Protocol for extracting data from Gate Keeper

Project overview

The aim of this 3-yr project (2015-2018) is to apply the principles of 'Big data' [the ability to draw useful information from large and diverse data sets] to provide best practice guidelines for sustainable soil management in horticulture. The Soil Management Information System (SMIS) will hold, manipulate and manage available sources of data/information pertaining to the specific effects of soil management practices on horticultural crop productivity and environmental protection (within a rotational context). The project will

- Use novel informatics techniques to create and then interrogate a rule base of soil management practices (and their outcomes) in different scenarios (soil type, crop, rotation, location, etc.)
- Provide users with a set of robust, evidence-based, best-practice soil management guidelines (and the likely consequences of applying them)
- Create an interactive platform, giving AHDB-Horticulture, and its growers, agronomists and land managers access to guidance on contextual, effective soil management practices

We already have whole farm (Gatekeeper) datasets from several growers, covering over 57,000 ha for the time span 2008-2015. This includes a number of horticultural crops [e.g. onions, celeriac, peas (vining and dried), French beans, spring beans] grown within a range of rotational contexts.

However, as highlighted by Martin Evans and John Chinn at the recent stakeholder workshop (June 2017), it is important that we bring in data from a larger number of field vegetable growers representing multiple grower associations.

Data requirement

We are looking for growers to provide us access to their Gatekeeper data. This will involve exporting specific data from your Gatekeeper data (process detailed in the Appendix) as an .xml file.

The specific data that we require pertains to soil management over multiple years (for a full rotational context) and should include: cropping history (including crop yields), soil analysis data, timing of all operations, products applied (and their timing), crop, variety, yield etc.

We have found that the majority of the data that we require can be found under the following options in the Gatekeeper 'Analysis' menu list: 'Field Operations' (see Appendix, Step 5.2) and 'Previous Cropping' (see Appendix, Step 2.2). Under which the selection of specified headings (see Appendix, Step 2.4 and 5.4) are required to capture the precise data needed.

The exceptions are where you hold some data e.g. field specific soil analyses or yields outside of the Gatekeeper system. In which case this data should be sent separately –either in hard copy format for later return, as scanned data or as an electronic copy.

It is important to note that your data will be held securely and anonymously within SMIS. No financial data is held.

Data output

In return your data (per field basis) will be linked to various outputs from LandIS (www.landis.org.uk), namely:

'Machinery Workability Days', the number of days when the land can be worked with acceptable risk of damage to soil structure.

'Susceptibility to compaction', the risk of soil compaction with each machinery/field operation based on the timing of operation relative to the machinery workability days.

'Susceptibility to topsoil slaking', a topsoil stability assessment indicating the risk of slaking (surface capping).

'Potential for natural regeneration', the potential for the soil to recover naturally from compaction and the time period (if any) over which this can be expected to occur.

APPENDIX

<u>**Gatekeeper data extraction**</u>

Step 1: Refreshing Gatekeeper

Step 1.1:

Click on the Analysis Icon (circled below). This will open up the 'Analysis window'.
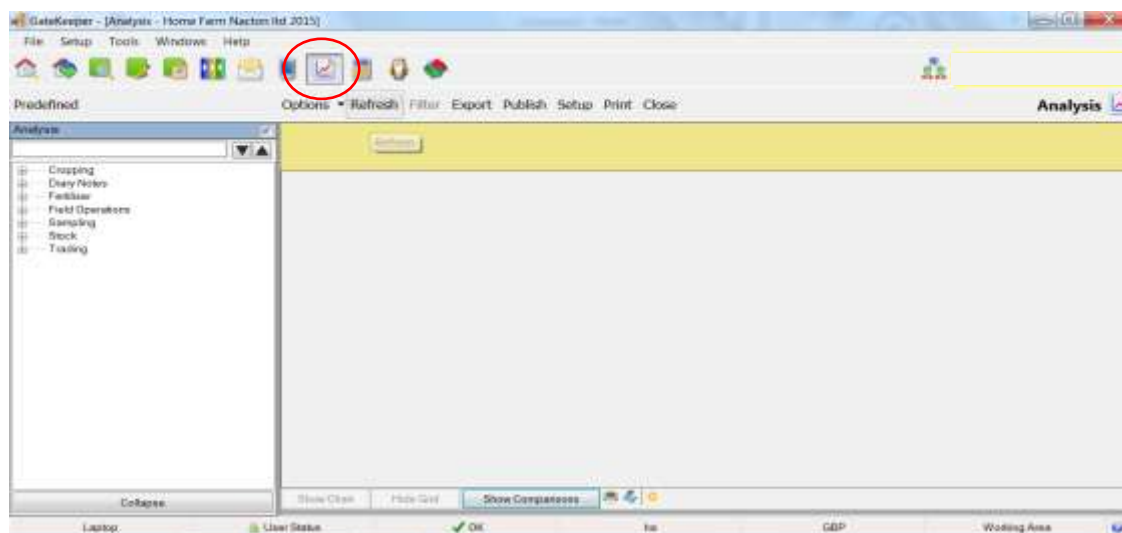
Step 1.2:

Click on the 'Refresh tab' (circled below). This opens the 'Refresh Analysis Summary Data' window (shown below). Select the years of interest that need to be 'Refreshed'.
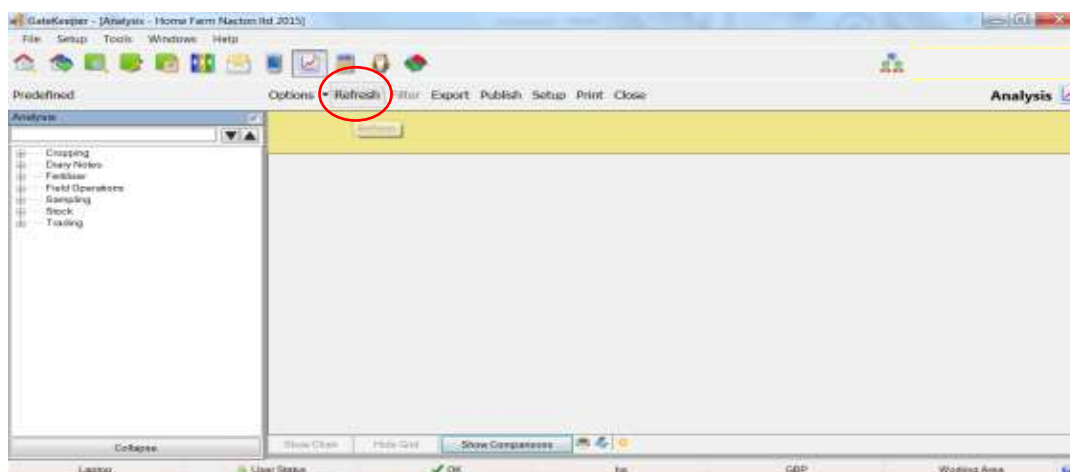
Step 1.3:

Once the years have been selected, click the 'Refresh button' on the bottom right of the window (circled below) to refresh.

Note: If the 'Refresh button' is hidden by the task bar (as is the case below) press the TAB button on the keyboard to select the 'Refresh button' and then press ENTER. This will finalise the refresh action.

Refresh Analysis Summary Data
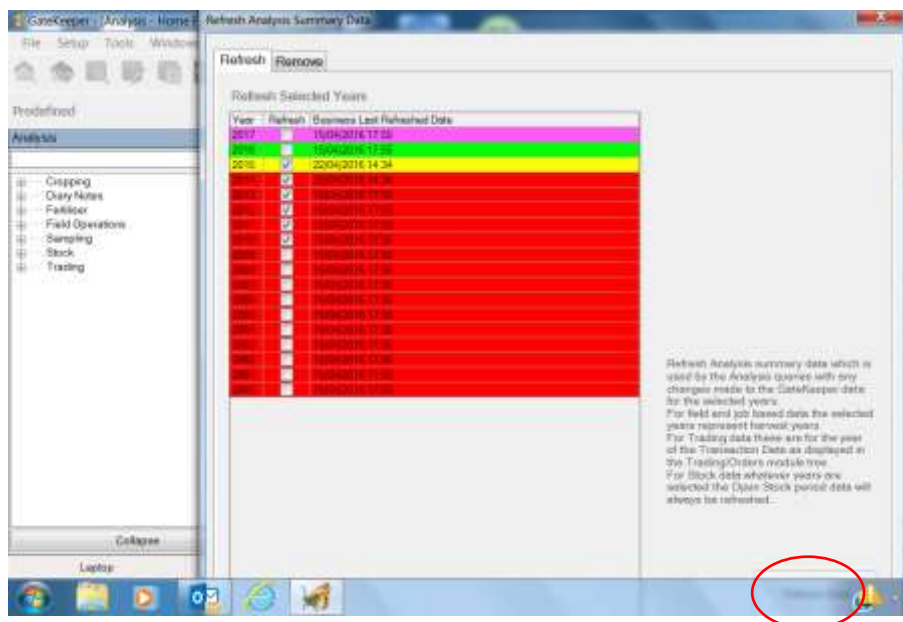
Step 1.4:

You can now close the 'Refresh Analysis Summary Data' window.

Step 2: Data preparation

Step 2.1:

You have now returned to the 'Analysis window'. Under the 'Analysis menu list' on the left hand side of the 'Analysis window' (highlighted below) select the 'analysis' item that contains the data of interest.

Step 2.2:

Click the cross icon to expand the 'Cropping' item. Under this list you will find and select 'Previous Cropping'. You then follow Step 2.3 below.
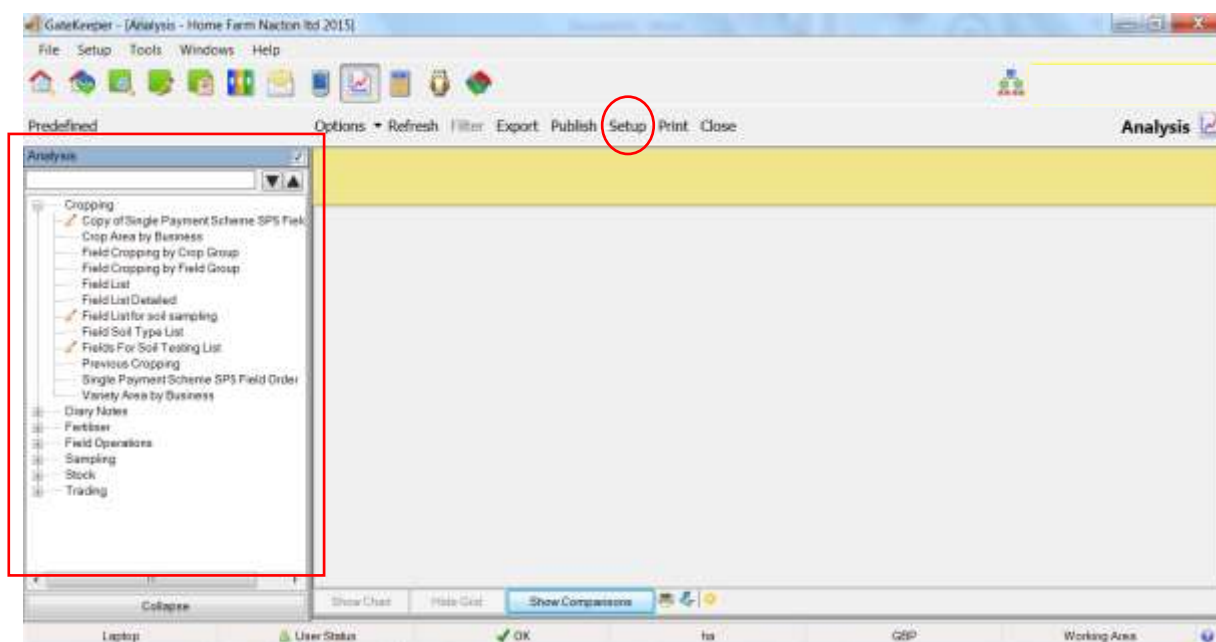
Analysis window and Setup tab

Step 2.3:

As the original item cannot be edited a copy of the item must be made. With 'Previous Cropping' selected click the 'Setup tab' (circled above) and create a copy. Once a copy has been made, click the 'OK' button. You will now see this listed on the left in the 'Analysis menu list' as 'copy of Previous Cropping'.

Step 2.4:

With the 'Copy of Previous Cropping' selected, click the 'Options tab' [*Note: Do not click the drop down arrow*]. You then need to select the following list of items, excluding all other items:

Actual/Issued Date	Heading	Product Name
Application Area ha	Heading Category	Quantity
Crop	Heading Group	Rate per Application Area ha
Crop Residue	Heading Type	Split Number
Crop Sequence	Map Sheet	Status
Descriptor	NG Number	Total Yield
Field Defined Name	Official Area ha	Units
Field Group	OS Area	Variety
Field Number	Parent Field Name	Year
Field Reference	Part Field Reference	Yield Units

Step 3: Data extraction

Step 3.1:

Select the 'Analysis menu list' item for extraction. For example, select 'Copy of Previous Cropping'. The pop up window below will appear.

3.2: Select the Years, Field Groups and Crops to be included.

 [*Note: To ensure that soil management can be evaluated in a full rotational context it is important to include as many years as possible. This also means that all crops are included such that a comprehensive range of rotational contexts can be captured. Depending on the processing capacity of the compute in use, it might be easier to extract the data in batches of only a few years at a time*]
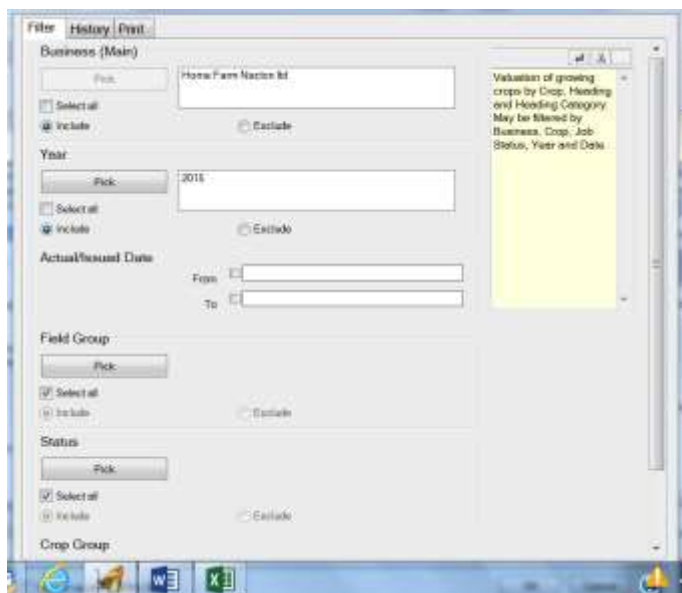
'Copy of Previous Cropping' pop up window

Step 3.3:

Once all the appropriate Years, Field Groups and Crops have been selected, click the 'OK button' at the bottom right corner. This will display the data available for export.

Note: If the 'OK button' is hidden by the task bar press the TAB button on the keyboard to select the 'OK button' and then press ENTER.

Step 3.4:

Click the 'Export tab' circled below and when prompted and save the file as an XML file.
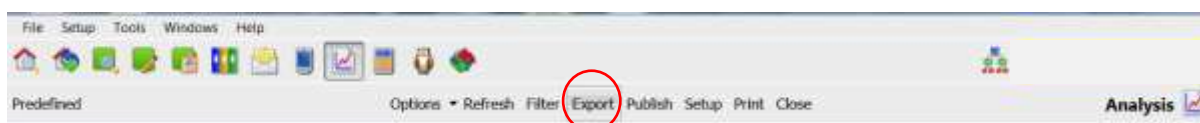
Step 4: Checking the exported data

Step 4.1:

Once saved, open the XML file from inside EXCEL to check the output. To do this first open EXCEL and call the XML file from the EXCEL open file option (see below)

Note: Double clicking the XML file will not work.
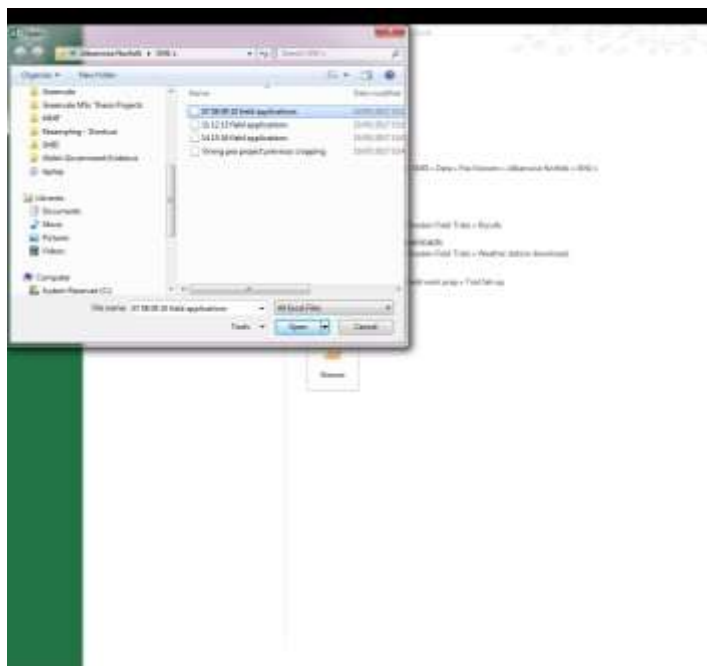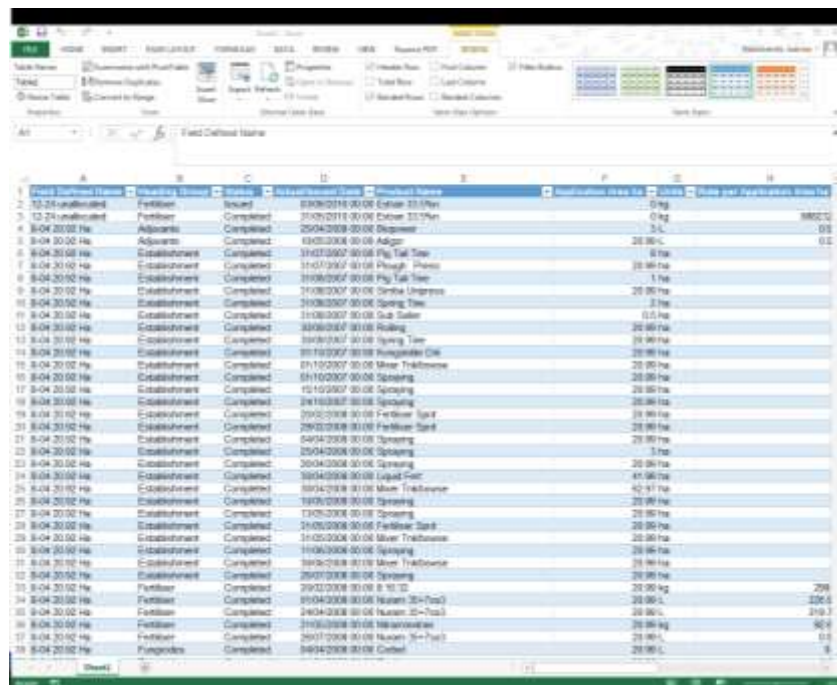
Step 4.2:

Select 'Open as an XML file' (see below).
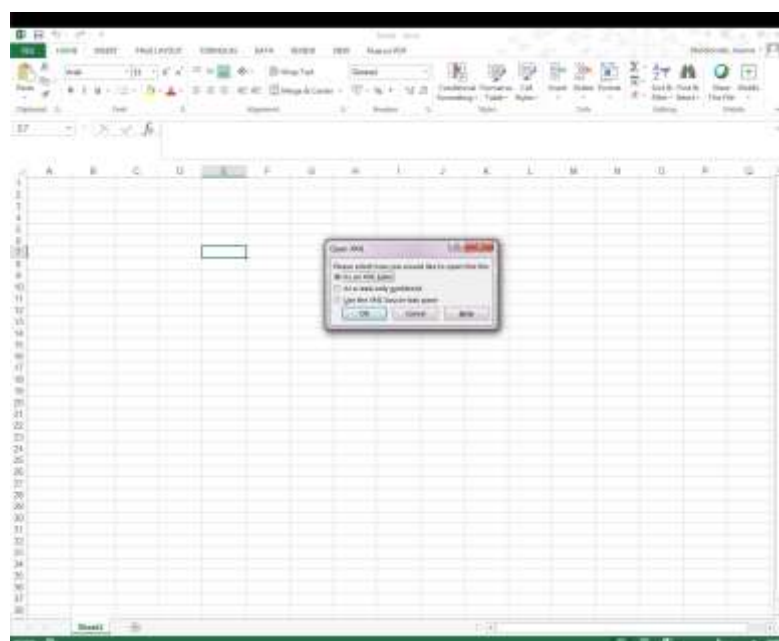
Step 4.3:

Save the file as a normal EXCEL workbook format.

Step 5: Preparing the remaining data

Step 5.1:

Return to the GateKeeper 'Analysis window'. Under the 'Analysis menu list' on the left hand side (highlighted below) of the 'Analysis window' select the 'analysis' item that contains the data of interest.

Step 5.2:

Click the cross icon to expand the 'Field Operations' item. Under this list you will find 'Field Applications list'.
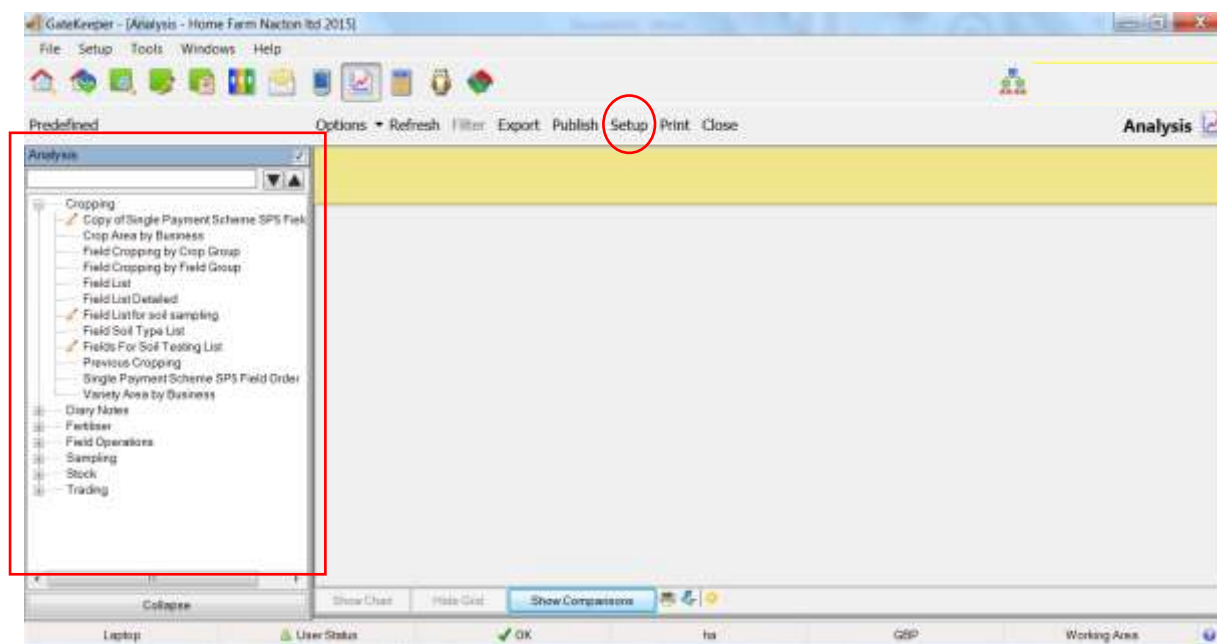
Step 5.3:

As the original item cannot be edited a copy of the item must be made. With 'Field applications list' selected click the 'Setup tab' (circled below) and create a copy. Once a copy has been made, click the 'OK' button. You will now see this listed on the left in the 'Analysis menu list' as 'Copy of Field applications list'.

Step 5.4:

With the 'copy of Field applications list' selected, click the 'Options tab' [*Note: Do not click the drop down arrow*]. You then need to select the following list of items, excluding all other items:

Actual/Issued Date	Heading	Product Name
Application Area ha	Heading Category	Quantity
Crop	Heading Group	Rate per Application Area ha
Crop Residue	Heading Type	Split Number
Crop Sequence	Map Sheet	Status
Descriptor	NG Number	Total Yield
Field Defined Name	Official Area ha	Units
Field Group	OS Area	Variety
Field Number	Parent Field Name	Year
Field Reference	Part Field Reference	Yield Units

Step 5.5:

You will now repeat Steps 3 to 4.