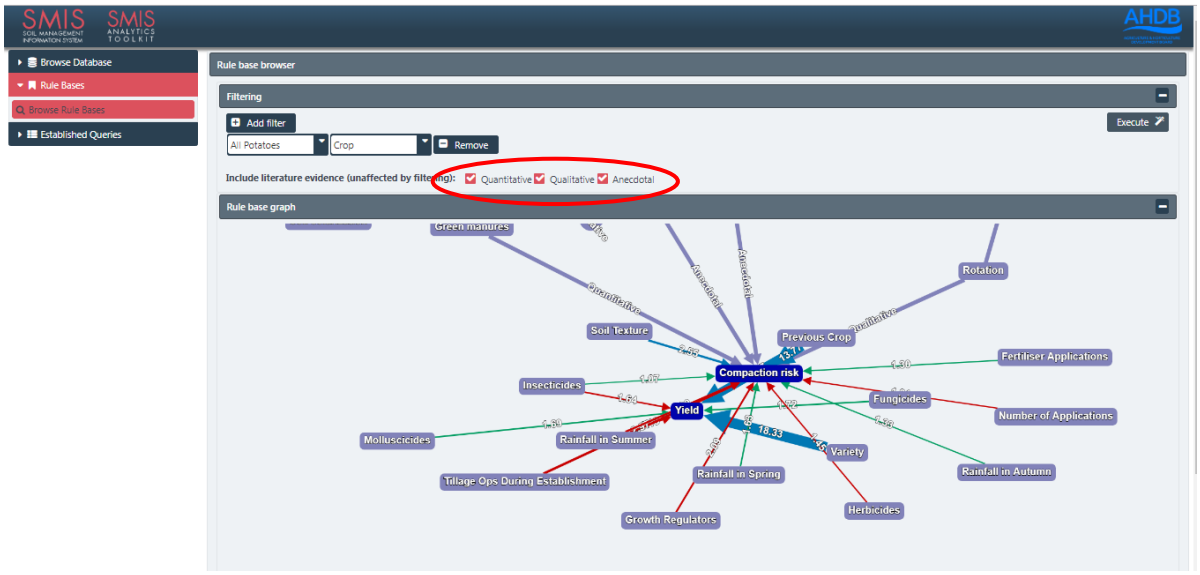

Figure 31. Factors affecting compaction risk in potatoes.

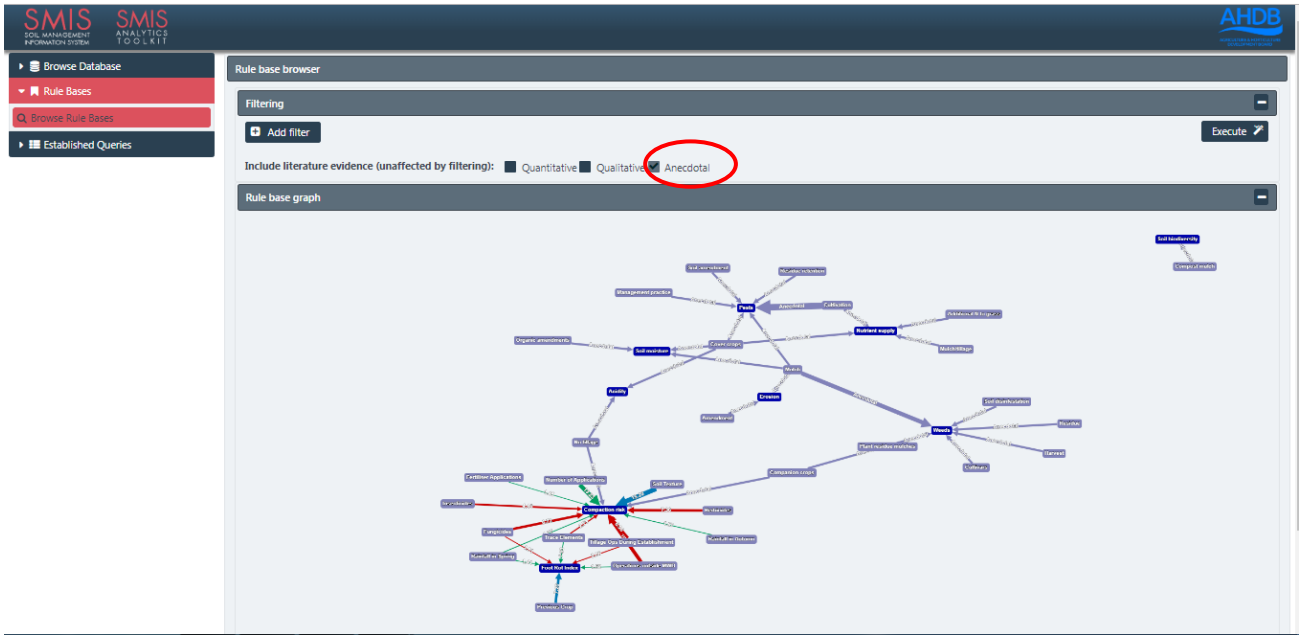

Figure 32. Visualisation of grower database with only anecdotal literature.

In the example of Figure 33, the three sources of data / information dealing with vining peas are not connected. This is because there are no instances of where the grower data, literature and experimental data have common linkages. In other words, 'Yield' (blue label in bottom right hand corner) is only found in the grower database: there are no records on vining pea yields related to soil management in the literature or experimental data. Similarly, there are no records in the literature or experimental data on soil management effects on Foot Rot Index (circled in green in Figure 33). This is explained by the narrow focus of the literature review (on horticultural crops only) and yet the grower database includes a lot of non-horticultural crops due to the inclusion of cross rotational data. These non horticultural crops were deliberately not in scope for the literature review. Also, as stated previously, there are few links between the experimental data and grower data, because: a) only 2 sites (fields) that appear in the Grower database have provided experimental data; b) there are no common (shared) outputs (e.g. yield) or variables predicting / explaining the outputs (e.g. bulk density; organic matter); and c) the dataset is too small for any statistical relationships (e.g. linear regression modelling) to be run.

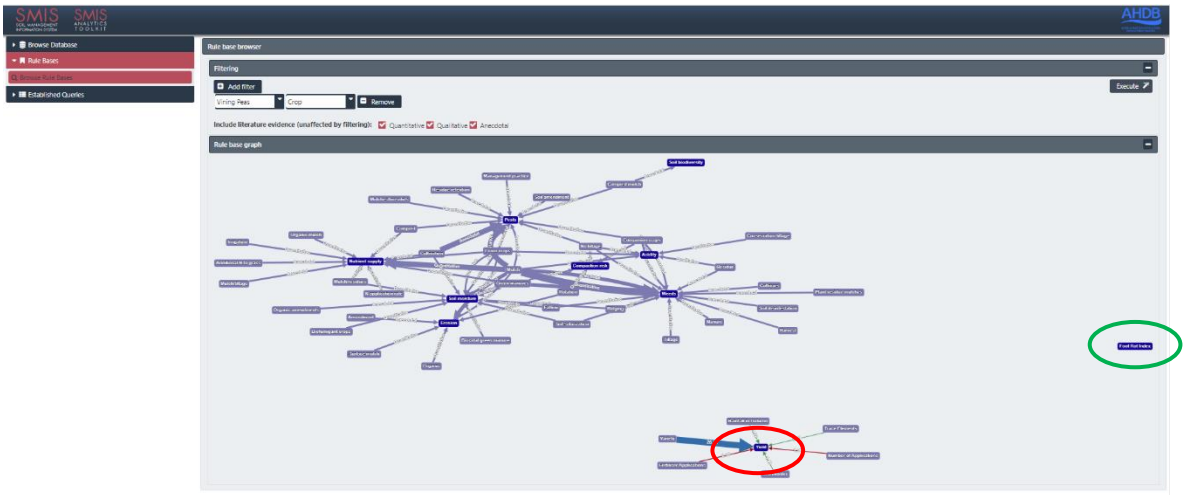

Figure 33. Visualisation of grower data and all literature data for vining peas.

The thickness of lines represents the strength of the relationships. For the literature data, clicking on the purple lines reveals the sources of literature that infer the relationship. For example, Figure 34 shows the number of items of literature (2) referring to the effect of mulches on soil erosion in carrots.
→ Clicking on the authors names in the Citation column then directs the end user to the original literature source (Figure 17).

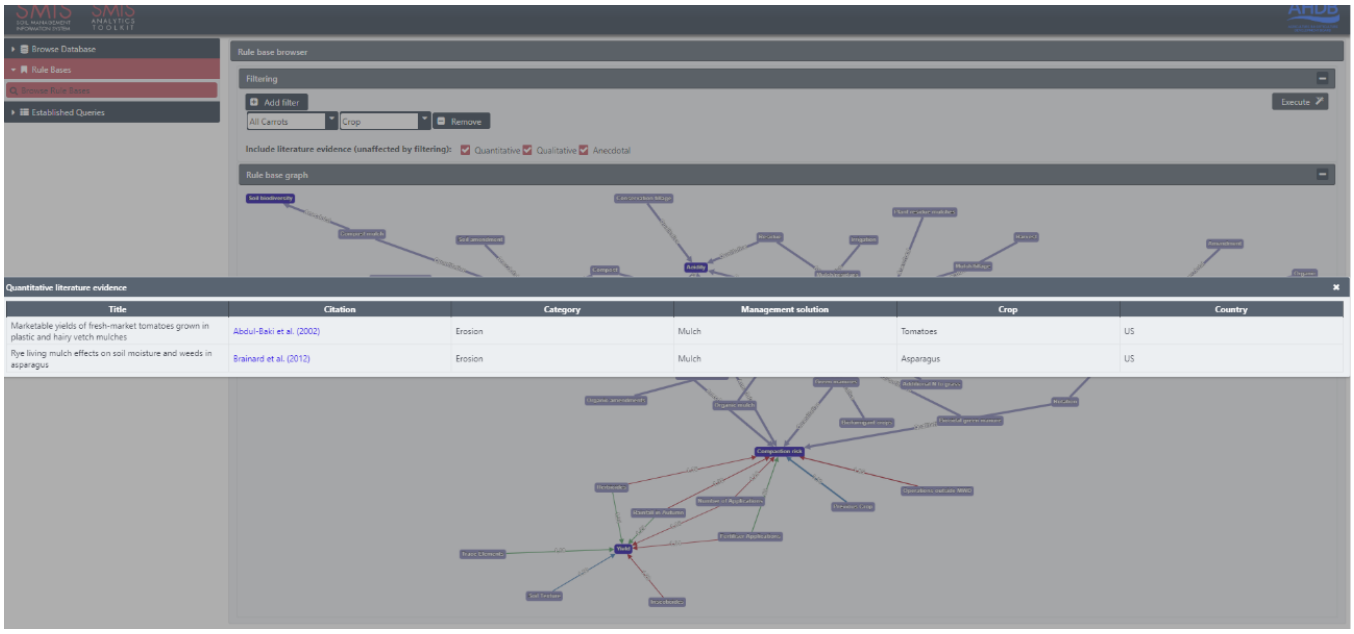

Figure 34. Identifying the literature sources that underpin the relationships between cause and effect.

4. Established Queries

The Established Queries function of SMIS aims to better understand and analyse the cause and effects of the horticultural industry's specific soil management challenges, as identified in the gap analysis of soil management research and knowledge transfer in horticulture by Rickson and Deeks (2013) and stakeholder feedback during the development of SMIS.

This application of SMIS can be activated by clicking on the 'Established Queries' tab on the left hand side of the SMIS home page (Figure 35). This reveals the list of established queries that can be run given the current data within SMIS. These were agreed by delegates at the Stakeholder Workshops in June and July 2018. They are:

- Factors affecting yield
- Compaction risk
- Foot rot index
- PCN levels
- Cavity spot.

It is important that the 'Execute' button on the right hand side of the page is clicked and refreshed when a new filter is added or removed.

The rows in the spreadsheet represent individual data entries from the grower database (Figure 35). The user can expand or decrease the number of rows shown (10, 20 or 30) using the drop down menu at the bottom of the spreadsheet. The user can scroll through the database using the arrow buttons, also at the bottom of the spreadsheet.

It should be noted that not all queries will generate output (in the form of pie charts, histograms etc.). This is because at present there is insufficient data in the system to generate the statistical relationships needed to generate these displays.

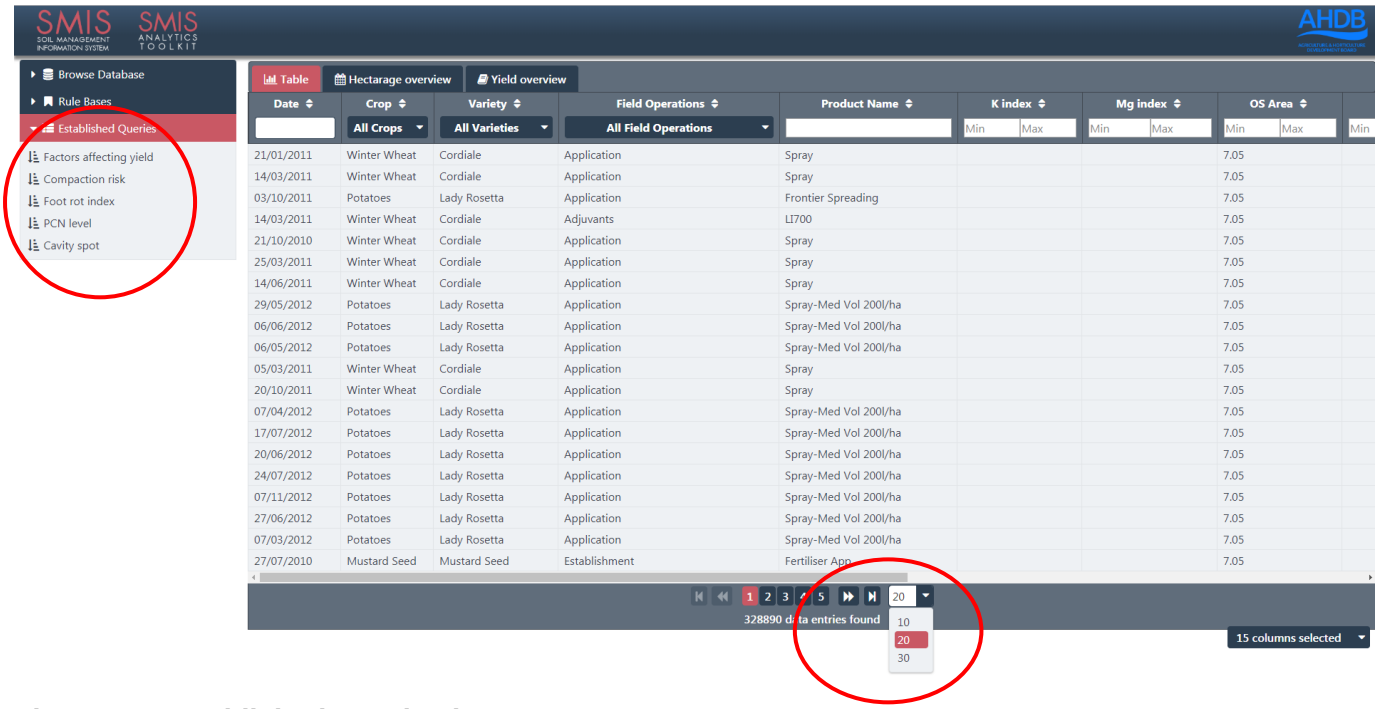

Figure 35. Established Queries home page.

4.1. Factors affecting yield

First, the crop of interest must be selected in the Query Constructor box. The down drop menu reveals all the crops in SMIS, as listed in Table 2. **It is important that the 'Execute' button on the right hand side of the page is clicked and refreshed when a new filter is added or removed**. If sufficient data are available for the analysis, a histogram will appear, with all the factors related to that crop's yield. For example, Figure 36 shows the display for Winter Wheat, listing: Previous crop; Variety; Soil Texture; Rainfall in spring; growth regulators; molluscicides; tillage operations during establishment; rainfall in autumn; herbicides and number of applications as being the most important factors affecting Winter Wheat yield. The Y axis (labelled 'Variable importance') has values from 0 – 100, where a figure of 100 would mean a particular predictor variable explains all of the variation in the output (here, yield). For example, in Figure 36, this means that 58% of the variation in wheat yield is due to the previous crop and 41% is associated with crop variety.

Hovering over the bars will reveal further information about that factor, usually in the form of a pie chart or histogram. For example, Figure 37 shows a pie chart of the breakdown of Previous Crops, with winter oilseed rape being most prevalent before Winter Wheat within the dataset.

Clicking on the Previous Crop bar then reveals which crops (across all soil types and years) have the most effect on Winter Wheat yield (Figure 38). The same procedure can be repeated to investigate the impact of crop variety on Winter Wheat yield (Figure 39). Varieties with 'Green' bars indicate those varieties in the database (across all soil types and years) that are associated with a positive influence on yields. Those varieties with no bar shown have neither a positive or negative effect on yield. Those varieties that are associated with comparatively poor yields are indicated by 'Red' bars in this instance, 'Reflection'. The Y axes labelled "Linear model coefficient" describe the relationship between a predictor variable (e.g. previous crop) and the response (e.g. yield of the following crop). The coefficient value represents the mean change in the response given a one unit change in the predictor. For example, in Figure 39, Variety 'Cordiale' has a positive effect on winter wheat yield, whereas 'Reflection' has a negative effect on winter wheat yield.

When new queries are to be run, it is important to 'untick' the option used in the previous query (unless it is to be included in the 'new' query).

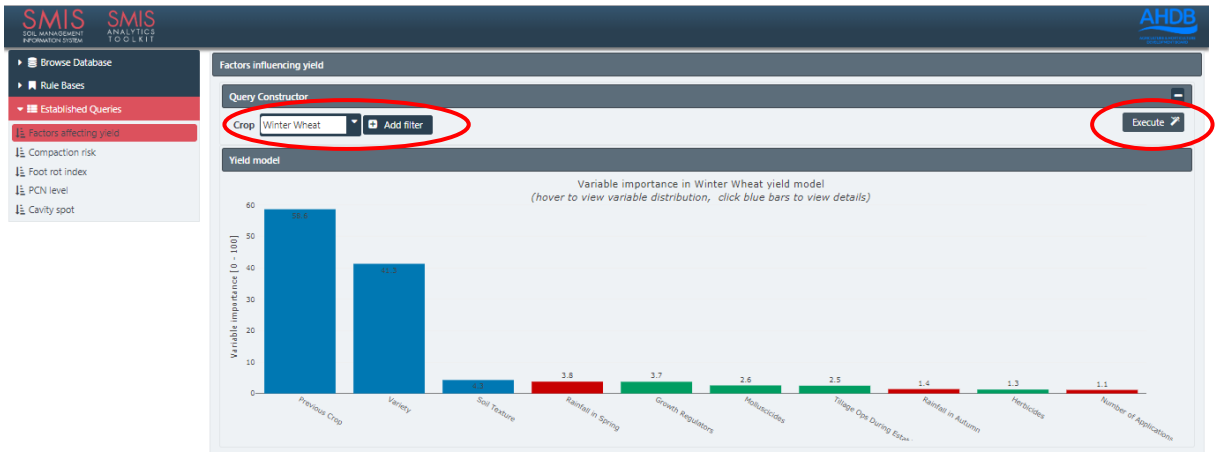

Figure 36. Factors affecting yield in the Established Queries function of SMIS.

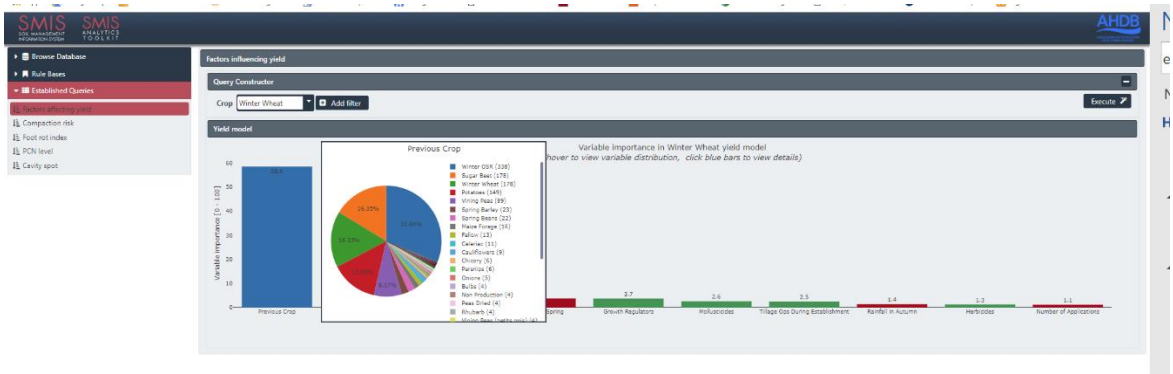

Figure 37. Previous crops grown before Winter Wheat have most effect on yield. Pie chart shows the prevalence of the different previous crops in the SMIS database.

Figure 38. Previous crop impact on Winter Wheat.

Figure 39. Impact of crop variety on yield.

In the Query Constructor, more filters can be added by clicking on the ‘Add filter’ button. The filters are:

- Previous crop;
- Crop variety;
- Soil Texture; and
- Year.

It is important that the ‘<u>Execute</u>’ button on the right hand side of the page is clicked and refreshed when a new filter is added or removed.

For example, Figure 40 shows the result of Winter Wheat yields for deep clay soils only. This reveals that ‘Variety’ rather than ‘Previous crop’ (as for all soils Figure 24) is the most important factor affecting yield. Again, hovering over the bars will reveal further information about that factor, usually in the form of a pie chart or histogram.

In the example of Winter Wheat on deep clay soils, clicking on the ‘Variety’ bar then reveals which varieties have the most positive (Green) or negative (Red) effect on Winter Wheat yield (Figure 41). The same procedure can be repeated to investigate the impact of crop variety on Winter Wheat yield (Figure 41). This indicates that across all years, that Humber, Cordial and Solo varieties perform comparatively poorly on deep clay soils with ‘Reflection associated with comparatively high yields. If the same procedure is

repeated for light sand soils, Viscount, Oakley and Target are considered poor yielding varieties (Figure 42).

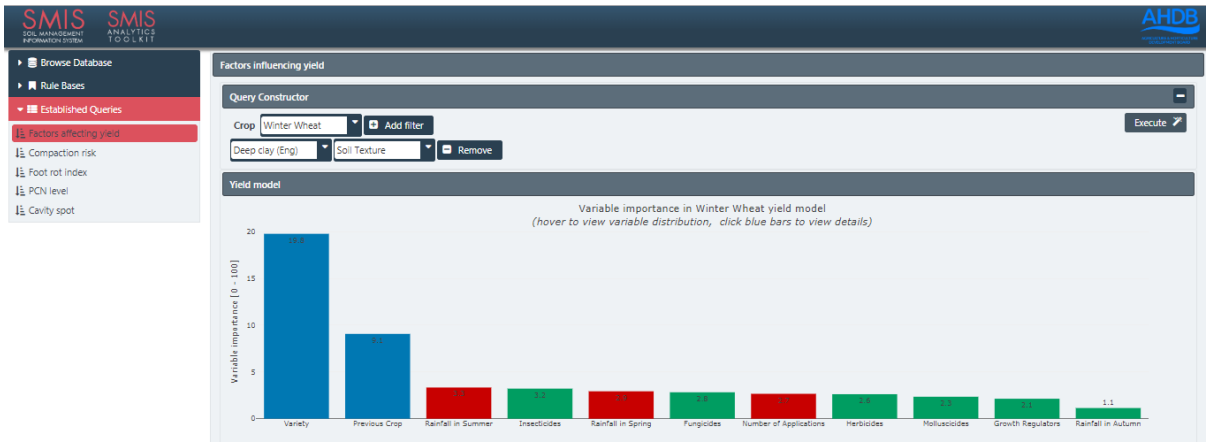

Figure 40. Winter wheat yields for deep clay soils only.

Figure 41. Effect of crop variety on Winter Wheat yields on deep clay soils.

Figure 42 . Effect of crop variety on Winter Wheat yields on light sand soils

Figure 43 and Figure 44 show how SMIS can reveal the difference in factors affecting yield for different years, when the 'Year' filter is used. In these examples, the crop selected is oilseed rape. The years are 2011 and 2012. In 2011, the data show several factors affecting yield: 'crop variety' being the most important. However, in 2012, a particularly wet summer and autumn, 'soil texture' was the most important factor affecting yield. Again, further information is revealed by clicking on the factor of interest. Figure 45 shows that in 2012, oilseed rape yields were positively affected by 'deep silt' soils (Green bar) but negatively affected (Red bar) by 'medium' soils. There is no bar for 'deep clay soils' suggesting insufficient data on oilseed rape yields in 2012 is available for that soil type in that year.

Figure 43. Factors affecting oilseed rape yields in 2011.

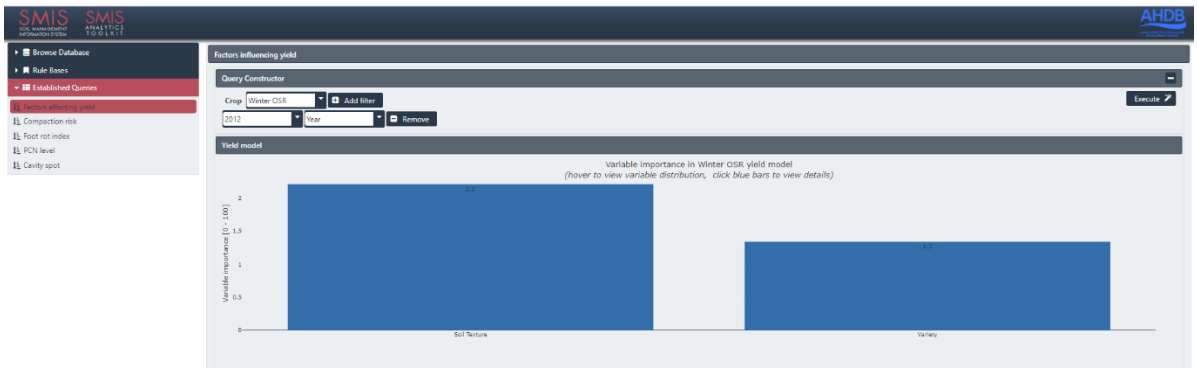

Figure 44. Factors affecting oilseed rape yields in 2012.

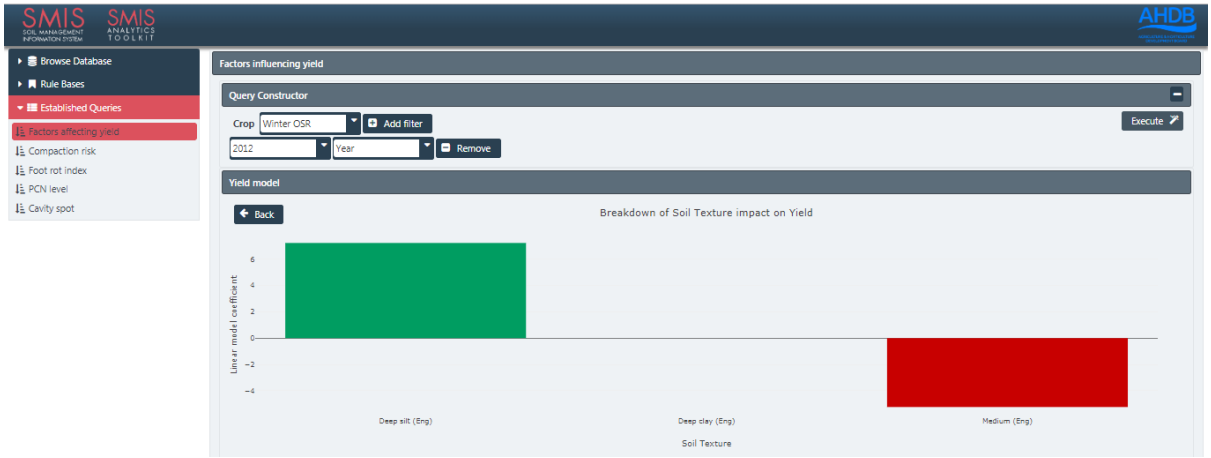

Figure 45. Influence of soil texture type on oilseed rape yield in 2012.

4.2. Compaction risk

Figure 46 shows the example of running the SMIS Established Query for soil compaction risk in carrots. It demonstrates that Previous Crop, Operations outside MWDs, Herbicides and Soil Texture have an effect on the risk of soil compaction in carrots (in this example). These results highlight what can be done to reduce compaction risk, so informing better soil management decisions in the future.

Figure 46. Factors affecting compaction risk in carrots.

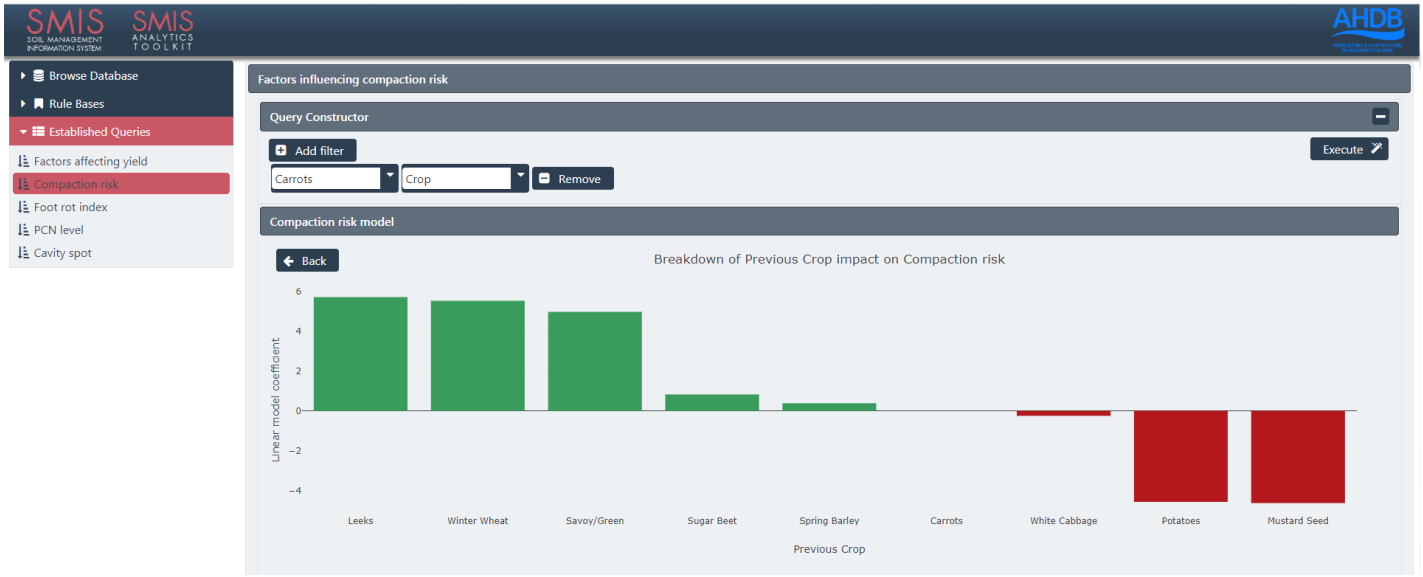

Figure 47. Effect of Previous Crop on compaction risk in carrots in the SMIS database.

→ Clicking on each bar will reveal further information about that factor. For example, clicking on the Previous crop bar in Figure 46 will reveal which previous crops contribute most to compaction risk (Figure 47). Figure 47 shows that leeks, winter wheat, savoy/green cabbage, sugar beet and spring barley as previous crops increased compaction risk in the following crop (green bars), whereas white cabbage, potatoes and mustard seed (red bars) had a negative effect on compaction risk. Carrots (no bar) appeared to have no effect on the level of compaction risk in the following crop. (This is a good example of SMIS raising more questions, e.g. why would growing savoy and green cabbage *increase* compaction risk in the following season, but white cabbage *reduces* it?).

4.3. Foot rot index

Figure 48 shows the factors affecting foot rot index according to the SMIS database. 'Previous crop' has the highest importance. Hovering over this bar then reveals the previous crops that precede instances of foot rot index (FRI) values (Figure 49). [It should be noted that sometimes FRI is measured 1-year or 2-yrs before the pea crop is grown].

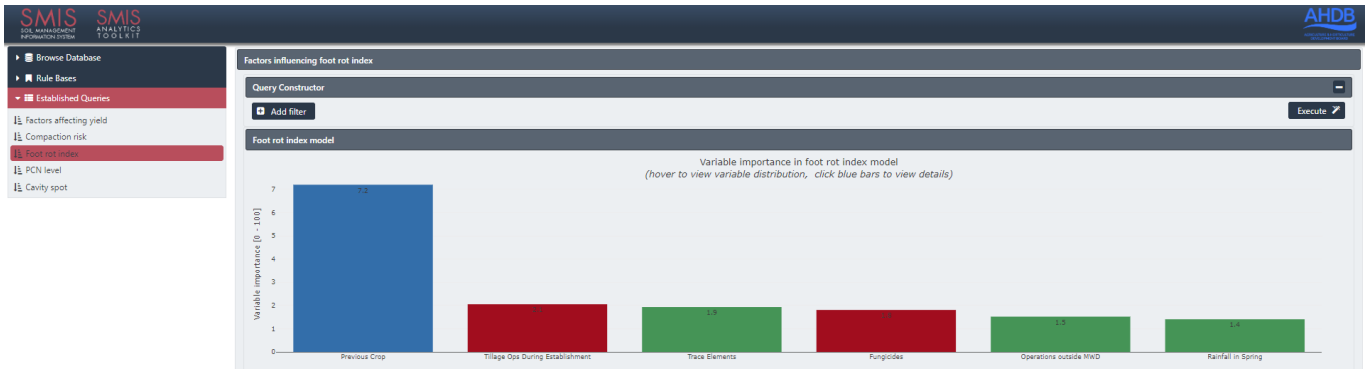

Figure 48. Factors affecting foot rot index (all crops)

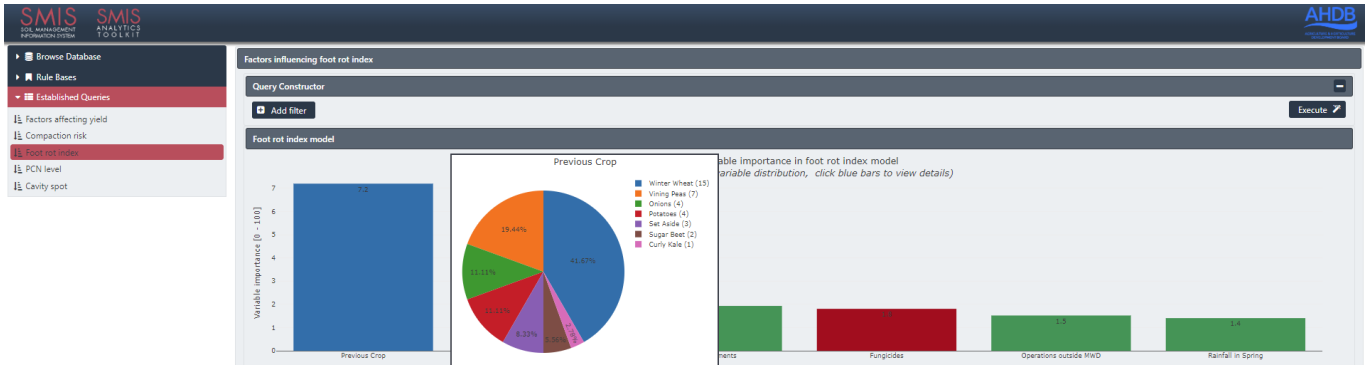

Figure 49. Hovering over 'Previous crop' reveals the crops that have preceded instances of foot rot index measurement.

By clicking on the 'Previous crop' bar, another histogram is displayed showing that of the previous crops, onion had the greatest (positive) relationship with foot rot index (Figure 50). In this example, potatoes had a negative effect on the foot rot index value. The negative effect of potatoes on the foot rot index value is in fact encouraging as it indicates a reduction in the foot rot index value. This could be in part due to the number of pre-emptive fungicide spray applications applied to potato crops, but that also have impact on foot rot risk in following crops.

Figure 50. A preceding crop of onions has the most impact on foot rot: potatoes have a negative effect on the foot rot index value indicating a positive response.

4.4. PCN level

Limited data on PCN counts was provided by participating growers. As such the outputs generated in Figure 51 and Figure 52 are to demonstrate SMIS functionality only and should not be taken as robust cause and effect relationships. As more data is added to SMIS with a specific focus on potatoes and PCN levels, the system architecture and functionalities are in place to develop and display any robust relationships.

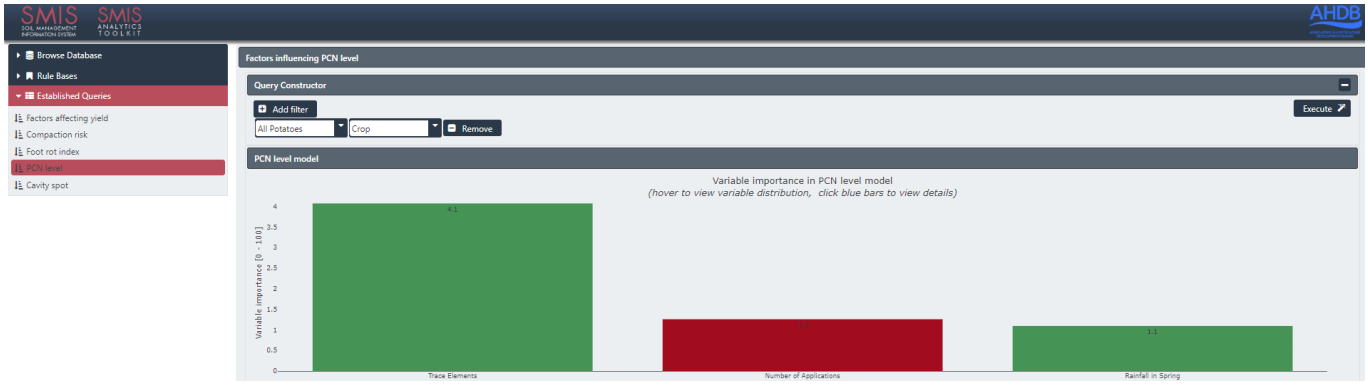

Figure 51. Factors affecting PCN levels in potatoes

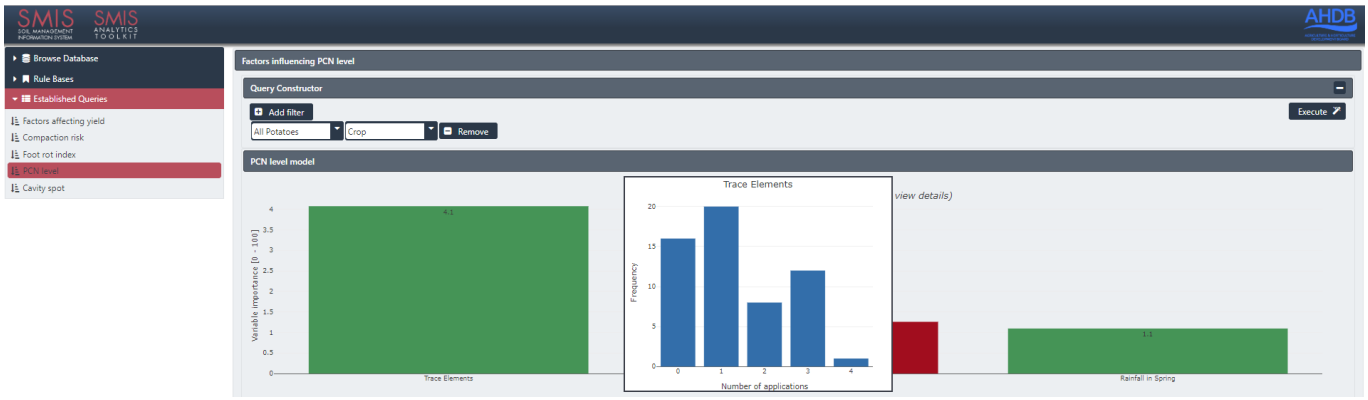

Figure 52. Number of applications of trace elements associated with incidence of PCN in potatoes.

4.5. Cavity spot

It should be noted that not all queries will generate output (in the form of pie charts, histograms, links to articles etc.). This is because at present there is insufficient data in the system to generate the statistical relationships needed to generate these displays. As more data is uploaded into SMIS, more scenarios can be developed as the system architecture and functionalities are already in place.

Figure 53 shows the display when there is insufficient data within SMIS to generate the cause and effect relationships (in this example, regarding cavity spot incidence). This is due to the fact that the incidence and severity of cavity spot is not recorded in Gatekeeper, but generated from grower pack house data management systems. This pack house data was not provided by participating growers due to complications in isolating individual fields and to protect commercial sensitivities.

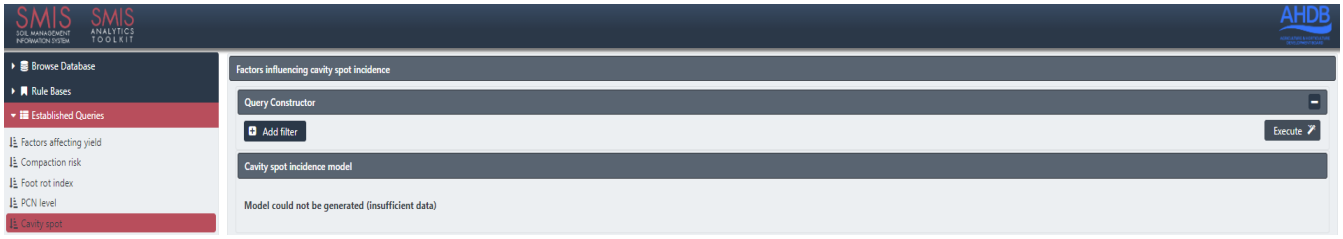

Figure 53. Cavity spot established query: insufficient data to generate a model or display.

Table 8 shows examples of queries that were run in the Established Queries component of SMIS during the June 2018 Stakeholder Workshop.

Table 8. Examples of queries run in the Established Queries component of SMIS.

Query	Vining peas?	Vining peas in 2015?	Oasis variety (all years)?	Oasis variety In deep clays?
What are the top 3 main factors affecting yield?	1. Variety 2. Fertiliser applications (-) 3. Number of applications (-)	1. Fertiliser applications (-) 2. Mollscicides (+) 3. Tillage operations during establishment (+)	1. Rainfall in spring (-) 2. Tillage operations during establishment (-) 3. Rainfall in autumn (-)	1. Rainfall in spring (+) 2. Number of applications (-) 3. Herbicides (+)

5. References

Rickson, R.J. and Deeks, L.K. (2013) A gap analysis of soil management research and knowledge transfer in horticulture to inform future research programmes. Final report to the Horticultural Development Company. 38pp. + 4 appendices. NSRI, Cranfield University, UK.