Project title:	Development of a Soil Management Information System (SMIS)
Project number:	CP107D
Project leader:	R J Rickson, Cranfield University
Report:	Annual (12 month) Report, December 2016
Previous report:	n/a
Key staff:	Steven Hallett, Rob Simmons, Ron Corstanje, Caroline Keay, Jason Carvalho, Joanna Niziolomski, Lynda Deeks
Location of project:	Cranfield University
Industry Representative:	Currently pending advice from AHDB. Previously Mr Guy Thallon (whilst Head of Research, Development and Innovation at Produce World)
Date project commenced:	01/11/15
Date project completed (or expected completion date):	31/10/18

DISCLAIMER

The results and conclusions in this report are based on an investigation conducted over a one-year period.

AUTHENTICATION

We declare that this work was done under our supervision according to the procedures described herein and that the report represents a true and accurate record of the results obtained.

[Name] R J Rickson

[Position] Chair of Soil Erosion and Conservation

[Organisation] Cranfield University

Signature .... Date ...19/12/16.............

[Name] R W Simmons

[Position] Senior Lecturer in Sustainable Soil Management

[Organisation] Cranfield University

Signature ....Date ...19/12/16....................

Report authorised by:

[Name] Professor L Terry

[Position] Head of AgriFood and Environment Theme

[Organisation] Cranfield University

Signature .... Date ....19/12/16..................

[Name]

[Position]

[Organisation]

Signature .... Date ....

CONTENTS

List of Figures

List of Tables

GROWER SUMMARY

1. Headline

Information and data on soil management issues and solutions are currently dispersed, but can be collated, harmonised and integrated in the Soil Management Information System (SMIS). Interrogation of SMIS generates outputs that give growers better soil management advice and guidance, to improve marketable crop yields and avoid soil degradation costs.

2. Background

Soil management is at the heart of sustainable intensification as it has the potential to improve crop yield (both quantity and quality), whilst protecting soil and water resources. In 2013, AHDB Horticulture commissioned a gap analysis of soil management research and knowledge transfer in horticulture to inform future research programmes (CP107). Incorporating growers' views and requirements, the final report identified a number of gaps in the research evidence, including the limitations of the experimental empirical base and the need for 'big data' approaches, especially given the unprecedented amount of data being generated by growers through on-farm data management software such as 'Gatekeeper' and 'Muddy Boots'. It was recognised that sources of information and data related to soil management in horticulture are currently unstructured, uncentralised and difficult to find and/or access. Growers tend to hold and record data in three categories: (1) paper records; (2) disparate, customised electronic formats, and (3) data held in formal farm records systems (e.g. Gatekeeper). According to the Gatekeeper website (https://farmplan.co.uk/crops/gatekeeper-grower/), there is "*frequently no real sense to the purpose of collecting it [data] and the ultimate aim. This is becoming very much the case in the agricultural industry today, with most growers amassing fairly sizeable sets of data on a whole variety of subject areas, <u>but how much is actually done with it</u>?*". As a result, it was recommended that future research should develop a soil management information system (SMIS) that will hold, manipulate and manage such data in a way that can be interrogated to provide advice and guidance on the benefits of soil management practices, with regard to crop productivity and environmental protection. This should include relationships between soil management practices and field and farm-level outcomes (e.g. economic costs and benefits; and environmental impacts). Development of the Soil Management Information System (SMIS) will meet these demands.

3. Summary

Currently, the horticultural sector lacks access to coherent information that will support decisions on sustainable soil management and horticultural production, and the analytical approaches to determine differences between productive and unproductive sites. This is turn severely compromises effective knowledge exchange (KE). The purpose of SMIS is to gather and collate diverse sources of information and data on soil management issues and their mitigation (irrespective of format or source) and harmonise it so that the overall database so created can be interrogated with data mining techniques (i.e. the SMIS Analytical Toolkit) to establish patterns in the data. This 'rule base' establishes the relationships between the intrinsic factors that affect soil condition (e.g. site conditions, soil type, crop, weather, timing of operations, machinery) and best management practice for soil management operations, together with anticipated outcomes of given actions (for all SMIS data sources).

SMIS will give growers, agronomists and land managers access to guidance and information on evidence-based, optimal soil management practices in horticulture. Critically, as the size and number of records in the database increases, SMIS can learn and reincorporate new data as it is introduced in the form of updated probabilities and likely outcomes. Knowledge which is not present in the form of data (expert or grower experiences; outcomes from the literature or other experimental results) can also be encoded and uploaded into SMIS, where it can be interrogated to find the best management practices for the given field / crop scenario.

Year 1 of the project has focussed on the collation of data, information and knowledge on soil management issues and their solutions from a wide range of sources, including farmer/grower data (e.g. as held in Gatekeeper farm management software); other relevant datasets (e.g. Met Office data; LandIS Soil Series data); research outcomes from on-going AHDB-Horticulture funded projects (i.e. CP107b and CP107c); expert knowledge from researchers and growers; and an extensive review of the literature (both scientific papers, research reports and 'grey' articles). Harmonisation and integration of these varied sources of data is required. A method to do this has been proposed and will be tested as an important component of Year 2 activities. The stages of building the SMIS and developing the eGuide are demonstrated using a case study of soil compaction in vining peas (Figure 1) and are explained in the Science section of this report.

The outputs of the project will be of direct benefit to levy payers because:

a) For the first time, information on soil management advice and guidance in horticulture will be available as one centralised resource;
b) SMIS outputs will provide advice on improving soil health (with associated improvements in crop production); and

c) SMIS outputs will provide advice on avoiding soil degradation and associated costs.

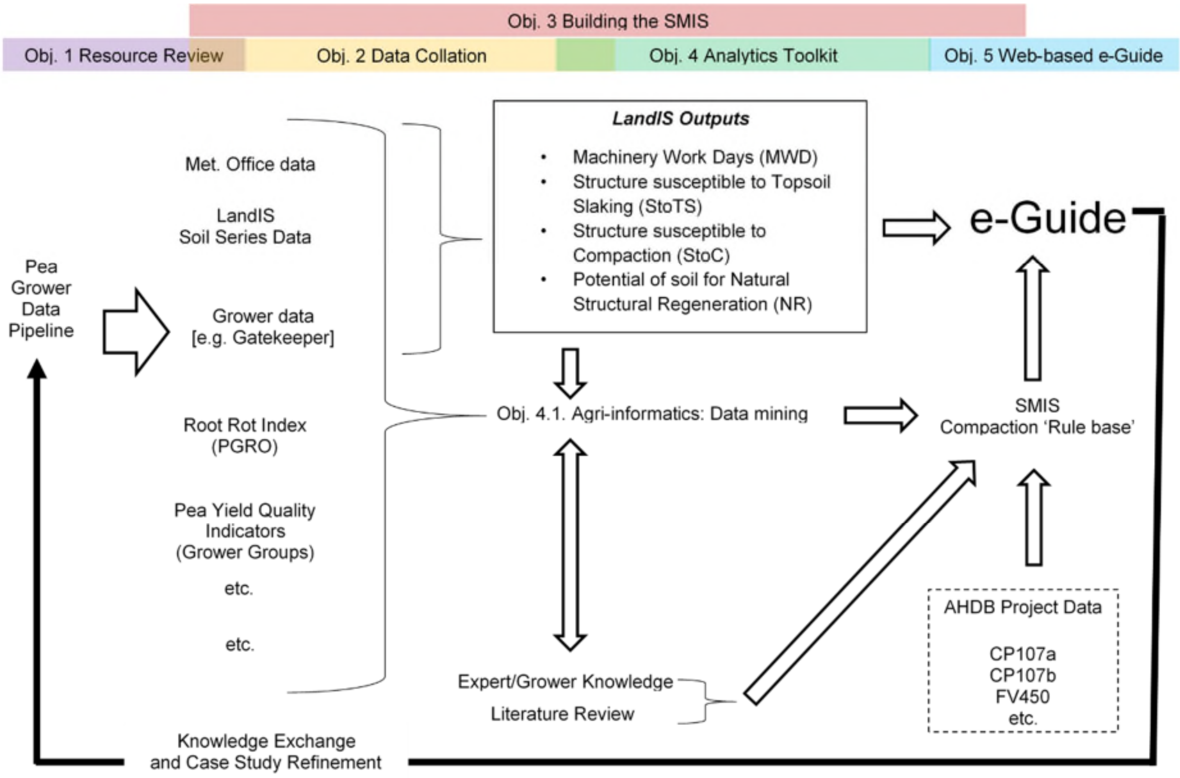

Figure 1. Stages in developing a Soil Management Information System and eGuide (demonstrated with a case study of soil compaction in vining peas).

4. Financial Benefits

The soil management advice and guidance given by the SMIS will bring financial benefits for levy payers in two ways. First, SMIS will help mitigate costs incurred by growers from soil degradation. Soil degradation affects yield and yield quality directly, as does the timing of tillage, planting and harvesting operations. Poor soil quality (e.g. compacted soil) leads to gaps in production continuity and critically to pinch points in product delivery. Such continuity gaps can exert significant financial impact on growers and increase the reliance on imports to meet customer requirements and to maintain national food security. Conservative estimates of the impacts of soil degradation on agricultural production are estimated at £212-270 million per annum (Graves et al., 2011; 2015). Costs to individual farmers/growers may include reseeding operations, subsoiling to alleviate compaction, relevelling land subject to erosion, fines incurred due to breaches of the Water Framework Directive (eroded soil in watercourses) or from the Highways Agency (mud on roads), additions of organic amendments, and poor yields. Loss of customer confidence due to the difficulty of delivering to time and specification can also have a significant longer-term impact on farm income.

Second, SMIS will provide more effective advice and guidance on soil management leading to improved soil health and system resilience. As well as increased outputs (yield quantity and quality; Table 1), well managed soils have lower input requirements (nutrients, water, agrochemicals), giving better financial margins in the short term and better soil quality / health in the long term.

Table 1. Increase in yields of crops grown in horticultural rotations due to improved soil and water management (2015 prices)

Crop	Yield increase associated with better soil health	Financial benefits to individual growers
Wheat	up to and over 10%	10% increase in yield would result in 1.2 t/ha increase @ £130/t
Potatoes	5%	based on 15,000t produced = 750t extra – contract price £165 /t = £123,750 income
Maize	5%	Improved yield means less land required. If 40 ha of land under maize @ growing cost per ha of £1550k = saving of £65,000. The 40 ha could be put to wheat = 528 tonnes = £68k income.
Lettuce	1.5%	Improved yields mean 1.5 million fewer heads per yr needed = 15 ha less land @ growing cost per ha of £8k = saving of £120,000
Onions	2.5%	based on 5000t produced = 125t extra yield – contract price £190 per ton = £23,750 income

5. Action Points

- To achieve the benefits of improved soil management advice and guidance, (e.g. enhancing crop production and environmental protection), it is necessary to provide data and information to the SMIS data repository. We will continue to source additional inputs to SMIS by working closely with our industrial partners, project collaborators (e.g. PGRO), other researchers (especially those undertaking CP107 projects) and developing additional contacts in the horticultural sector.
- Inputs will be combined with other datasets such as Met Office data and LandIS Soil Series data to present specific soil management issues faced by growers, such as soil compaction.
- Issues of data integrity, reliability and accuracy will be addressed as new sources are incorporated into SMIS. This will include how to manage missing data – is it possible to use proxy data instead? For example, soil bulk density measurements are not always available (as an indicator of soil compaction). However, knowing the soil type, weather conditions and machinery used, the risk of soil compaction can be estimated.

- The large datasets within SMIS will require the use of complex data management techniques and advanced computational skills. We will continue to develop analytical methods and statistical modelling, drawing across the body of data assembled, allowing comparative assessment and benchmarking against available grower and case-study data. Statistical interpretation of grower datasets within SMIS will provide a more scientific basis for guidance on a wide range of soil management issues.
- A rules-base for functional relationships between data members will be established based on expert opinion, established AHDB guidance documents and weight of evidence in the literature. This will form the basis of a suite of expert knowledge and hypothesis driven statistical analyses.
- As the database within SMIS develops, it is envisaged that data mining techniques will provide useful insights to address AHDB Horticulture Panels' 2015-2018 priorities.
- We will continue to promote SMIS in project knowledge exchange activities.

SCIENCE SECTION

1. Introduction

The Horticultural Soil Management Information System (SMIS) Project (CP107d) will result in the development of a Soil Management Information System, which will hold, manipulate and represent a wide range of available sources of data and information pertaining to the specific effects of soil management practices on horticultural crop productivity and environmental protection. The key output of the project will be knowledge exchange through an e-Guide toolkit, able to interrogate SMIS to provide AHDB users with a set of robust, empirically-based, best-practice guidelines, and the likely consequences of applying them on crop production and environmental protection. As a result, the overall outcome of the project will be better managed, healthier soils that underpin sustainable economic production, whilst being resilient in the face of external pressures such as extreme weather events.

The stages of developing the SMIS are shown in Figure 2 and are reflected in the project deliverables (Appendix 1). To date, the Resource Review has been completed, with additional knowledge and data being inputted to SMIS throughout the project duration (and beyond). The other tasks (Data handling and Validation, and Analysis Toolkit) are also underway, as described in the report section below. These activities will lead to the creation of the eGuide soil management toolkit.

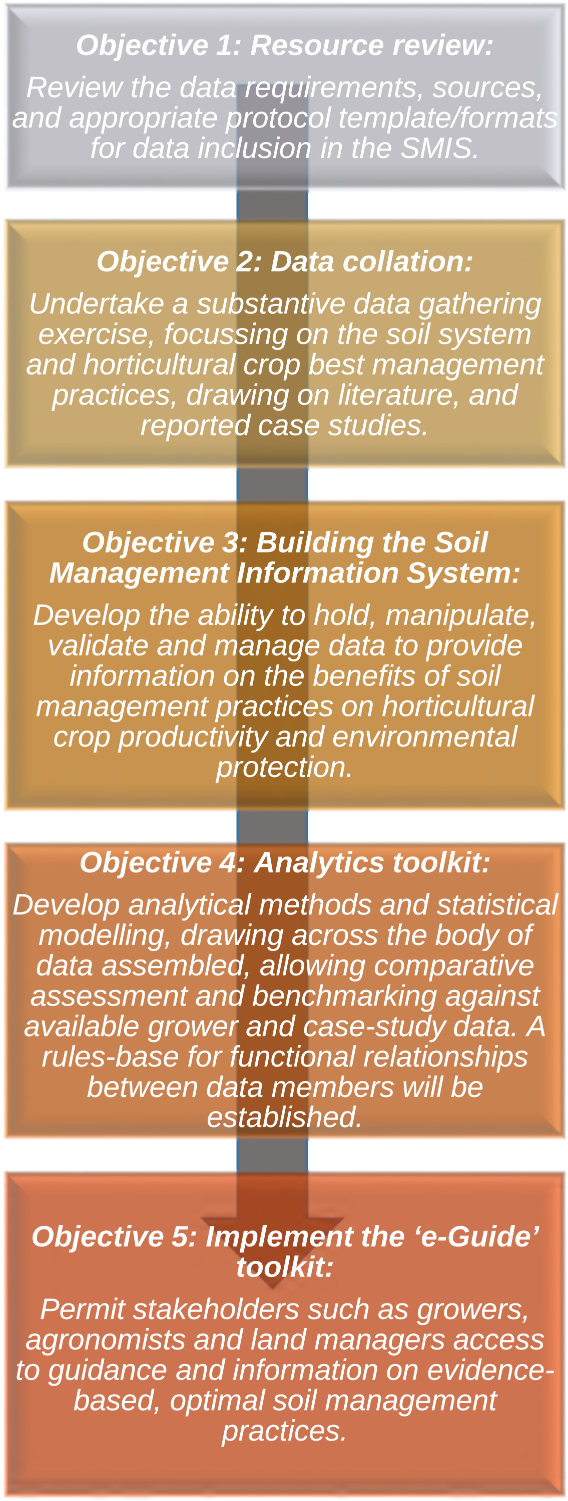

Figure 2. Stages of developing the SMIS and eGuide

The structure of this report reflects the project objectives and activities to date. Section 2 describes the approach taken to build the SMIS. Section 3 – 6 describe the varied data sources on which the SMIS will be based. Section 7 presents how the data sources can be harmonised and the current architecture of SMIS. Sections 8 and 9 discuss the project outputs to date. Section 10 covers the project's knowledge exchange activities to date. Sections 11 and 12 are a project glossary and reference list respectively.

2. Materials and methods

This section provides an overview of the methods used to develop and assemble the SMIS. This includes identifying a) what farmers and growers need in a soil management information system (Requirements Analysis); and b) the sources of data, information and knowledge of soil management issues, practices and their outcomes that will form the basis of the SMIS.

2.1. Requirements Analysis

In 2012, AHDB Horticulture commissioned a gap analysis of soil management research and knowledge transfer in horticulture to inform future research programmes (CP107a; Rickson and Deeks, 2013). This included a questionnaire of farmers and growers, aimed at gathering individuals' perceptions of the key soil management issues in horticulture (and rotations that include horticultural crops). The responses were used in the current project to identify the key soil management issues faced by respondents, and the soil management practice options used to address them. This scoping exercise identified a number of intrinsic (i.e. site) factors and extrinsic (i.e. management options) factors related to soil management issues and solutions (Appendix 2). This preliminary exercise identified the key inputs and outputs of soil management practices in horticulture that have to be captured in the SMIS and eGuide.

2.2. Sources of data, information and knowledge

The SMIS is an innovative 'knowledge repository', designed to hold and manage linkages between the diverse and currently disparate sources of information on soil management issues, practices and outcomes. The SMIS repository provides the foundation for the development of an 'e-Guide' toolkit – designed to provide AHDB-Horticulture (and the growers, agronomists and land managers it advises) access to guidance on optimal soil management practices.

The four SMIS 'Data / Information / Knowledge Sources' capture soil management issues, practices and their outcomes in widely differing formats, as described in Table 2. As such, the data is '*unstructured*'[1]. Collating, harmonising and managing this 'heterogeneity' requires

[1] '*Unstructured*', in that there is not a single comprehensive database design that encompasses all facets of all of the data.

innovative approaches, as described in Section 7.1 below. First, each source is presented separately in the sections below.

Table 2. Data / information / knowledge sources for SMIS

1. Farm/Grower Records	Growers tend to hold and record quantitative dataset records in 3 ways: (1) paper records; (2) disparate, customised electronic formats, and (3) data held in formal farm records systems (e.g. Gatekeeper)
2. Research project outputs (e.g. CP107b and c outputs)	AHDB-Horticulture funded research outputs, which may have either (a) quantitative data or (b) qualitative information. A stakeholder form has been developed to help source this data, see Appendix 8)
3. Literature	Scientific and grey literature concerning soil management issues, practices and outcomes
4. Knowledge	The expert and informed opinion of leading practitioners in the field. This knowledge of soil management issues, practices and outcomes is captured in the form of Fuzzy Logic Cognitive Models (FCM).

The SMIS will permit users to navigate between these diverse sources of information and extract guidance on best soil management practices for given site conditions using the eGuide toolkit.

3. Data / information / knowledge source: Farm/grower records systems

Farmers and growers are collecting significant amounts of data from their field and farm operations. Much of this is pertinent to soil management challenges such as soil compaction, which can directly impact on crop yield, financial returns and environmental quality. Data is also available on the practices that have been used to prevent or remediate these issues. The purpose of SMIS is to gather and collate this data (irrespective of format or source in the first instance) and harmonise it so that the overall database so created can be interrogated with data mining techniques (i.e. the SMIS Analytical Toolkit) to find best soil management outcomes, given the field conditions at any given site (e.g. soil type, crop, timing of operations, machinery). At this stage of the project, we have tested and validated this approach with extensive data records from growers who use a market leading farm software product, the Gatekeeper farm records system (https://farmplan.co.uk/crops/).The intention is that other data sources will be incorporated into the SMIS database in the same manner during the project and beyond.

3.1. Intrinsic value of Gatekeeper datasets in SMIS

To demonstrate how farm records can be used as input to the SMIS, a case study of soil management issues in vining peas is used as an example, starting with the initial consultations with pea growers, through to the data mining techniques used on the assembled data sets. The latter will extract causal relationships between site properties, soil management practices and outcomes in terms of yields and possible risks of soil degradation (here, soil compaction). The data available is in a format that incorporates a whole farm, cross-rotational context. This means the SMIS approach could be relevant to other crops in the pea rotation, so having potential application to multiple AHDB (Cereals and Oilseeds, Potatoes and Horticulture) and non-AHDB (MGA, BBRO) sectors.

Following consultations with CP107d project partner Becky Ward at PGRO, two Pea Growers Groups were identified as potential data sources for SMIS. Consequently, meetings were held with Richard Fitzpatrick of HMC Peas (10th February 2016) and Ian Watson of Stemgold (29th February 2016). HMC Peas are a grower group consisting of 31 growers (90% are on Gatekeeper) who on average grow 1,740 ha of peas per annum. Stemgold consists of 21 growers who cultivate circa 1,620 ha per annum, aiming to achieve production of 6000 t of frozen and 1000 t of canned peas.

The objectives of these meetings were to

- Introduce the aims and objectives of the AHDB-Horticulture SMIS Project
- Elicit detailed expert knowledge and experiential evidence regarding the key soil management challenges of their grower groups and soil management options that have been adopted to address these challenges (with a specific focus on peas).
- Identify individual growers who would be willing to provide data to SMIS.

As a result, key soil management issues on the silts and light sands associated with the HMC and Stemgold growers were identified as:

- Soil workability/trafficability issues associated with optimising drilling and seedling development/vigour whilst minimising compaction risk
- Undertaking sub-soiling operations at appropriate soil moisture conditions to optimise sub-soiling efficiency
- Experiential evidence of a link between soil compaction and Foot Rot
- Topsoil slaking/slumping and capping prior to and post emergence
- Avoiding growing peas following late harvest crops such as sugarbeet and maize to avoid loss in yield associated with legacy-compaction.
- Key soil borne diseases, including Pea Cyst Nematode (Pea^{CN}) caused by *Heterodera gottingiana* and Foot Rot (*Fusarium solani* f. sp. pisi, *Phoma medicaginis* var. *pinodella*).
- The need to extend rotations to 8 years to avoid build-up of key soil borne diseases impacting on yields (current pea rotation is 5-6 years).

This expert knowledge and experiential evidence of key soil management challenges and management options will be used to develop the SMIS 'rule base' and will form the basis of key questions to be addressed in the data mining component of the Analysis Toolkit (Figure 2; Section 3.3.4). As a consequence of these meetings, whole farm, full rotation Gatekeeper datasets (2010-2015) have been obtained to date from four HMC pea growers namely Worth Farms, Caley Farms, Jack Buck Farms and Hay Farming (circa 3,000 ha in total). Of these, the Worth Farms dataset (circa 1,200 ha) is the most coherent, and as such has been used as the initial dataset to demonstrate the feasibility of using grower Gatekeeper data for incorporation into SMIS. Table 3 demonstrates that the Worth Farms Gatekeeper records include a comprehensive suite of data identified as pertinent in the SMIS Requirements Analysis (section 2.1; Appendix 2).

The remaining datasets from the other farms will be incorporated into SMIS in early Year 2 (2017) in accordance with agreed project milestones. Additional datasets will be brought in from other Pea Grower Groups and from other horticultural sectors (including the project's industrial partners) to support the grower data pipeline for SMIS (See Section 8.1, Next Steps).

Table 3. Worth Farms Gatekeeper derived data (2010-15).

SMIS Requirements Analysis	Data	Description	Comments
Intrinsic site factors/properties	Location	National Grid (NG) Reference Codes, Farmer field name/code, Area (ha) etc.	Used to derived Unique field identifier onto which all other data will be assigned. NG Codes used to link to Rural Payments Agency (RPA) field boundary shape file NG Codes used to link to Cranfield University's Land Information System, (LandIS) Soil Series data and derived soil attributes (LandIS http://www.landis.org.uk/)
	Year/data	Cropping year as well as date (dd/mm/yr) of all operations/activities undertaken at a field scale	Used to derive cropping histories, timing of operations (particularly tillage and harvesting operations) Timing of tillage and harvesting operations linked to LandIS derived 'Workability Windows' to calculate 'compaction risk'
	Crops	Calabrese, cauliflowers, forage maize, mustard seed, peas vining, potatoes (main crop), winter rye, salads, sugar beet, winter wheat	***Note: Crops grown also link to multiple AHDB Sectors*** • Forage Maize = Maize Growers Association (MGA) • Winter Rye and Wheat = AHDB Cereals & Oilseeds • Potatoes = AHDB Potatoes • Calabrese, Cauliflowers, Vining Peas and Salads = AHDB Horticulture • Sugar Beet = British Beet Research Organization (BBRO)
	Seeding/Plants*	Crop, variety, rates, timing of drilling	Information on varieties grown as well as seeding rates can be linked to yield as related to all other attributes recorded.
	Grower Soil Testing Data	B, K, Mg, Na, P, PCN (Cysts 100 g soil, Eggs g soil, Total Cysts per 100g soil), pH, S.	Field specific analyses of a number of key biological and chemical soil quality indicators. Note: Physical soil quality indicators not measured

	Rotation context	Previous cropping history (up to 5yrs) and associated field operations	Can be used to assess impact of previous cropping history and field operations on current season / year's crop.
Extrinsic factor/property	Machinery* Operations	Tillage equipment used during establishment, harvest, spray and fertiliser application operations	Can be used to timeliness of operations and type of tillage e.g. deep (sub-soiling), shallow, rotational or inversion
	Fertiliser* Inputs	B, CaO, Co, Cu, Fe, K_2O, Manure (NH4-N, K_2O, MgO, N, P_2O_5, SO_3), MgO, Mn, Mo, N, Na_2O, P_2O_5, S (elemental), SO_3 and Zn	Timing, and rates applied
	Pesticides*	Various Herbicides, Insecticides and Fungicides	Crop specific timing of application, frequency and rates applied. Can be used as an indicator of disease incidence
	Cover Cops	Mustard used prior to Salads and Potatoes (Main crop) and Oats used prior to Peas Vining, (limited fields for 2014 and 2015)	Timing of cover cropping windows and associated crops
	Organic Amendments	AD_Solid Waste applied on selected fields (2014 and 2015)	Timing, and rates applied
Outcomes	Yields	Crops include Calabrese, Cauliflowers, Forage Maize, Mustard Seed, Peas Vining, Potatoes (Main crop), Winter Rye, Salads, Sugar beet, Winter wheat	Yield (t ha^{-1}). Used as basis on which to access soil management practices adopted.

* Timing of fertiliser application and spray operations can be used to identify soil compaction risk if operation undertaken outside of 'Workability' windows.

In addition, two further farm based datasets dealing with Foot Rot Index and Yield Quality (Table 4) have been successfully combined with the Worth Farms Gatekeeper data (and will be included as data variables in future data mining activities (part of the Analysis Toolkit)).

Table 4. Foot Rot Index and Pea Yield Quality datasets

SMIS Requirements Analysis	Data	Description	Comments
Intrinsic site factors/properties	[1]Foot Rot Index	Foot Rot Index for Worth Farms Fields (2009-2015)	Agri-informatics approaches to be used to identify factors affecting Foot Rot Index
Outcomes	[2]Pea yield and yield quality	Yield (t ha^{-1}), Tenderness Rating (TR), field yield, frozen yield and wastage	Agri-informatics approaches to be used to identify factors affecting yield quality attributes

[1]PGRO: Data input by Mark White of PGRO; [2]HMC Peas.

3.2. Gatekeeper data integrity checking

Worth Farms data was exported from Gatekeeper as a series of XML files. Data quality and coherence was assessed prior to individual data files being combined to a 'flatfile' format required for subsequent datamining (Section 3.1.3). Key data quality / coherence issues and solutions adopted are listed in Table 5 below. The integrity of Caley Farms, Jack Buck Farms and Hay Farming data is also under investigation, and missing data fields have been requested.

Table 5 Managing Gatekeeper data integrity

Problem	***Solution***
No consistent identifier of field across all the tables. Between 2-7% of NGC code or (Mapsheet & NG Number)	Unique ID created for fields, and add to all data tables
Unable to link to FieldUniq in Fixed Costs, as NGC code blank for 4% of data	Add field number to All Fields tables, based on Yields table
Crop Names not consistent across all tables	Crop names updated to ensure consistency across data files.
80% of Fixed costs data also in Variable Costs table	Drop Fixed costs table
Fields in Yields table contain duplicate yield data.	Duplicate fields removed

Several fields were consistently split between 2010 and 2015 and so were identified separately	The 'field defined name', which has split field details, was re-defined to link to relevant data fields enabling comparison with crops (and all associated data fields) grown in previous years.

3.3. Utility of Gatekeeper data: Compaction Case Study

It is important to bear in mind the intended outputs of the SMIS, when managing the data sources used to 'populate' the system. Both Richard Fitzpatrick (HMC Peas) and Ian Watson (Stemgold) identified soil compaction as a key soil management challenge. This is in line with the findings of other growers surveyed in the CP107a Gap Analysis Report. Further, the observed decline in soil structure means that pea growers within the HMC Peas and Stemgold grower groups are routinely turning to mechanical approaches such as sub-soiling to artificially generate 'soil structure' (Personal Communication: Simon Day, Worth Farms; Richard Fitzpatrick HMC Peas).

Compaction: Causes and Consequences

Primary causes of soil compaction in agricultural systems include the trafficking of the land by vehicles, and the tilling of soil at inappropriate times and/or repeat tillage at the same depth (Batey, 2009). Soil compaction, is estimated to effect >30 M ha of agricultural land in Europe (van dan Akker & Canarache 2001). There is an increasing trend in the degradation of soil structure arising from highly intensive agricultural systems involving excessive and inappropriately timed trafficking, increased machinery mass and loss of soil C (Ball et al. 1997; Chamen 2006; Batey 2009), compounded by an increased frequency of extreme weather events (IPCC 2007; Gornall et al. 2010).

This systematic degradation of soil structure can severely restrict root development (Clark et al. 2003; Whalley et al. 2006; Grzesiak et al. 2013) and compromise the ability of crop plants to access water (White & Kirkegaard, 2010) and nutrients (Seymour et al. 2012), increase susceptibility to disease and pest damage with direct impacts on yield, yield quality and production costs. The susceptibility of the soil to compaction (termed workability or trafficability) depends on the interaction between climate and soil physical properties (Batey, 2009). Wet soils are more susceptible than dryer soils to compaction. Soils at field capacity (FC) should not be trafficked or worked as they are at greatest risk of compaction. The clay content of soil influences workability and trafficability, in relation to its ability to retain water, with increasing clay content reducing the number of workable days in a year.

Monitoring Compaction: Missing Metrics

However, monitoring the level of compaction is not easy. Methods of determining compaction include measuring bulk density (Batey, 2009), total porosity, air-filled porosity (Ball et al., 1997), penetration resistance (Duiker, 2002; Batey, 2009) and visual soil assessment (Askari et al., 2013; Batey, 2009; Mueller et al., 2009; Shepherd, 2003).

Further, these properties do not just vary with degree of compaction but also in relation to soil texture (Archer and Smith, 1972) and in the case of penetrometer readings in relation to soil water content. *Consequently, growers do not routinely measure soil compaction*.

Utilising Gatekeeper data to determine soil compaction risk

However, data relating to timing and type of machinery operations (as accessed in Gatekeeper records) can be used as indicators of compaction risk.

The following section demonstrates how Gatekeeper data can be combined with other data sources within SMIS to provide an assessment of soil compaction and offer solutions in terms of best management practices.

3.3.1. Use of the SMIS to assess soil compaction risk: Worth Farms Case Study

To facilitate an assessment of compaction risk, two external datasets (from LandIS and the Meteorological Office) relating to soil and climatic properties (Table 6) have been linked to the Worth Farms Gatekeeper data.

Table 6. Linking to LandIS and climatic datasets

SMIS Requirements Analysis	Data	Description	Comments
Intrinsic site factors/properties	[1]LandIS	LandIS HORIZONfundamentals database Standardised soil profile data giving texture (sand. Silt and clay %w/w), pH and carbon, calcareous status	Used as input parameters to calculate soil Wetness Class required to determine • Machinery Work Days (MWD) • ‘Structure susceptible to Topsoil Slaking (StoTS) • Structure susceptible to Compaction (StoC) • Potential of soil for Natural Structural Regeneration (NR)

	[2]Climate	Rainfall and temperature data for Eastern England	Used as input parameters to MWD calculations and compaction risk

Source: [1]LandIS http://www.landis.org.uk/ ; [2]Met Office East of England historical data 1940-2015.

Compaction risk is based on 'Good' Machinery Work Days (MWDs). This workability scheme devised by Thomasson (1982) combines Field Capacity (FC) soil moisture data with soil structure, permeability and soil water regime (wetness classification) to predict the periods in spring (1 January to 30 April) and autumn (1 September to 31 December) suitable for land work (i.e. tillage, harvesting and trafficking associated with field operations). Cultivation of well drained coarse textured soils is often possible within the general FC period without harmful effects on soil structure (i.e. compaction), whereas under the same climate, clayey or otherwise slowly permeable soils (such as light silts or soils associated with impeded drainage due to the presence of a plough-pan) are usually impassable and remain too wet for longer than the FC period.

In organising the data, first the Worth Farms fields' National Grid Codes derived from the Gatekeeper dataset were used to locate the farm and delineate the RPA field boundaries for each field. Subsequently, the 1:50,000 'Boston and Spalding' digitised Map Sheet was overlain in Arc-GIS to identify the Soil Series associated with the Worth Farms landbank (Figure 3). These soil series could be arranged into two distinct groups: light silty marine alluviums (Romney, Rockcliffe and Wisbech) and medium silty marine alluviums (Agney, Tanvats, Stockwith).

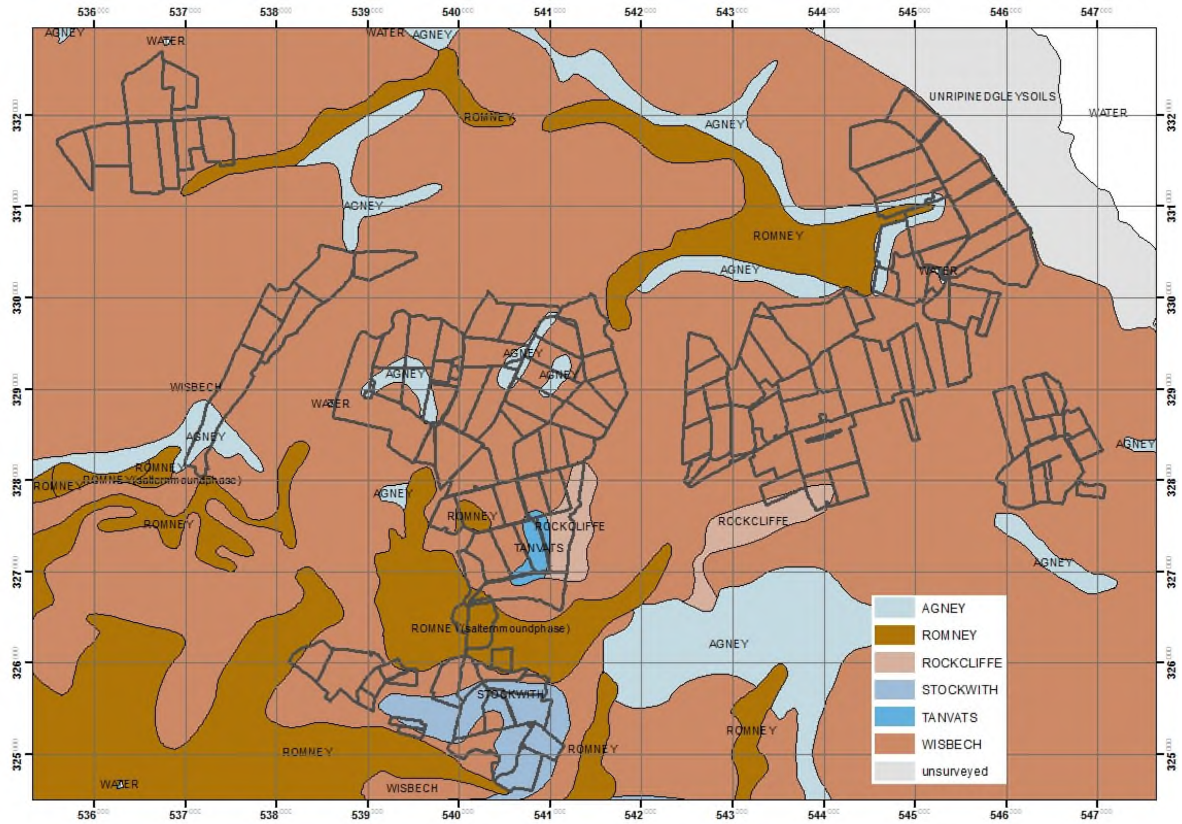

Figure 3. Soil Series associated with the Worth Farms land bank

Consequently, soil properties from LandIS were used to derive the Wetness Class and the MWD for the Worth Farms soils (Table 7). Figure 4 visually represents the windows for 'safe' access to land (e.g. when soil compaction risk is low) for the dominant Wisbech Soil Association under 'normal' and 'wet' conditions. It is envisaged that a similar visualisation will form a component of the proposed SMIS e-Guide.

Soil series	Soil assess-ment	Type of Year	M.W.D's	AUTUMN			WINTER			SPRING		M.W.D's
				SEP	OCT	NOV	DEC	JAN	FEB	MAR	APR	
Wisbech, Rockcliffe & Romney	a	Normal	122									43
		Wet	104									27

M.W.D's : Number of good machinery work days during the period indicated

Frequent opportunities for Autmn landwork

Frequent opportunities for Spring landwork

Little opportunity for landwork

Figure 4. Workability Days (MWD) for Wisbech Soil Association (Un-adjusted).

Table 7. Period of Field Capacity for Worth Farms Soils as derived from National Soils Map and LandIS Soil Series data.

			Return date of Field Capacity*/**		End date of Field Capacity*/**	
Year and Soil Type	**WA**	**WC**	**Min**	**Max**	**Min**	**Max**
2010 (dry spring/wet autumn)						
light silty	a	I	30-Jan	03-Feb	28-Mar	31-Mar
medium silty	c	III	11-Dec	15-Dec	23-Apr	24-Apr
2011 (dry spring & autumn)						
light silty	a	I	20-Dec	23-Dec	28-Mar	31-Mar
medium silty	c	III	01-Nov	03-Nov	23-Apr	24-Apr
2012 (wet spring & autumn)						
light silty	a	I	30-Jan	03-Feb	22-Feb	23-Feb
medium silty	c	III	11-Dec	15-Dec	19-Mar	20-Mar
2013 (median spring & autumn)						
light silty	a	I	09-Jan	12-Jan	08-Mar	10-Mar
medium silty	c	III	20-Nov	23-Nov	03-Apr	04-Apr
2014 (wet spring/median autumn)						
light silty	a	I	09-Jan	12-Jan	22-Feb	23-Feb
medium silty	c	III	20-Nov	23-Nov	19-Mar	20-Mar
2015 (dry spring/autumn)						
light silty	a	I	09-Jan	12-Jan	08-Mar	10-Mar
medium silty	c	III	20-Nov	23-Nov	03-Apr	04-Apr

WA = Series based assessment of workability ranging from *aa* to *f* where *aa* is used for very coarse well drained soils and *f* is used for extremely wet, heavy or peaty soils. (Thomasson 1982)
WC = Soil water regime described by the system of wetness classes grading from Wetness Class I, well drained to Wetness Class VI, almost permanently waterlogged within 40 cm depth (Hodgson, 1976)
*Based on Met Office Climate Data (1940-1970)
**Adjusted based on whether for the years 2010-2015 the autumn and spring periods were associated with rainfall considered to fall within the Dry (<25th Percentile), Median (25th -75th Percentile) or Wet Quartile (75th Percentile) relative to the 1940-1970 average.

3.3.2. Dataset derivation and classification of compaction risk and natural compaction regeneration potential

Utilising the Worth Farms Gatekeeper data and LandIS data can provide current best estimates for three elements of soil structural degradation and remediation, i.e. topsoil slaking, compaction and natural structural regeneration. In each case, a 4-tier ranking of structural stability for cultivated soils is given (Table 8):

- Structure susceptible to Topsoil Slaking, (StoTS) classes S1 to S4: very unstable, unstable, moderately stable and stable;
- Structure susceptible to Compaction (StoC) classes C1 to C4: very susceptible, moderately susceptible, slightly susceptible and very slightly susceptible;
- Potential of soil for Natural Structural Regeneration (NR) classes R1 to R4: little potential, Slight potential – likely to take at least 5 years; Moderate potential for natural regeneration but could take up to 5 years; and Large potential for natural recovery over 1 season (autumn to spring).

Further research undertaken during the course of this project (linking to the Literature Review (Section 6) and interrogation of additional Pea Grower datasets) will inevitably improve the understanding of these processes and hence the interpretations for individual Soil Series will be improved.

For all Soil Series for Worth Farms, an initial assessment was made of the following factors:

- Wetness Class – this can change according to climatic differences across the country and the 'Field Capacity' dataset, supplied as part of the workability/trafficability assessment has been used to apply the most commonly applicable Wetness Class value to each soil series (see Table 7);
- Soil texture – has been derived from the LandIS Soil Series description;
- Soil organic carbon values – extracted from the LandIS HORIZON fundamentals dataset for the 410 Soil Series occurring on the National Soil Map (NATMAP); for the remainder, mean values were taken from the National Soil Inventory (NSI) dataset or by using expert judgment.

For some Soil Series (notably calcareous and iron-enriched (ferruginous) soils), reference has been made to the Soil Series definition to help identify these factors.

Table 8. Applying soil structural stability classes to soils on Worth Farms for soil management issues: topsoil slaking, susceptibility to compaction and potential for natural regeneration.

Series Name	Subgroup	% Area of Worth Farms	Susceptibility to top soil slaking	Susceptibility to compaction	Potential for natural structural regeneration
AGNEY	8.12	4.3%	S4	C3	R3
ROCKCLIFFE	8.11	1.8%	S1	C2	R1
ROMNEY	5.32	2.4%	S2	C3	R1
STOCKWITH	8.12	4.5%	S4	C3	R3
TANVATS	8.11	1.0%	S3	C3	R3
WISBECH	8.12	85.9%	S1	C2	R1

Figure 5, Figure 6 and Figure 7 visually represent the susceptibility of the Worth Farm landbank soils to Topsoil Slaking, (StoTS) and Compaction (StoC) and the potential for Natural Structural Regeneration (NR).

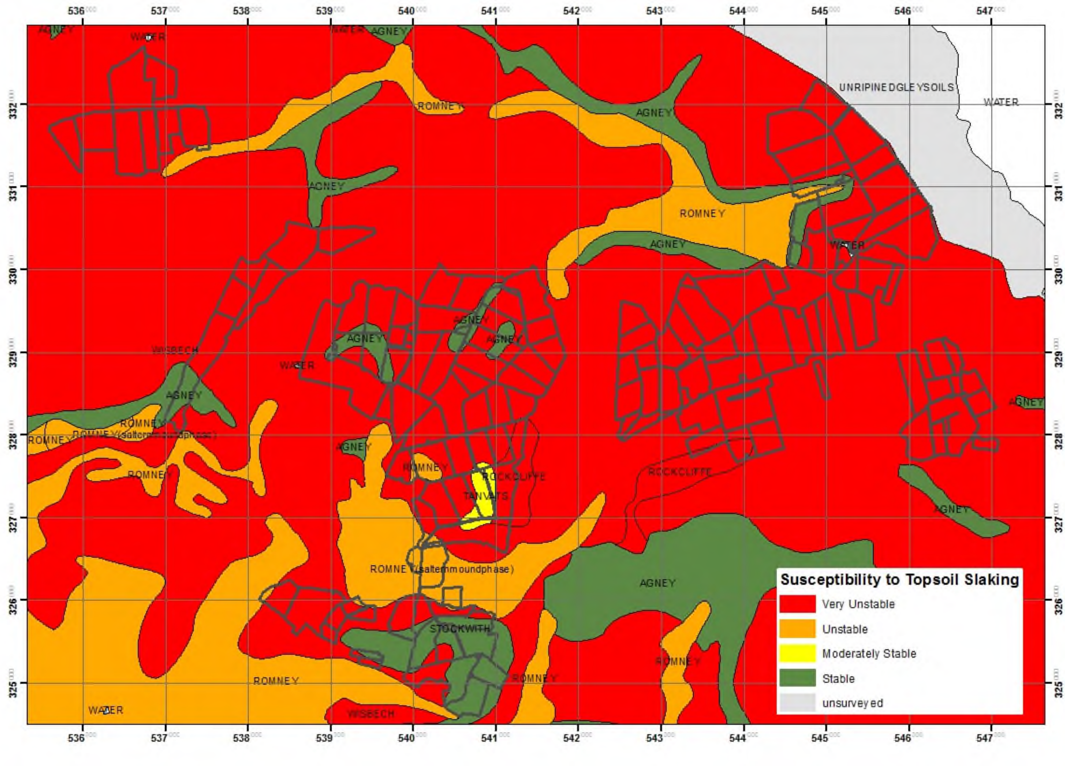

Figure 5. Worth Farms landbank (2010-15): Susceptibility to Topsoil slaking

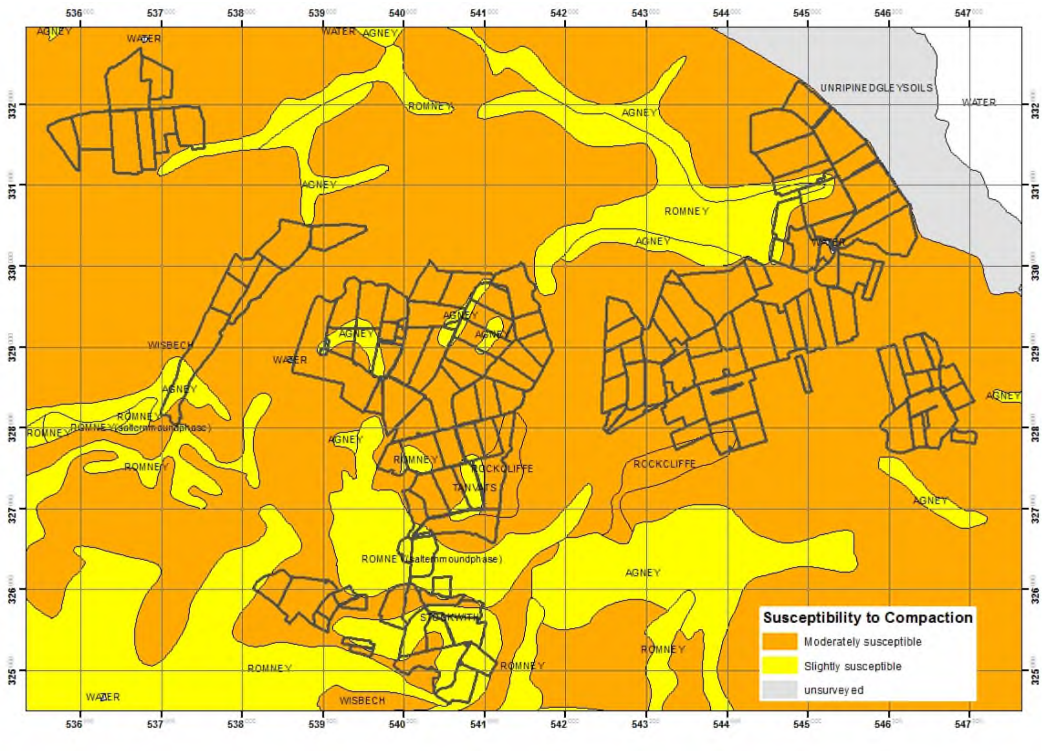

Figure 6. Worth Farms landbank (2010-15): Susceptibility to compaction

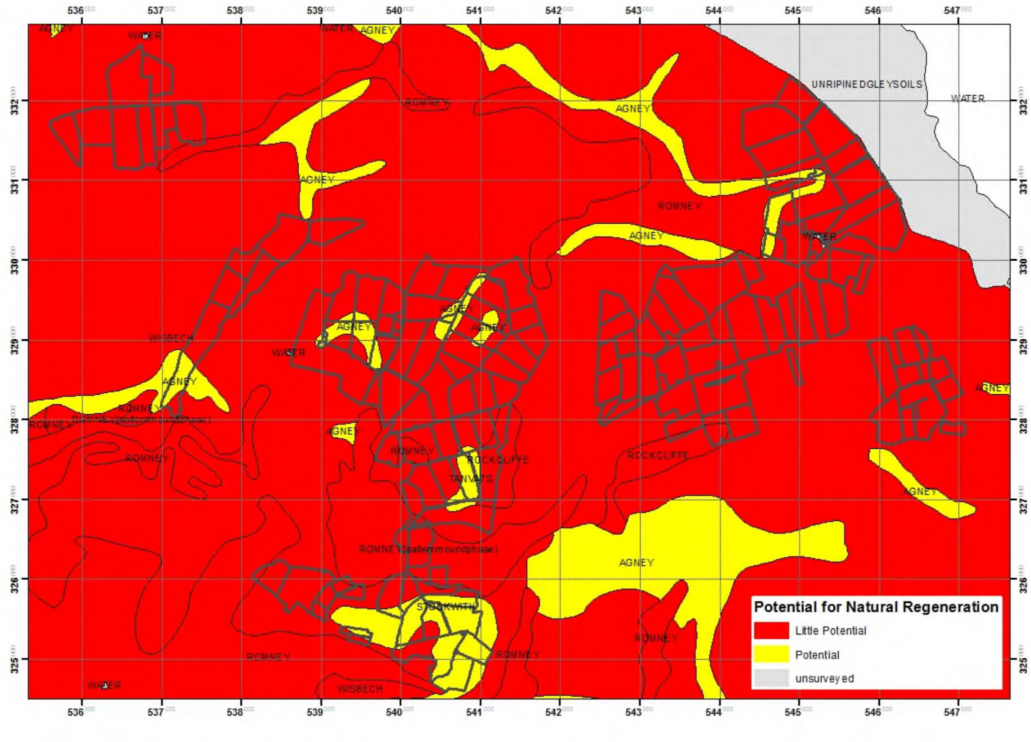

Figure 7. Worth Farms landbank (2010-15): Potential of soil for Natural structural Regeneration (NR).

3.3.3. Using Workability Days (MWDs) as a tool for assessing compaction risk

Interrogation of the Worth Farms dataset (2010-2015) indicates that on a crop specific basis, the majority of machinery/field operations were undertaken at appropriate times e.g. within the MWDs (Figure 4). For brevity, only data for soil management operations in Cauliflower, Vining Peas, Sugar Beet and Winter Wheat are shown (Table 9). However, the results also indicate that a percentage of machinery/field operations were also undertaken outside of the MWDs, thus presenting (for the Worth Farms soils) a high compaction risk (Table 9).

Table 9. Percentage of Worth Farms machinery/field operations undertaken outside of MWDs (2010-2015).

Crop	Machinery Operation Recorded	No. of times operation recorded	No. of times operation occurred outside MWD	% of operations undertaken outside MWD
Cauliflowers	3 Leg Buster	2	2	**100%**
	Disc and Press	1	0	0.0%
	Flatlifting	3	1	**33.3%**
	Plough	37	7	**18.9%**

	Sumo Trio	3	0	0%
	Top Down	3	2	**66.7%**
	Spray Operation	51	5	**9.8%**
Vining Peas	3 Leg Buster	1	0	0%
	Bust Tramlines	1	0	0%
	Disc & Press	17	0	0%
	Fill in Furrows	3	0	0%
	Flatlifting	29	0	0%
	Pea Vining (Harvest)	65	0	0%
	Plough	93	3	3.2%
	Preperator	47	0	0%
	Roll	47	0	0%
	Simba Culti Press	75	0	0%
	Sumo Trio	22	1	**4.5%**
	Top Down	5	2	**40.0%**
	Vardestadt Drill 6m	85	0	0%
	WFL Pea Drill+Cult	16	0	0%
	WFL Spray	511	4	**0.8%**
Sugar Beet	3 Leg Buster	2	0	0%
	Bust Tramlines	5	0	0%
	Disc & Press	5	0	0%
	Flatlifting	47	0	0%
	HCCT Fert Appliction	15	0	0%
	LFP Beet Drill	94	0	0%
	LFP Beet Harvest	112	0	0%
	Omex Application	76	4	**5.3%**
	Plough	106	8	**7.5%**
	Preperator	75	0	0%
	Roll	3	0	0%
	Simba Culti Press	5	2	**40%**
	Beet Drill + Cultivate	10	9	**90%**
	Sumo Trio	21	0	0%
	Top Down	7	0	0%
	WFL Beet Cultivate	19	0	0%
	WFL Fert Application	19	0	0%
	WFL Liquid Fertiliser	79	0	0%
	WFL Spray	776	10	**1.3%**
Winter Wheat	3 Leg Buster	1	0	0%
	AHW Combi Drill(4m)	6	0	0%
	Combination Drill	49	0	0%
	Cultivate	54	0	0%
	Disc & Press	12	0	0%
	Flatlifting	55	0	0%
	HCCT Fert Apptn.	9	2	**22%**
	Plough	58	0	0%
	Power Harrow	6	0	0%
	Roll	52	11	**21%**

	Sumo Trio	14	0	0%
	Top Down	158	0	0%
	Vardestadt Drill 6m	217	0	0%
	WFL Combine	268	0	0%
	WFL Digestate Apptn.	8	5	**62%**
	WFL Fert Application	126	47	**37%**
	WFL Liquid Fertiliser	477	59	**12%**
	WFL Spray	1689	70	**4.0%**

Note: Shaded cells denote tillage operations associated with the mechanical alleviation of compaction (sub-soiling). This demonstrates that compaction is a cross-rotational issue.

Having compiled the data from Gatekeeper and relating this to meteorological and soils data for Worth Farms, data-mining approaches will be undertaken (see Section 3.3.4) to:

- identify the year in which the operation was undertaken
- assess potential consequences on crops yields/yield quality
- identify if additional soil management approaches were adopted in the follow-on crop to mitigate ‘potential compaction’, and their efficacy (in terms of follow-on crop yield)
- determine correlations between soil management practices used and Foot Rot Index

3.3.4. Preliminary Datamining: Results

3.3.4.1. Bayesian Belief Network for Vining Peas yield

Having collated a number of relevant data sets from Worth Farms, (e.g. Gatekeeper, LandIS, yield data, Foot Rot Index, meteorological data), the datamining process and outputs are demonstrated using vining peas as a case study. Bayesian Belief Networks (BBN) were constructed to describe the relationships and drivers of vining pea yield in Worth Farm fields. The BBN is the final output of a set of datamining steps in which the dataset is interrogated using numerous datamining tools to determine the underlying data structure. In this case, we initially applied ‘random forests’ to better understand the factors affecting vining pea yield (i.e. driving variables), and the consequential probabilities and driving variables were used to inform the BBN.

Bayesian Belief Networks are graphical probabilistic models which link a set of nodes (random variables) with a set of directed edges (representing conditional dependencies between nodes). This produces a set of cause and effect relationships between variables in the system being modelled. The inherent uncertainty surrounding these relationships is represented with a probability, which can be determined by either expert knowledge or empirical data (in this case, data was used). An example, Figure 8 presents the layout of the BBN developed to predict pea yield, which includes a number of soil management related factors.

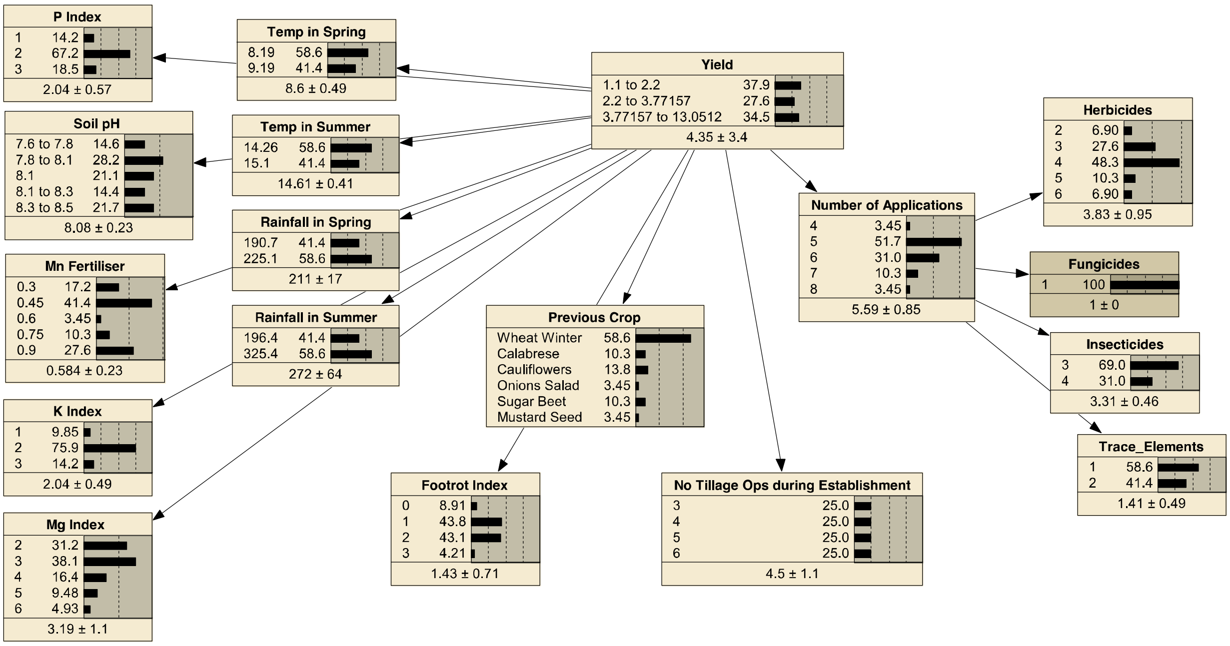

Figure 8. The layout of a Bayesian Belief Network explaining yield of vining peas using Worth Farms data (from Gatekeeper).

The variables used to predict yield of vining peas are explained in Figure 9. They represent only a subset of possible variables (see Table 3), as initially filtered through the datamining process, but contain variables associated to cropping history, soil management and the application of fertiliser (K, Mg-, P- indices, Mn application); variables associated with the weather (temperature and rainfall during spring and summer); disease; and management practices. Bayesian Belief Networks can order covariates according to their explanatory power, and these are given in Figure 9. Here we see that the main factors describing yield are the previous crop, followed by soil pH, soil nutritional aspects and incidence of crop disease, as shown by Foot Rot index for example.

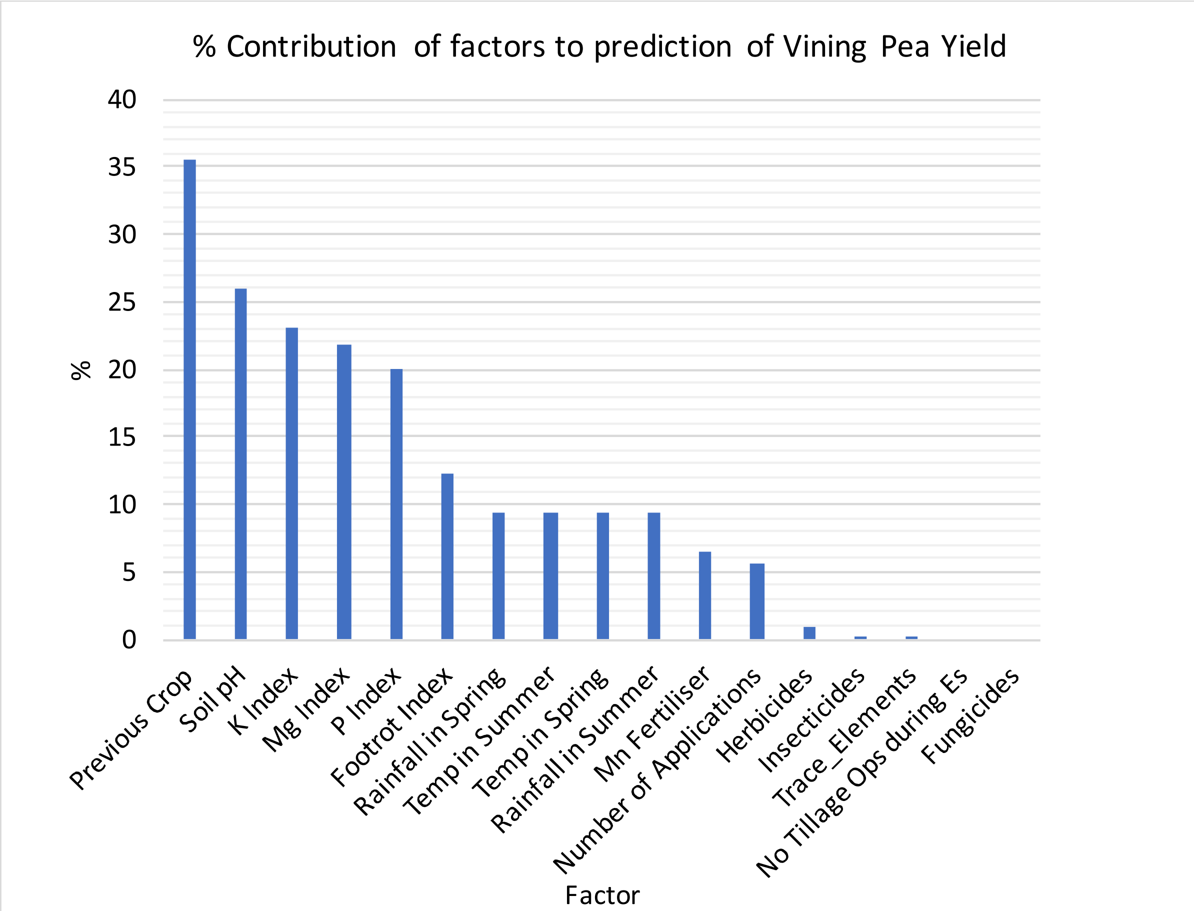

Figure 9. Factor analysis of the Bayesian Belief Network model presented in Figure 1.

This example forms part of the outcome of the datamining exercise, in that it seeks, observes and codifies relationships found in the data contained in the database, but also forms then the basis of the decision making process around SMIS, in that it can be deployed as a model to assist in the decision making rules within the SMIS system.

Critically, as the size and number of records in the database increases (See section 3.3.6) , BBN's can learn and reincorporate new data as it is introduced in the form of updated probabilities and likely outcomes. Knowledge which is not present in the form of data (expert or grower experiences; outcomes from the literature or other experimental results) can also be encoded and included in the BBN probability structure.

3.3.5. Future utility of Gatekeeper datasets: Critical evaluation of management responses to alleviate compaction

The high risk of soil compaction and low potential for natural regeneration has resulted in Worth Farms undertaking sub-soiling to artificially generate 'soil structure' (Personal Communication: Simon Day, Worth Farms; Richard Fitzpatrick HMC Peas). However, sub-soiling, often vaunted as a solution to soil compaction, is a short term, costly and largely ineffectual 'fix' in the longer term (Evans, 1996; Hamza & Anderson, 2005). The ineffectual nature of sub-soiling is demonstrated in the Worth Farms dataset, as sub-soiling is now considered a 'Fixed Cost' and a routine tillage operation for all crops grown.

As a result, Worth Farms (and more widely HMC Peas) have since 2014 started to incorporate cover crops within rotations to bioengineer soil structure (Kirkegaard et al., 2008; Cresswell & Kirkegaard, 1995; Chen and Weil, 2011; Seymour et al. 2012)), as well as for bio-fumigation (the suppression of soil-borne fungal diseases including Fusarium sp.) by isothiocynates released by brassica cover crops (Smolinska et al., 2003). Going forwards, there is the potential for the efficacy of this change in practice across the HMC grower group to be investigated using agri-informatics approaches in terms of, for example:

- Cross rotational increases in crop yield/crop quality
- Reduced frequency of mechanical sub-soiling operations (reduced fuel usage and farm costs)
- Reduced Foot Rot Index
- Cross-rotational reductions in PCN levels (particularly associated with use of mustard cover crops)

4. Data / information / knowledge source: AHDB Horticulture project data

A key objective of the project is to enable the inclusion of data, information and knowledge arising from other on-going soil related research projects being sponsored by AHDB–Horticulture. There are 2 specific projects:

- CP 107b – Growing Resilient Efficient And Thriving (GREAT) Soils, led by the Soil Association, Organic Research Centre & Earthcare Technical Ltd. (01-04-15 to 30-03-18); and
- CP 107c –The application of precision farming technologies to drive sustainable intensification in horticulture cropping systems, led by ADAS with input from SRUC (01-04-15 to 30-03-18).

4.1. CP 107b – Growing Resilient Efficient And Thriving (GREAT) Soils

CP107b is a knowledge transfer project with the overall aim to inspire and support fruit and vegetable growers (primarily) to develop the ability and confidence to assess the health of their soils and take practical action to improve their soil management strategies. Equipped with an improved understanding of soil health, they will be able to choose appropriate methods to enhance and maintain the health of their soils; which is key to good crop yield and quality. In consultation with the PIs, the outputs of this research will be incorporated within SMIS (Year 2 and 3), namely a range of strategies to maintain and enhance soil health in UK horticultural systems, guidance notes on costs and benefits of improving soil health at the farm level; and methods of soil health measurement and management. As well as the quantitative outputs of CP107b, one challenge going forward is to express the more qualitative outputs of CP107b in the SMIS database. SMIS project team are working with Ben Raskin and Simone Osborn in developing strong links and lines of communication.

4.2. CP 107c –The application of precision farming technologies to drive sustainable intensification in horticulture cropping systems

CP107c is a research project with strong industry liaison and guidance. The overall aim is to evaluate the current and future potential of precision farming techniques to optimise soil and nutrient management for improved profitability and sustainable intensification of horticulture crop production systems. A range of annual, biennial and perennial crops is included. The project will provide data into SMIS through its various activities including: assessments of the structural condition of horticultural soils (intrinsic factors of the Requirements Analysis, Section 2.1) and baseline information on typical soil management practices across a range of horticultural crops. The soil structure survey will focus on topsoil and subsoil condition (typically to a depth of 60 cm) and will be carried out both pre- and post-planting/drilling. Field data is being collected into a schema which is compatible with SMIS inputs, although it is envisaged that some transformation of the data will be required to enable its harmonised usage. The provisional data structure that is anticipated from this work is shown in Table 10. We are in close communication with Lizzie Sagoo and Paul Newell Price (ADAS) to ensure synergy between the 2 projects.

Table 10. Provisional data structure from the ADAS CP107c project

Phase	Soil depth	Factor
Baseline	Topsoil (0.7.5cm)	Soil Organic Carbon
		Total-N
		Particle Size Distribution
Pre-Planting		Penetration Resistance (PR) 0-25cm depth
		PR 40-60cm depth
	Bulk Density (BD) mid-topsoil (10-15cm)	Max (PR)
		Median (PR)
		Min (PR)
	BD Upper Sub-soil (c. 35cm)	Max (PR)
		Median (PR)
		Min (PR)
	BD Deeper Sub-soil (c. 40-60cm)	Max (PR)
		Median (PR)
		Min (PR)
		Visual Soil Assessment (VSA; 0-25cm)
		Visual Evaluation of Soil Structure (VESS; (40-

		60cm)
Post-Planting		PR 0-25cm depth
		PR 40-60cm depth
	BD mid-topsoil (10-15cm)	Max (PR)
		Median (PR)
		Min (PR)
	BD Upper Sub-soil (c. 35cm)	Max (PR)
		Median (PR)
		Min (PR)
	BD Deeper Sub-soil (c. 40-60cm)	Max (PR)
		Median (PR)
		Min (PR)
		VSA (0-25cm)
		VESS (40-60cm)

Data relating to quantitative and qualitative metrics of soil structure captured under the ADAS CP107c project will be incorporated into the SMIS system in the form of field specific data. This data will contribute to the development of the SMIS 'rule base'.

Under the CP107c project, field soil structural condition of horticultural soils will be carried out twice (pre- and post-planting/drilling) in a limited number of fields (60 fields across 30 holdings covering the main horticultural regions of the UK). The CP107c survey is stratified by crop type (perennial, biennial and annual), soil type, and land tenure (owned and rented), with the stratification by crop type reflecting the importance of each crop group in terms of production area and levy contribution e.g. 10-15 brassica fields, 10 carrot fields and 10-15 onion fields etc.

The soil structure survey focuses on topsoil and subsoil condition (typically to a depth of 60 cm) and will be carried out both pre- and post-planting/drilling. Metrics of soil structural condition recorded under the CP107c project include i) Penetration resistance, ii) Bulk Density and iii) Visual Assessments from the mid topsoil (at 10-15 cm), upper subsoil (at *c*.35 cm) and deeper subsoil (40-60 cm) layers. In addition, topsoil samples will be taken (0-7.5 cm) and analysed for soil organic carbon, total nitrogen and particle size distribution (i.e. sand, silt and clay content) to characterise each field.

The utility of the CP107c data to SMIS will be substantially enhanced if the Gatekeeper (or equivalent) data for the CP107c fields can be obtained such that the metrics measured can be put into the context of the cropping, operational and climatic history of the field sampled (e.g. relevant intrinsic and extrinsic factors are covered). Capturing Gatekeeper (or equivalent) data for some of the CP107c fields will need to be undertaken going forwards.

To date, no final data from CP107c is currently available.

4.3. Other research projects

A search of relevant research project reports has also been conducted by sifting through projects listed on funder websites. To date, this has included Defra projects under the Crop and Horticultural Business Unit, and AHDB Horticultural projects spanning the last 10 years. As per the project timelines (Appendix 1), more analysis of these projects will be undertaken in Year 2 and 3. For example, information from the following AHDB projects will be used to provide input to SMIS:

FV 361 (SA Link SA563/LK09130) Sustainable Arable Link: Reducing the impact of Sclerotinia disease on arable rotations, vegetable crops and land use.

FV 429: Vining peas: Development of a laboratory based assay for the detection of Common Root Rot (*Aphanomyces euteiches*).

FV 345: Establishing Best Practice for determining soil nitrogen supply - addition of field Veg sites to HGCA project 3425

FV 380: Vining Peas: Identification of critical soil P levels

FV 428: Vining peas: The effect of soil phosphate levels on rhizobial population.

FV450 – Asparagus: Sustainable soil management for stand longevity and yield optimization

We have also listed BBSRC projects that contain the search terms: soil* AND plant* (n = 441); soil* AND crop* (n = 340); soil* AND hort* (n = 36), based on research grants, institute projects and fellowships, including studentships and training grants current or completed between 2001 and the present. Primary investigator details for relevant projects of the latter have been captured so that contact can be made to request datasets that can be used to inform SMIS (Appendix 5). This activity will be undertaken at the beginning of Year 2 of the project.

5. Data / information / knowledge source: Capturing expert knowledge

On behalf of the HDC, Rickson and Deeks (2013) conducted a review of soil management challenges and practices in horticulture (CP107a). The purpose of that review was to identify key gaps in research and in knowledge transfer mechanisms which hinder the development and implementation of best practice guidelines for sustainable soil management. As a result of the grower survey undertaken as part of the gap analysis, some 11 specific challenges were identified (Table 11).

Table 11. Soil management challenges in horticulture, identified in CP107a (Rickson and Deeks, 2013)

Soil Management Challenges in Horticulture		
Soil erosion by water	Drought	Yield quantity
Soil erosion by wind	Drainage	Yield quality
Soil compaction	Accessing wet soils	Yield reliability
Too little organic matter	Soil-borne diseases	

To reflect end user requirements of SMIS, all these issues (and advice on how to manage them) must be at the heart of the SMIS and eGuide. Therefore, to build the structures of the SMIS database (that can be interrogated at a later stage), it is necessary to visualise (or 'map') the causal factors behind these challenges and the outcomes of any measures used to mitigate them (i.e. the effectiveness of soil management practices). To do this, a series of workshops and questionnaires were undertaken, following a 'Fuzzy-logic Cognitive Mapping' (FCM) approach. The FCM was designed to identify and map the (expected) 'cause and effect' relationships between:

- intrinsic site factors (soil, climate, aspect etc.);
- soil management issues identified (erosion, incidence of soil borne diseases and pests, compaction etc.);
- inputs (soil management solutions / approaches e.g. minimum tillage); and
- consequences of applying these solutions (e.g. less soil erosion).

FCM is a semi-quantitative and dynamic method used to structure expert knowledge. It uses graphical representations of a system to illustrate relationships and boundaries between discrete nodes (e.g. in this research causes and consequences of soil management challenges). Cognitive linkages and flow directions that represent direct and indirect causality, are defined to build a community understanding of the defined issue/concept. The resulting cognitive maps are suggested to be an external representation of internal mental models derived from experience and expert knowledge (Jones et al. 2011), helping exchange of individual understanding. Fuzzy-logic Cognitive Mapping (FCM) was deemed an appropriate methodology to capture expert information / knowledge on soil management and then organise cause and effect relationships in a structured way that can be incorporated into the SMIS (See Section 8.3). This recognised technique was chosen as it has already been used successfully to model dynamic systems (including agricultural applications; Table 12), representing causal relationships between diverse system variables and analyse inference patterns (van Vliet et al., 2016; Khadra et al., 2011).

Table 12. A selection of agri-environmental projects that have applied 'Fuzzy-logic Cognitive Mapping' (FCM) successfully

Project and location	Reference
Water use efficiency in agriculture, Malta.	Cranfield University staff
Sustainable wetland management, Uganda.	Bosma et al. (2017)
Vineyard grape production, Greece.	Groumpos and Groumpos (2016)
Environmental management systems for mineral works, South Africa.	Mbele and Masinde (2016)
Sugar cane yield predictions, India.	Natarajan et al. (2016)
Improving agricultural policy design, Scotland.	Christen et al. (2015)
Predicting yield in cotton crop production, Greece.	Papageorgiou et al. (2011)
Community water management, Italy.	Giordano et al. (2005)
Ecosystem management planning for Uluabat Lake, Turkey.	Özesmi and Özesmi (2003)
Conservation of the Sultan Marsh ecosystems, Turkey.	Dadaser and Özesmi (2002)

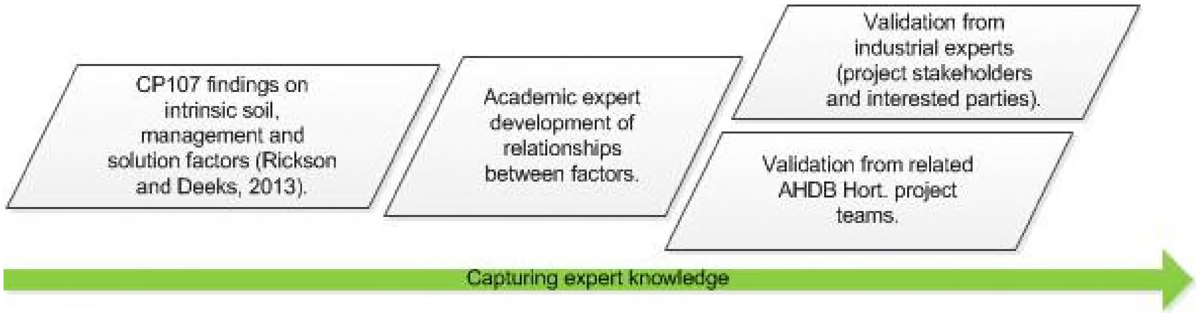

Figure 10. The breadth of expert knowledge gathered to date that informs the FCM output.

FCM was used to capture both academic and industrial expert knowledge (Figure 10). The FCM was initially trialled in an internal workshop held over 2 days utilising the expertise of Cranfield staff, spanning a breadth of soil management disciplines. The aim of the first workshop session was to identify the intrinsic factors pertinent to soil management in horticulture (e.g. rainfall, soil type, crop etc.), the problems that can arise (e.g. soil compaction) and the solutions available (e.g. subsoiling), as well as their intra- and inter-relationships (Figure 11). The second workshop session aimed to develop the FCM structure further by describing any causal relationships, in terms of their direction (positive or negative relationship), cardinality and strength. A weighting system was used to quantify the strength of each relationship. A positive three-scale weight (+, ++, +++) was assigned to indicate a positive association between factors (Khadra et al., 2011) e.g. the addition of organic matter has a positive relationship (++) with soil structure. A negative three-scale weight (-, --, ---) indicates a negative association (Khadra et al., 2011) e.g. soil capping/sealing inhibits (--) crop emergence.

Beyond the workshop, the direction and weighting of each relationship within the FCM can be corroborated by the other sources of data / information / knowledge (e.g. from grower data (Section 3), research projects (Section 4) and the literature (Section 6)). We will also consult with industry partners to ensure all relevant soil management issues and their mitigation continue to be captured.

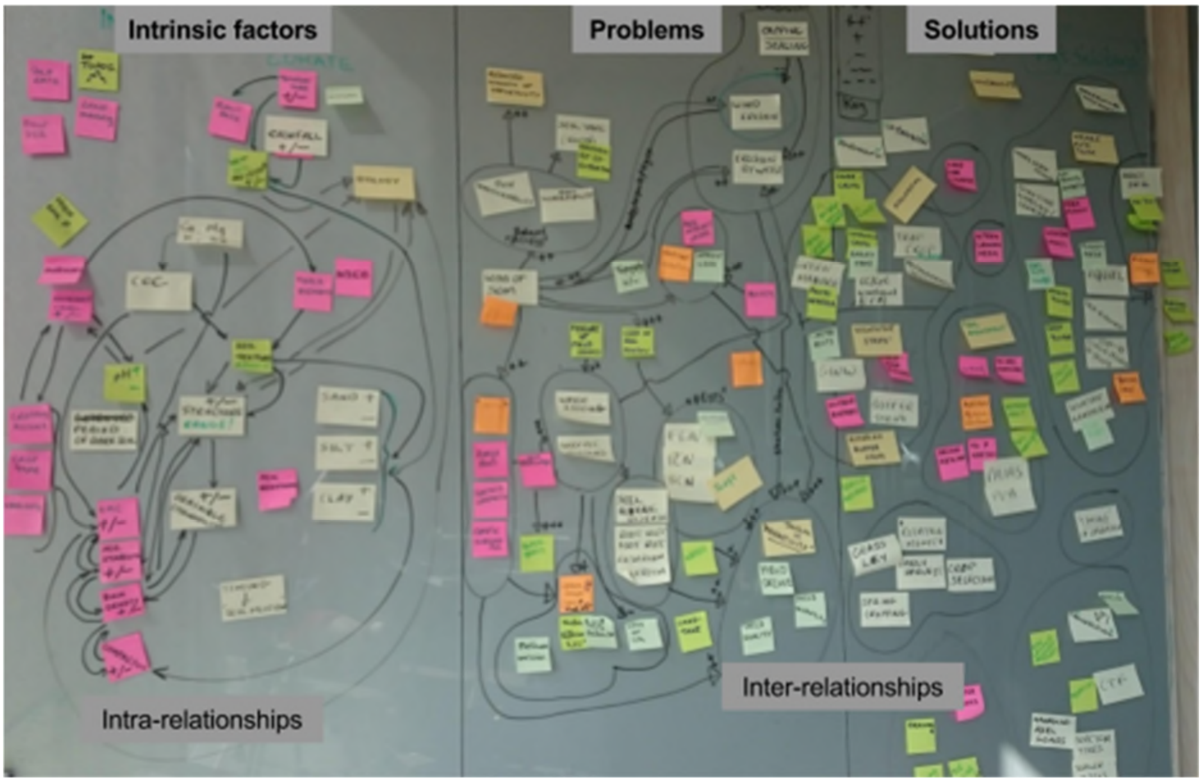

Figure 11. FCM development during the internal workshop.

Individual ‘cause and effect’ relationships were mapped for some of the identified soil management issues; compaction (Figure 12), erosion by water (Figure 13), erosion by wind (Figure 14) and soil-borne disease (Figure 15); for each of these, first, second and third order factors are shown for a) inherent factors affecting the issue, and b) management solutions to the issue. The workshop FCM output was also used to build a framework around soil management issues that would ultimately lead to a decline in crop productivity (quantity, quality and reliability) (Figure 16). These ‘knowledge maps’ were subsequently sent out for initial validation by industrial experts (Table 13), including the project stakeholder group and other horticultural industry contacts (including the Processing and Growers Research Organisation members). Copies were also sent out to the project teams involved with related AHDB Horticulture projects; CP107b “Growing Resilient Efficient And Thriving (GREAT) Soils” and CP107c “The application of precision farming technologies to drive sustainable intensification in horticulture cropping systems”.

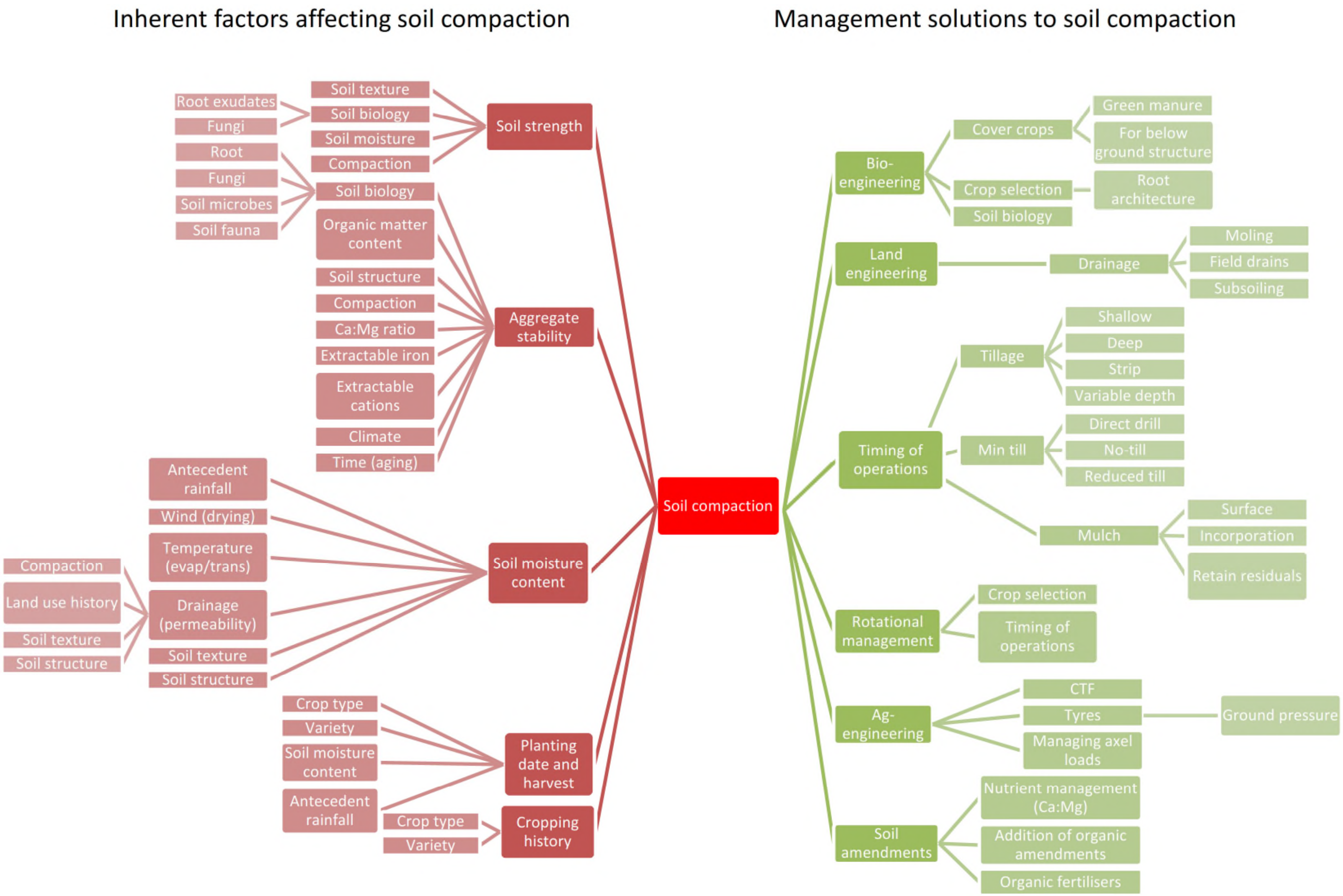

Figure 12. Soil compaction knowledge map.

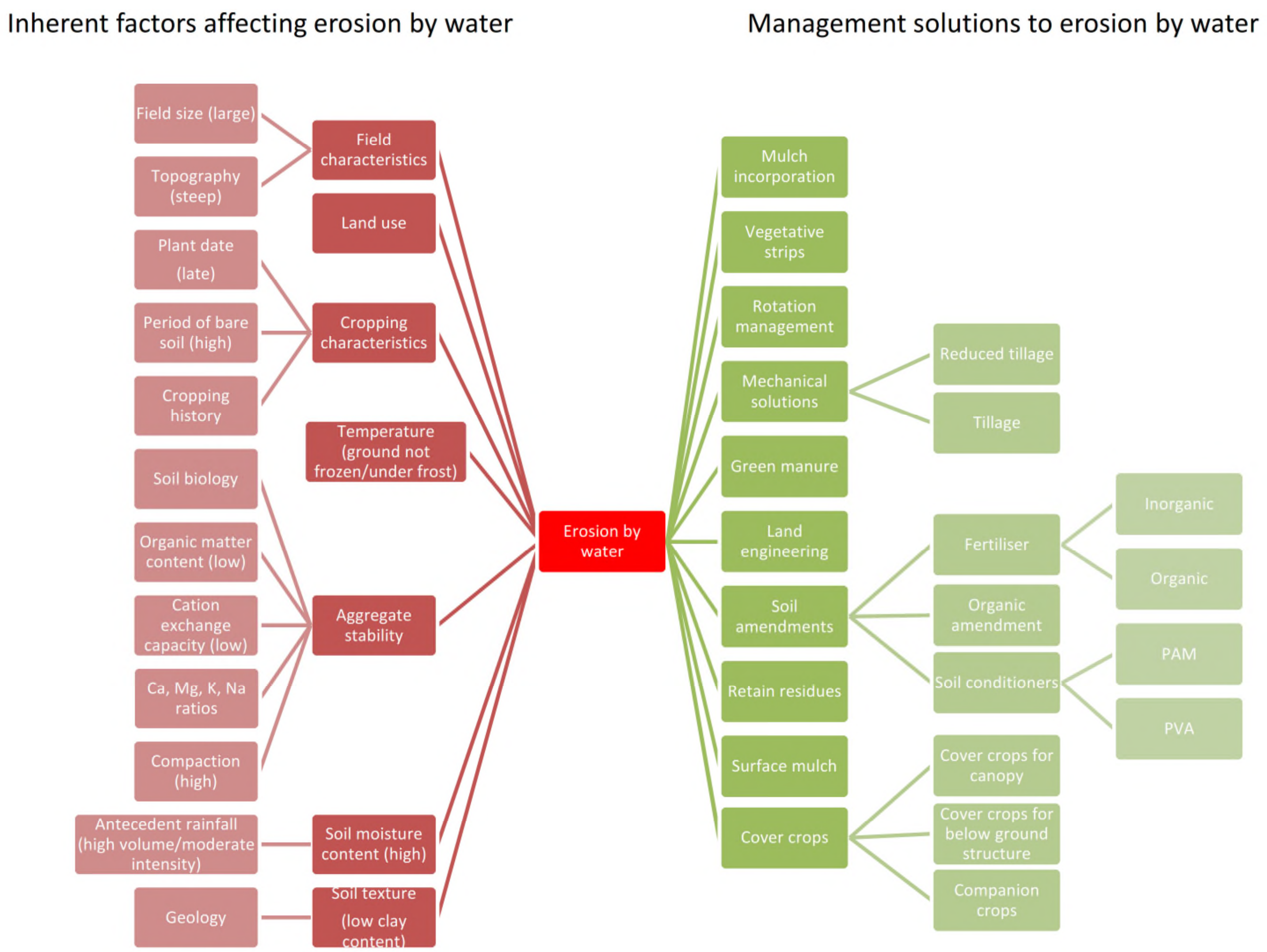

Figure 13. Erosion by water knowledge map.

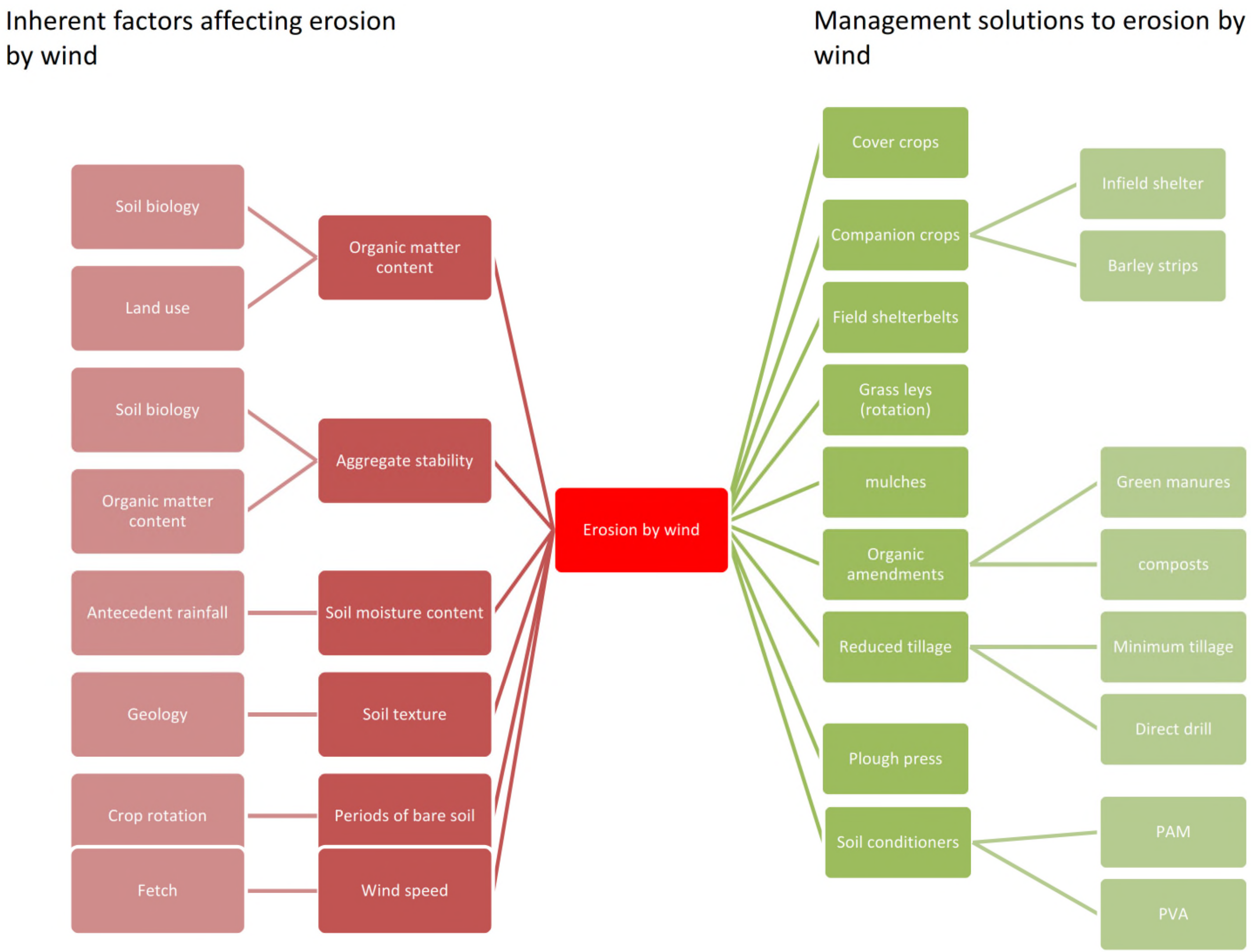

Figure 14. Erosion by wind knowledge map.

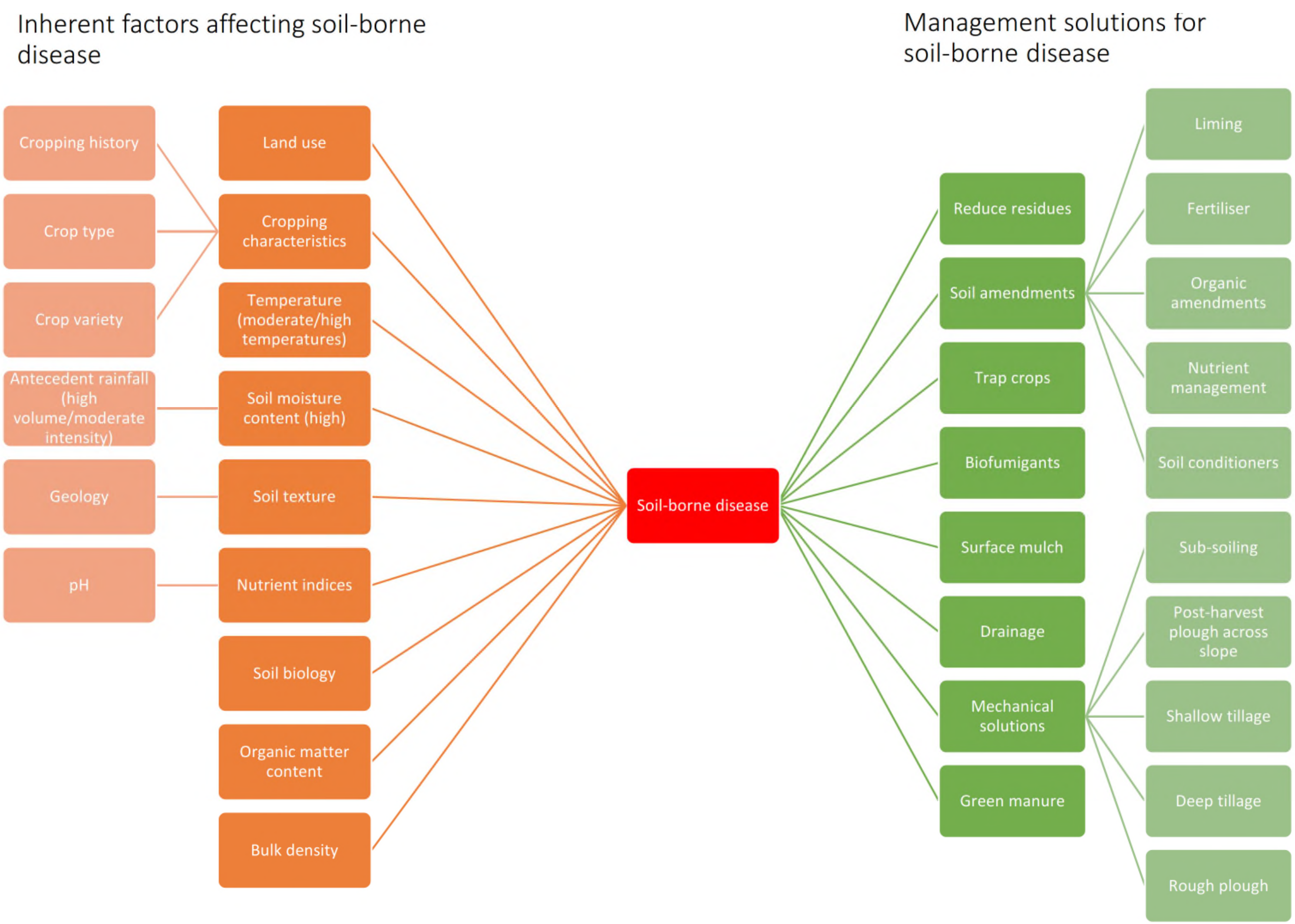

Figure 15. Soil-borne disease knowledge map.

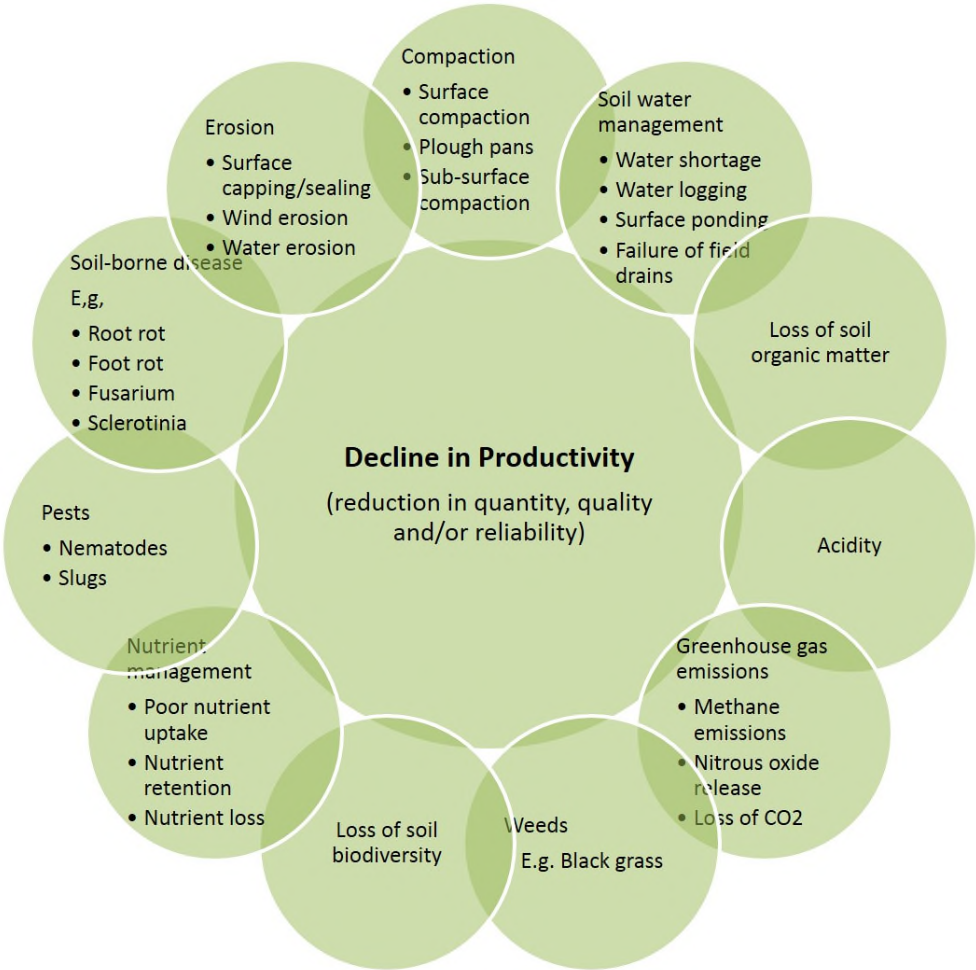

Figure 16. Soil management issues leading to a decline in productivity.

Table 13. List of industry experts to validate the knowledge maps.

Contact	Organisation
Andrew Lenson	Velcourt Farms Ltd
Andrew Letham	Scottish Borders Produce Limited
Andrew Whiting	Birds Eye
Andy Beach	Anglian Pea Growers
Ben Raskin	CP107b project team
Charles Shropshire	Gs Fresh
Chris Chinn	British Asparagus Growers Association and Cobrey Farms
Chris Stoate	Allerton Trust / Game and Wildlife Conservancy Trust

Claire Donkin	British Herb Growers Association
Emma Garfield	Gs Fresh
Ewan Findlay	Stemgold Peas
Fred Richardson	East Coast Viners
Guy Thallon	Produce World
Ian Watson	Farm Care Ltd and Stemgold Peas
James Thompson	J W Grant Farmers
John Chinn	Cobrey Farms
John Williams	ADAS
Lionel Murfet	Princes
Matthew Hayward	Swaythorpe Gorwers
Neil Cairns	Barfoots
Neil Murray	Bruce Farms
Paul Newell-Price and Lizzie Sagoo	CP107c project team
Philip Dodds	Herbsunlimited
Richard Byass	Scottish Borders Produce Limited
Richard Fitzpatrick	HMC Peas
Robin Buck	Buck Farms Ltd
Russell Corfield	Aylsham Growers
Stephen Francis	Fen Peas
Stuart Ashton	Pinguin Foods
Tom Davies	Malvern Herbs
William Brady	Green Pea co.

5.1. Feedback on soil management issues from industry experts

Contact with industry was made electronically, with the option of a hard-copy to be sent out if required. The knowledge maps were annotated with a number of questions (Table 14). Recipients were given 25 days (from 21st June 2016) to respond and were issued a reminder on the 11th July 2016. Responses received after this date were still captured. To date, 8 out of 31 responses have been received, giving a 26% response rate.

Table 14. Supporting questions sent out with the FCM's to stakeholders/industry contacts.

Question type	Question
General	Which is the most important soil management issue for your business
Productivity specific	Have we captured all the soil management issues that can result in a decline in productivity?
FCM specific	Have we captured all the factors causing the soil management issue presented?
FCM specific	Have we captured all the factors that affect the soil management issue presented?
FCM specific	Which management solutions work/don't work for you?

Respondents represented a range of horticultural sectors (Figure 17). Feedback was positive and informative (Appendix 3) and enabled further improvements to be made to each knowledge map. It also helped validate that the soil management issues identified to date were industry relevant (Figure 18). Particularly useful feedback also included insight into currently practiced soil management options (Table 15; Table 16), as well as knowledge/research gaps (alongside those identified by Rickson and Deeks, 2013) that SMIS can help to address.

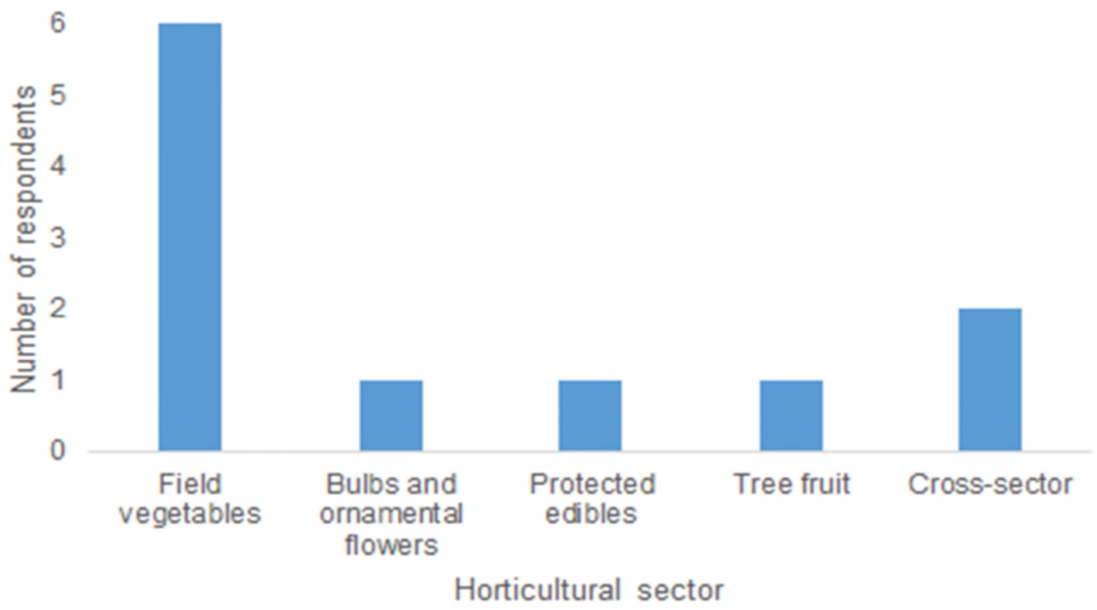

Figure 17. Principal horticultural sectors represented by the respondents.

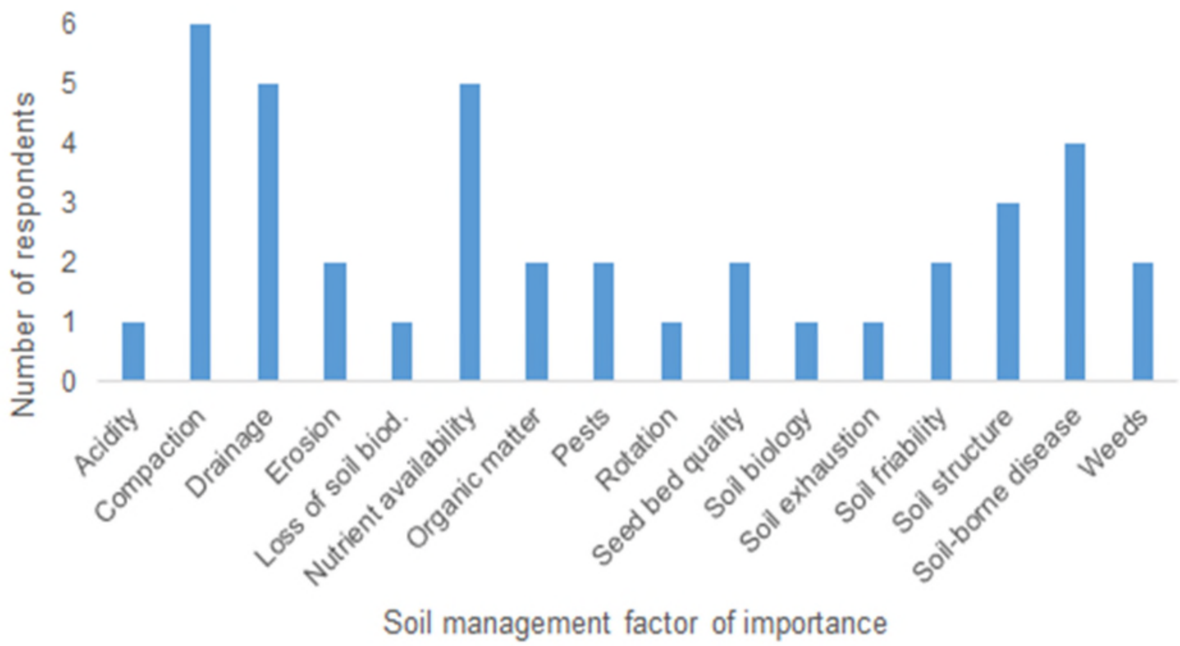

Figure 18. Soil management factors deemed to be important by respondents.

Table 15. Industrial insights gained from respondents concerning tried-and-tested solutions to soil management issues.

Management issue	Experience of management practices

	Positive	Negative
Compaction	-Adequate harvest resource for unusual times -Cover crops -Companion crops -Land drainage -Sub-soiling -Field choice and layout -Minimal soil movement (including ridges) -Increased organic matter content	-Controlled traffic farming on a bed system does not work due to the narrow tyres required to operate between beds. -Minimum-tillage is not always an option
Erosion	-Timing of operations -Row direction relative to prevailing wind -Inorganic fertiliser encouraging early season vegetative growth -Increase organic matter content (long term) -Vegetative strips -Row direction relative to slope -Cover crops -Companion crops -Field choice -Minimum soil movement (including ridges) -Slope management	-Minimum-tillage is not always an option -There is a conflict between winter cover crops and frost mould creation for good seedbed creation for spring crops
Soil-borne disease	-Irrigation management (prevent spread of disease) -Maintaining plant health -Drainage	

Table 16. Respondent feedback highlighting how SMIS can help inform.

Management issue	Feedback (as given)
Erosion (water)	No knowledge of soil conditioners
Erosion (water)	More work needed on cover cropping
Compaction	Adding organic matter in rows would help but there is no perceivable cost benefit to this so few, if any, growers take it up (it is done successfully in Canada/USA)

5.2. Utilising FCM outputs and knowledge maps

The FCM captured the causal factors of soil management issues and their mitigation, as identified by academic and practitioner experts. As well as providing knowledge and information to be built into SMIS (Section 7.1.), the output from the validated knowledge maps was used to structure the literature review (Section 6). The management issues identified formed the search criteria for new literature and the categories for the literature already identified by Rickson and Deeks (2013). Inherent factors and management solutions were used as initial sub-categories for literature searching within each management issue.

Integrating the outputs of the FCM exercise and other sources of data / information / knowledge in the SMIS will be undertaken in Years 2 and 3 of the project.

6. Data / information / knowledge source: Literature review

As well as grower data (Section 3), expert knowledge (Section 5) and research project outputs (Section 7), the database of SMIS will incorporate the robust evidence-base of different soil management issues and their solutions from the academic and practitioner literature.

6.1. Academic literature

The review of academic literature had two phases. The first phase required the collation of literature relevant to horticultural soil management. This brought together literature from two sources:

- the literature identified in the gap-analysis review conducted by Rickson and Deeks (2013).
- a new literature search (detailed below) that updated Rickson and Deeks' (2013) literature with relevant published papers between 2013 and 2016.

The second phase was then to review the collated literature from both sources in order to identify, categorise and catalogue case studies that will provide data, information and/or knowledge to be incorporated into SMIS, specifically on soil management issues and best management practices for horticultural crops (Figure 19). To date, this has been undertaken fully on the 2013 to 2016 literature (Appendix 4 and detailed below).

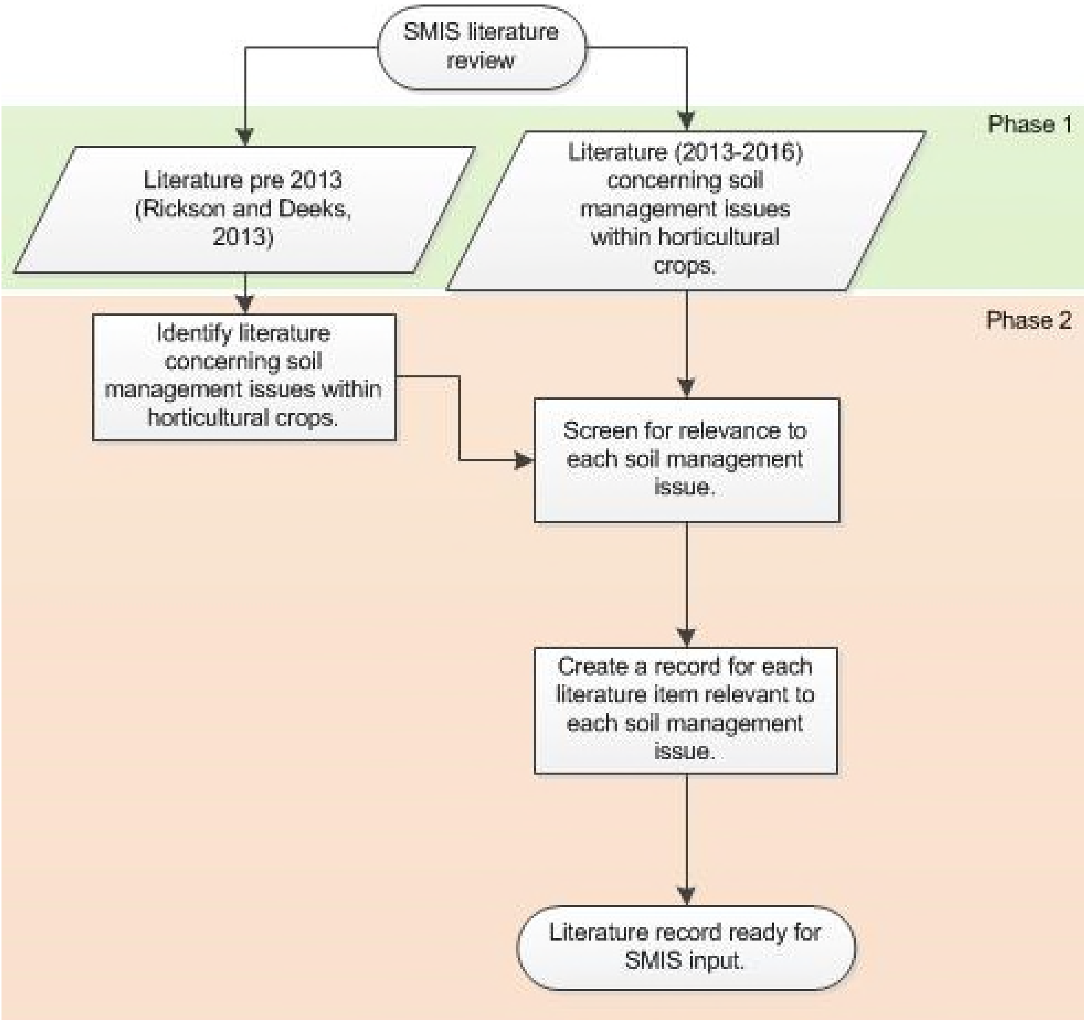

Figure 19. Literature review process

6.2. Updating the literature review (for 2013-16)

An academic literature search was undertaken using Scopus (http://www.scopus.com/home.url) for literature published between 2013 and 2016 (Appendix 4). The search terms trialled for this search are presented in Table 17. The aim was to select a search term that was not too broad (exhaustive and time consuming) and not too restrictive (with the risk that not all relevant articles would be captured). The keyword search results were considered to be most representative of relevant literature (Table 17). The citations and abstracts were exported into Mendeley, a reference management software.

Table 17. Search terms trialled in Scopus (August 2016) and the number of hits. *denotes wild cards.

Search term(s)	In abstract	In title	In keywords	Comment
(("Soil") AND ("horticultur*"))	639	21	52	Search term too broad in abstract and titles, but will capture all relevant papers
(("Soil") AND ("manage*") AND ("horticultur*"))	174	1	13	Search term too broad, but with no papers with search term in the title
(("Soil manage*") AND ("horticultur*"))	16	0	1	Search term reasonable, but with no papers with search term in the title
(("Soil") AND ("manage*") AND ("vegetable*"))	473	9	68	Search term too broad in abstract, with too few papers with relevant title
(("Soil") AND ("manage*") AND ("fruit*"))	676	5	71	Search term too broad, with too few papers with relevant title
(("Soil") AND ("manage*") AND ("mushroom*"))	25	0	4	Search term too narrow in abstract, with no papers with relevant title
(("Soil") AND ("manage*") AND ("protected crop*"))	5	0	0	No papers with search term in the abstract or title
Totals	2008	36	206	

These papers were categorised by soil management issue derived from Rickson and Deeks' report (2013) and the FCM exercise (Section 5; Table 18). This was considered an appropriate search as these papers were likely to include information on best management practices used to mitigate these issues.

Table 18. Soil management issues and their respective search terms.

Soil management issue affecting crop productivity*	Search term
Erosion (by water and wind)	Erosion

Soil management issue affecting crop productivity*	**Search term**
Compaction	Compaction
Soil-borne disease	Disease
Pests	Pest
Nutrient supply	Nutrient
Soil biodiversity	Biodivers / diversity
Weeds	Weed
Acidity	pH
Soil organic matter decline	Organic matter / carbon
Soil drainage	Water

*Based on Rickson and Deeks (2013).

The search for the management issues (Table 18) was conducted in Mendeley. Articles containing the selected search terms were separated into folders relevant to each soil management issue. Each folder was then scrutinised in turn to identify whether the search term 'hit' was relevant to horticultural crop best management practices. Articles deemed to be irrelevant were removed from the folder. Articles that were relevant, but concerned cropping conditions not prevalent in the UK (e.g. rice paddies) were also removed. This left just 42 relevant papers published between 2013-16; most addressed organic matter decline and nutrient supply more than soil compaction, soil erosion (by water and wind), pests and weeds (Figure 20).

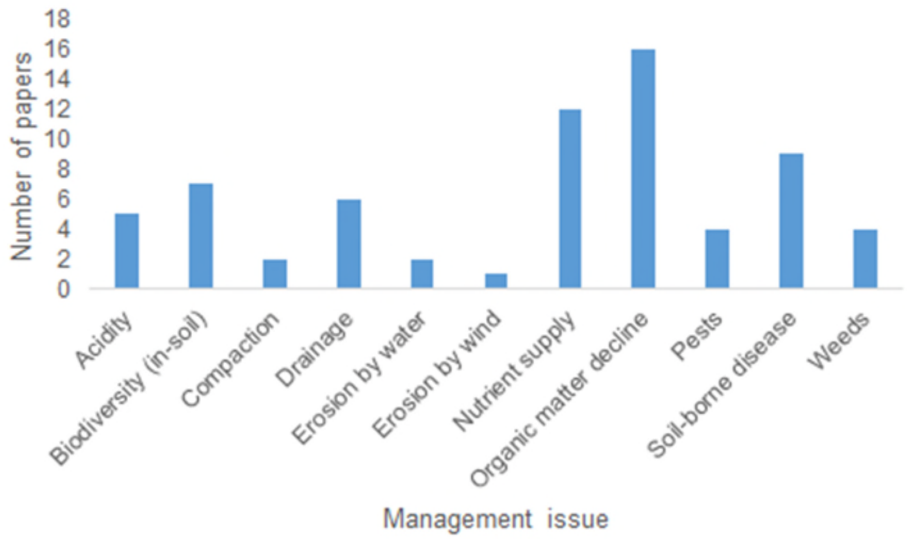

Figure 20. The distribution of relevant papers across soil management issue based on the search criteria indicated in Table 17 and Table 18. N.B. The total number of papers exceeds 42 as papers often dealt with more than one soil management issue.

A record was created for each soil management issue, detailing the relevant papers. Details recorded were specific to each soil management issue and covered the inherent factors of the site under investigation, the soil management solutions described, the crop investigated, the country in which the work was undertaken and the nature of the research (e.g. qualitative, quantitative or anecdotal). An example of capturing one soil management issue (loss of soil biodiversity) is given in Table 19. Others are shown in Appendix 4.

Table 19. Catalogue of articles (2013-2016) relevant to Soil Biodiversity.

Citation	Crop	Inherent Factor	Management solution	Country	Study type	Note
Carron et al., 2015	Palm oil		Organic amendment	Indonesia	Quant.	Field trial. On mature plantation
Dorias & Alsanius, 2015	Fruit and veg		Organic management	Global	Qual.	Review
Doring et al., 2015	Grapes		Organic management	Germany	Anecd.	Quoting from another source

Citation	Crop	Inherent Factor	Management solution	Country	Study type	Note
Ge et al., 2013	Field veg	SOC, N, CEC	Organic management	China	Quant.	Field trial
Gonzalez et al., 2014	Field veg	Soil C		China	Quant.	Field trial. N-fixing organisms
Liu et al., 2015	Tomato		Organic fertiliser	China	Quant.	Field trial
Shen et al., 2015	Banana		Organic amendment	China	Quant.	Field trial. Manure and organic fertiliser

A complete list of identified literature and their associated soil management issues can be found in Appendix 4.

Paper screening demonstrated that there was a wide distribution of soil management practices used to address these issues, covering 24 broad categories (Figure 21). Fewer papers were found that specifically addressed inherent site factors (e.g. organic matter content), and how they both contribute to, and are altered by, soil management practices (Figure 22).

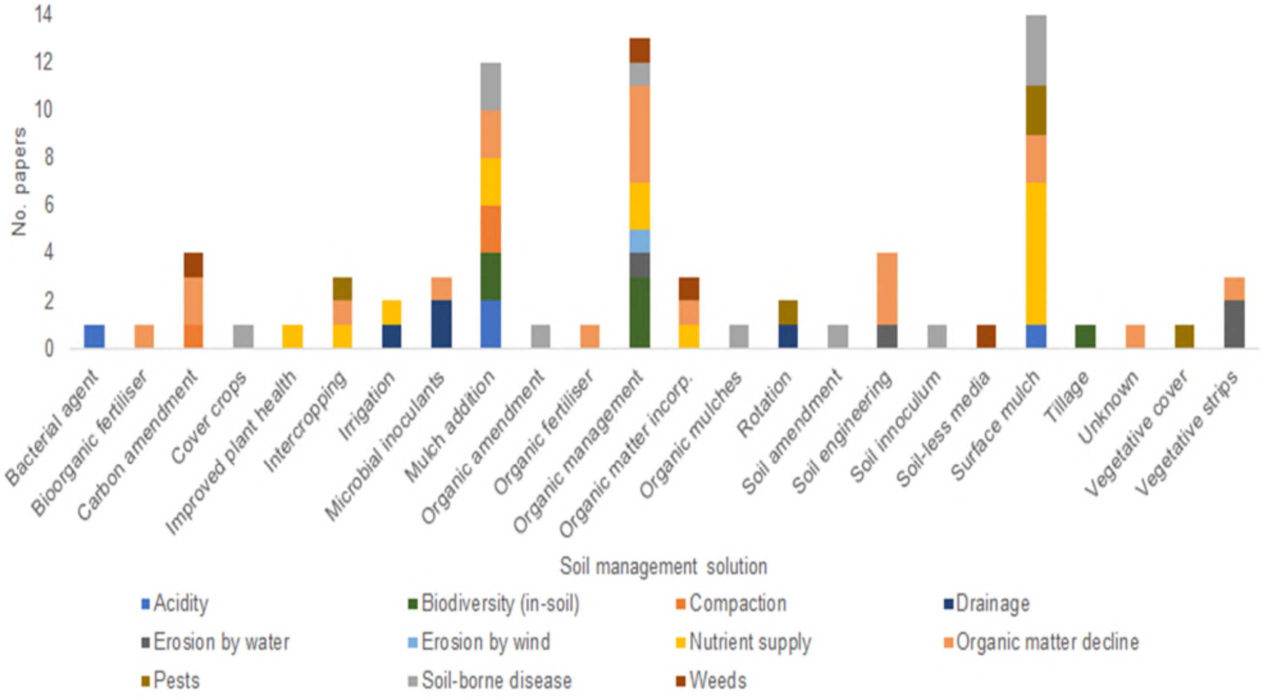

Figure 21. The number of relevant papers that address broadly classified solutions to the soil management issues.

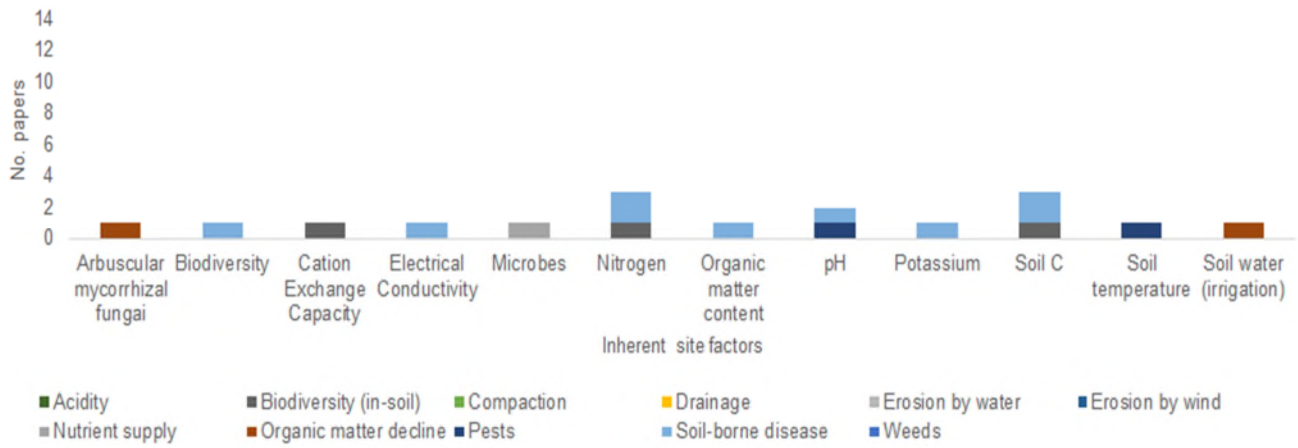

Figure 22. The number of relevant papers found to address different inherent site factors that can result in soil management issues.

6.3. Literature incorporation into SMIS

The research outputs (whether as data, information or knowledge) from each paper will sit in SMIS in record form, containing (where available) the 6 criteria outlined in Figure 25. An example of such a record using Berg (1984) is given in Table 21. The outputs will be used to develop the soil management ‘rule base’ (described in Section 8.3 below). Then, when interrogated with a particular soil management issue (e.g. soil compaction), SMIS will allow the relevant literature to be identified and drawn out through the SMIS switchboard (Section 8.3.1).

7. Results

Appendix 1 (Project Deliverables and Milestones) shows the key milestones for delivery of the project. Most fall after the date of this report (December 2016). The Requirements Analysis is completed (Section 2.1), which identified the required scope and purpose of the SMIS and eGuide. Systems analysis and the literature review are well underway and will be delivered at the end of January and February 2017 respectively. Construction of the SMIS system itself is due in January 2018. Data loading and testing will begin in May 2017, supported by case study assembly. The analytics toolkit will be developed during Years 2 and 3 of the project.

The progress on the development of SMIS to date is reported below.

7.1. Building the SMIS

At the centre of the project activity is the development of the Soil Management Information System (SMIS), which is able to receive and hold data, information and knowledge (Sections 3 - 6), and provide the means for its transformation into soil management guidance and advice, as delivered by the eGuide. Year 1 activity (see Appendix 1) in this area has involved both the development of technical approaches, and the collation of the types and extent of knowledge sources (Sections 3 - 6) which are the foundation of the SMIS data base. Presented in this report is a summary of the principle steps and outcomes of this developmental phase of the project. Key to this are activities related to:

- the classification and categorisation of factors pertaining to soil management (and their interdependencies), as extracted and elaborated from the data and expert judgement; and
- the various outcomes anticipated from datamining approaches (Section 3.3.4)

Doing this will allow the e-Guide to support queries that cover all aspects of soil management issues, practices and outcomes, which are currently not available in one centralised repository.

The anticipated user base for the e-Guide (who will draw information from the SMIS) is assumed to be the technical staff of AHDB Horticulture who will be able to operate the final system delivered, and who will interact with the unique knowledge base collated in this project. Example applications for the e-Guide are shown in Table 20.

Table 20. Intended use of e-Guide and SMIS.

AHDB Horticulture staff seeking to plan directed research programmes to address gap analysis.
AHDB Horticulture staff seeking to interact with, and influence, grower groups and land based industries.
AHDB Horticulture staff responding to specific questions from growers, or more likely, in preparing communications (e.g. guidance flyers and brochures) to growers.

As the SMIS project proposal outlined, the team identified the importance of incorporating information from the extensive academic and practitioner ('grey') literature available (Section 6). The large extent of the relevant literature lent itself to the use of text mining and natural language processing (NLP) techniques to aid extraction of key data, and prioritisation of papers that were most relevant in identifying soil management issues and the solutions used to mitigate these. These techniques could potentially extract keywords from the literature to allow automatic searching and collation of relevant sources of information. To undertake this, a series of prototype code scripts were written in the language 'R', incorporating the R text mining library 'TM'. Appendix 6 provides examples of the outcomes of this work. Unfortunately, due to the breath of subject matter to be investigated related to soil management issues and solutions, this approach was unsuccessful.

7.2. Overview of the SMIS architecture and its data management approach

The following section describes how the data, information and knowledge sources described in previous sections are being used to develop the soil management 'rule base' within SMIS.

7.2.1. SMIS Architecture and Use Case

The SMIS knowledge base is designed to support user queries that can apply across all of the multiple sources of information. This 'federated search' function will permit retrieval of items from the knowledge base that are related to a common set of criteria (Figure 23).

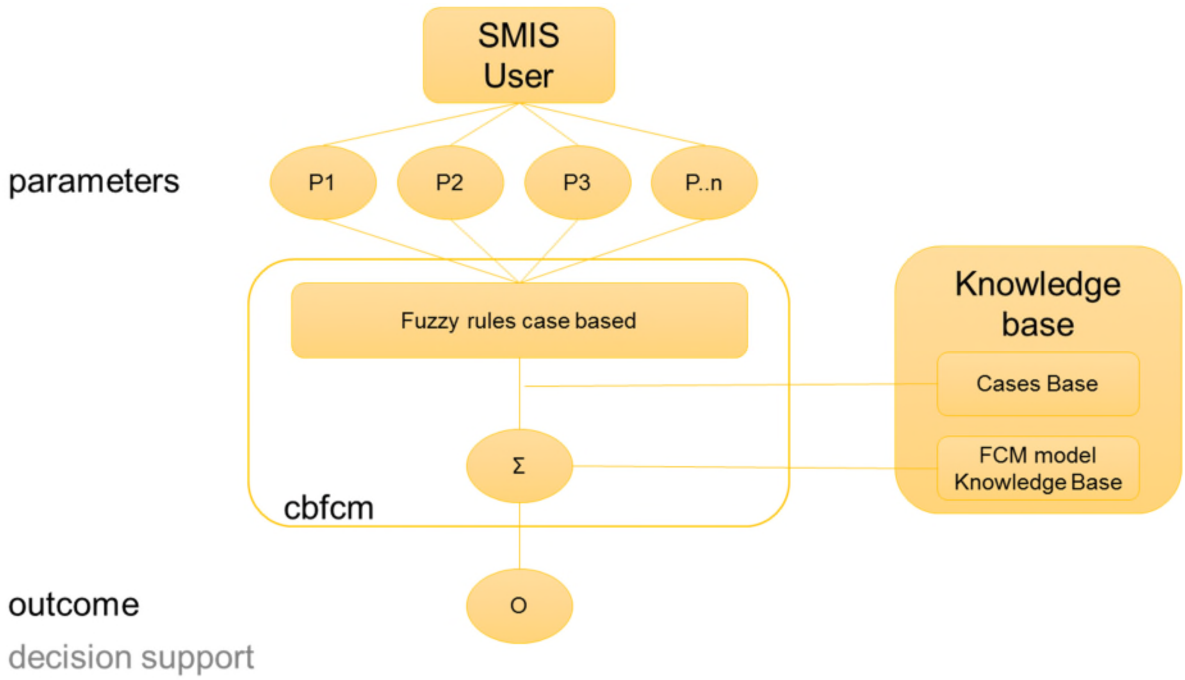

Figure 23. SMIS Architecture

7.3. Switchboard to integrate different data/ information/ knowledge sources

In order to achieve the 'mapping' between, and integration of, the different knowledge sources (Section 3 – 6), we have developed a high level meta-structure, to act as a '*switchboard*'. This permits each data / information / knowledge source to be classified and mapped, so allowing its integration and juxtaposition with other sources. The central function of this structure is to permit navigation between distinct data sources. When fully implemented in the final SMIS system, this will to permit users to identify, for any specific piece of information (e.g. soil management practice), related items within the full range of data sources. The 'Switchboard' (Figure 24**)** will thus permit soil management issues (and their solutions) to be drawn from the variety of data sources and related to each other.

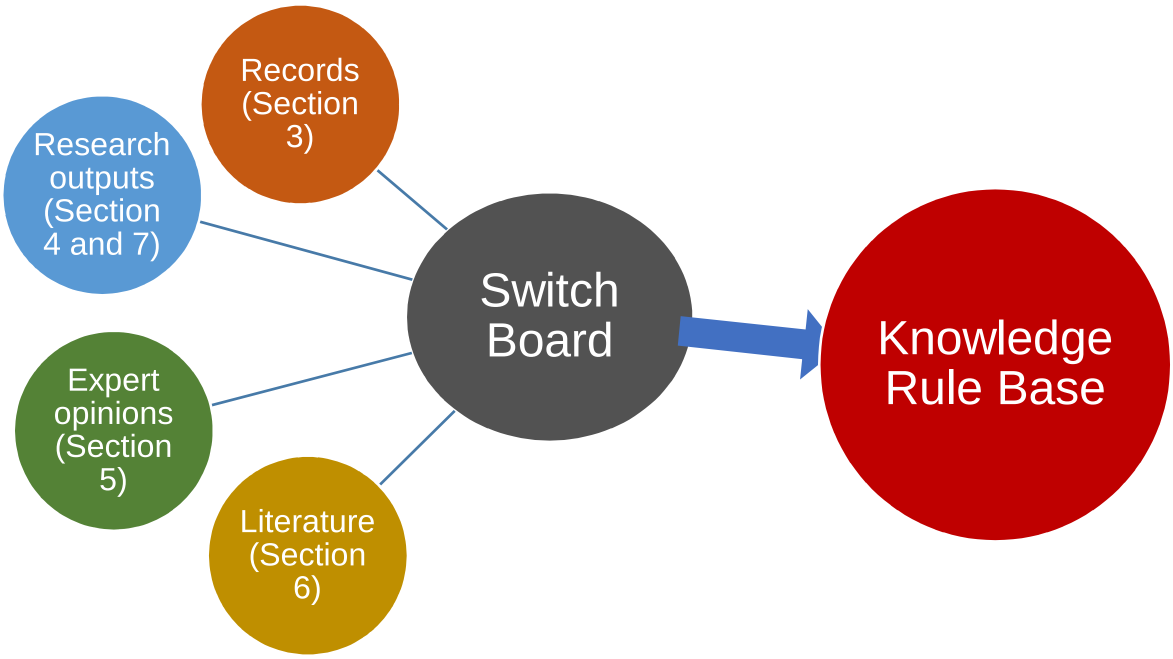

Figure 24. The SMIS Switchboard

7.3.1. Criteria and metadata

For the concept of the 'switchboard' to work, there needs to be a common descriptive form adopted referring to any piece of information held in the SMIS. Here the metadata is used as a 'key' to relate disparate sources of information in the SMIS. To achieve this, the high-level structure of SMIS has six elements, or '*criteria*', that are seen to encompass the range of conditions of interest to AHDB Horticulture. Each source of data / information / knowledge therefore undergoes a mapping of its own specific characteristics to this central 'switchboard', drawing on these meta-criteria. A series of six common criteria have been selected as representing the best balance of detail and practicality (Figure 25). For each of the six criterion, a series of subset 'criteria' are shown (note that these are not exclusive at this stage of development). Therefore, by example, when a source of information for SMIS is considered, a geographical 'place' may be extracted from it, for instance in the form of a field code, or an easting and northing. In the case study of grower data from Worth Farms (Section 3.3), this would be the National Grid coordinates (Table 3).

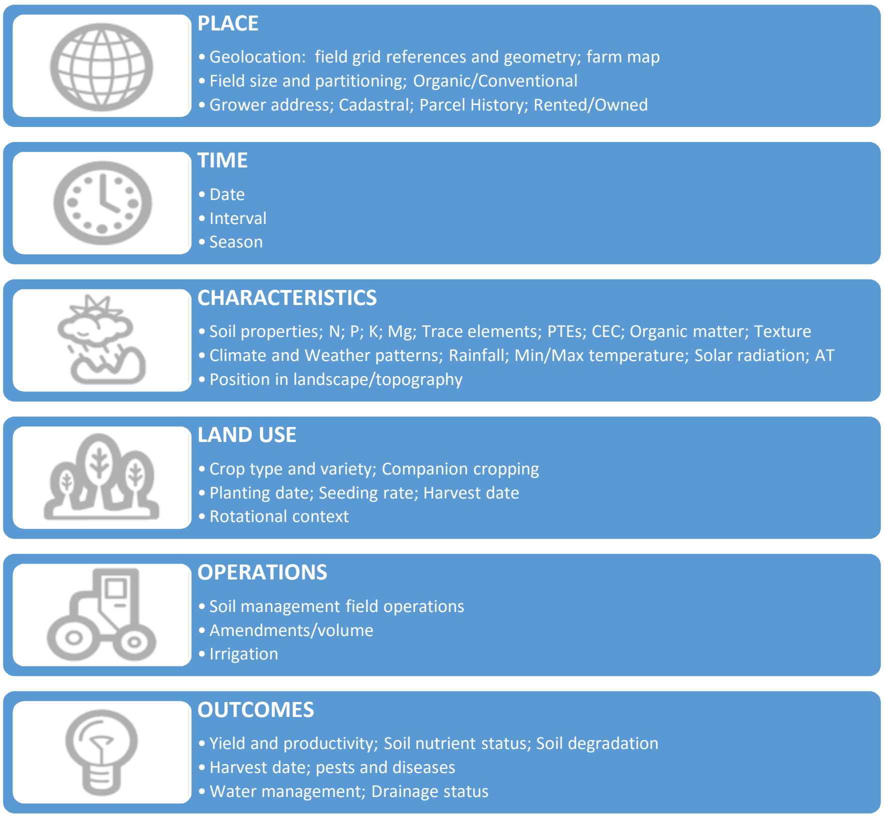

Figure 25. The meta-criteria adopted for SMIS

Another example to illustrate this is given below for an academic paper by Berg (1984) concerning maize cultivation (Table 21). Note that 'other' local meta-information is recorded alongside the 6 meta-criteria adopted from SMIS.

Table 21. Meta-data example for an academic paper

Attribute	Description
Citation	Berg, 1984
Criteria	*Place xxxx; Time xxxx; Characteristics xxxx; Land use xxxxx; Operations xxxx; Outcomes xxxxx*
Application	Agriculture (maize)

Treatment	Straw at three application rates and a control. 0.6 mt ha^{-1}, 1.2 mt ha^{-1}, 2.2 mt ha^{-1}
Media	Setting: Field; Corn furrows (study 1 and 2), Bean furrows (study 3 and 4). Slope: Study 1: 4 %, Study 2: 1.5 %, Study 3: 7 %. Soil type: Erodible silt loam. Reps: 3. Rainfall: Irrigated 9.5 l min^{-1}. Measuring: sediment concentration, infiltration and lateral wetting. Sampling: 1 l samples top and bottom with initial steady runoff then at non-specified regular intervals.
Results	Straw decreased depth of furrow erosion, increased lateral furrow wetting and improved infiltration by 50 %. Sediment was eliminated on the first irrigation, and even retained 10 to 30 % of eroded sediment from furrows above the plots. Straw length; short straw tended to float down the furrows and bunch up when the water was applied. Longer straw tended to stick into the sides of the furrow and become embedded when water was applied. Straw may stay in place better if it were pressed lightly into the furrow after application.

7.3.2. Linking data

To permit SMIS to link together the source-specific data elements, across all sources of data, a set of overarching meta-criteria are used (Figure 25). Where the source is a spreadsheet, this may mean, for example, relating one or more columns of data to a given criterion (e.g. spreadsheet columns ‘a’ to ‘d’ relate to the locational ‘Place’ criterion).

Each given data source will have a range of data items stored pertaining to its specific application and requirement; these data can arise in a range of forms and formats. For SMIS it is critical to recognise this range and, by implication, the heterogeneity of data sources in SMIS that will be used to serve and inform the e-Guide. For each of these data ‘schema’, common links will be established to the criteria of the switchboard as described.

Once all data / information / knowledge arising from each of the sources adopted is mapped to the switchboard criteria, source-spanning cause and effect factors can be investigated to develop a ‘rule base’.

7.3.3. Linking criteria to specific content

The SMIS contains a common reference framework to permit the alignment of sources of data, knowledge and information. Where data from disparate sources on field operations require a coherent operation, a common data structure is required. Appendix 9 presents further details. Contextual single criteria data deriving from an independent source is also collected as background information, e.g. a detailed field soil map, or a set of on-farm meteorological conditions/records.

The data sources contain additional data elements over and above these six criteria, e.g. data that cannot be 'mapped'. This is to be expected and reflects the scope of the SMIS project, as well as the diversity of data from specific AHDB projects as a function of the project specific objectives.

The means to '*drill down*' into specific content is inherently different from data source to data source. Figure 26 shows, schematically, how in SMIS different approaches are used. For example, the SMIS knowledge base itself is encoded in as a 'Resource Description Framework', or RDF[2], a formal specification approach designed as a metadata data model (World Wide Web Consortium; W3C).

[2] See https://www.w3.org/RDF/ and http://www.w3.org/TR/rdf11-concepts/. RDF uses the concept of 'graphs', containing sets of subject-predicate-object triplets.

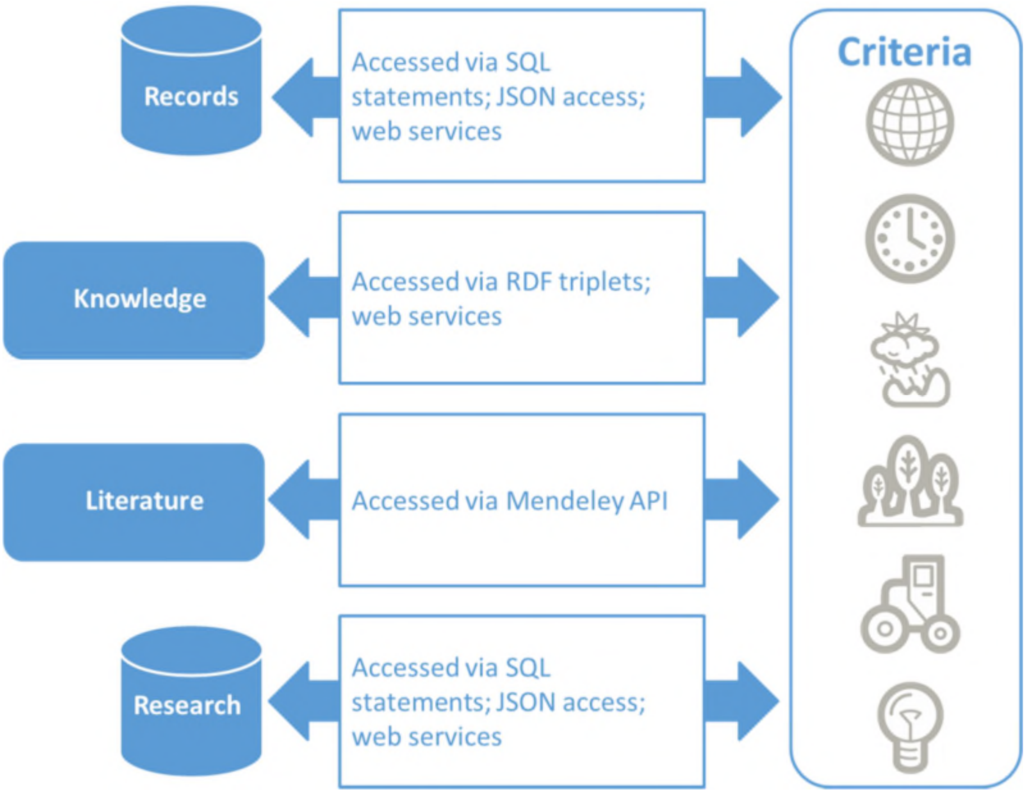

Figure 26. Linking different data sources

Any data source (for example a record describing an academic paper) has linkages to the expressed criteria. Some records have complete linkages across all criteria, and some only partial links to a subset of criteria. In the SMIS project, for a given item, the sum of each of these linkages is termed an '*arrangement*'. Figure 27 shows in columns the six criteria, and in rows beneath entries for four fictional items (for example four academic papers). Each arrangement will relate to its particular source record. The 'dots' shown in Figure 27 are symbolic only here – in reality, there will be a range of data captured for each source record. As well as a means of linking the diverse data sources, we will use these arrangements in the e-Guide to aid a graphical visualisation for the user in observing linked data.

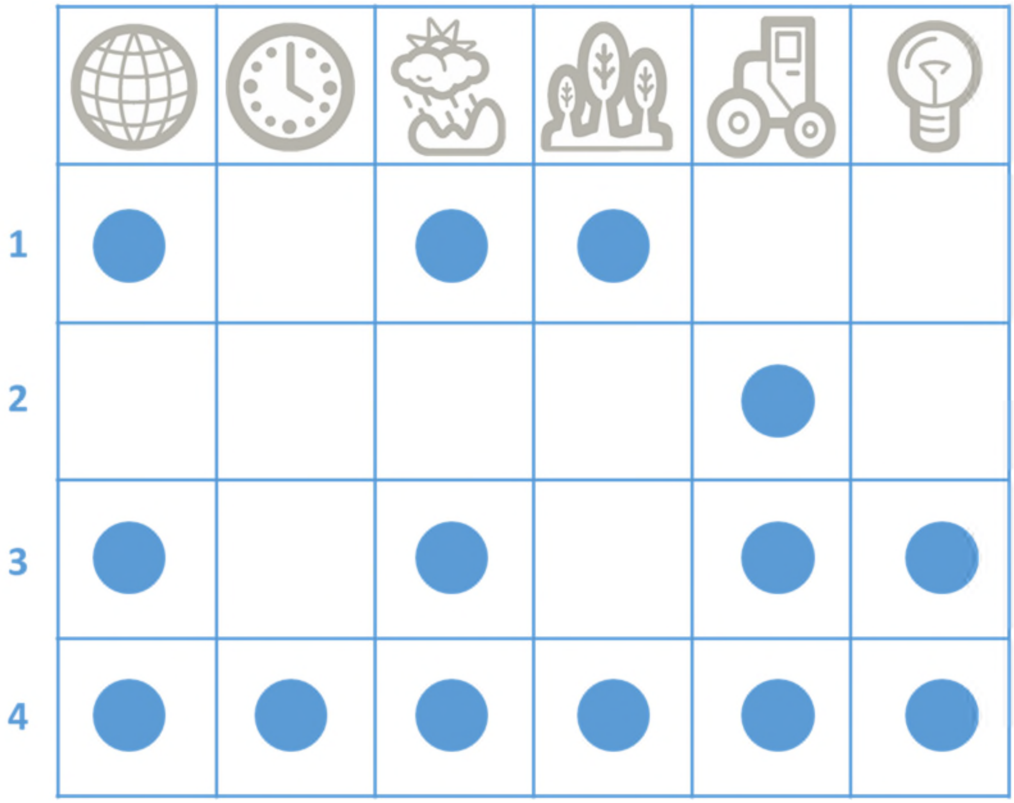

Figure 27. Arrangements (Four shown by example)

7.4. The SMIS rule base

The development and encoding of a 'rule base' is key to the methodology followed in the project. The rule base establishes the relationships between the intrinsic factors that affect soil condition (e.g. site conditions, soil type, weather) and best management practice for soil management operations, together with anticipated outcomes of given actions (for all SMIS data sources). The soil compaction case study in vining peas (Section 3.3) can be used as a simple example to illustrate this, where there will be a 'rule' regarding when cultivations can be undertaken without incurring soil compaction (i.e. when the soil is not too wet; Section 3.3.3). Once the data sources are inputted to SMIS (using the common criteria framework), the rule base will be established from all sources of data / information and knowledge as captured in SMIS.

7.5. Preliminary Use Case

The following section presents an anticipated use of the SMIS and e-Guide Toolkit, with each step being described in successive 'phases'. The user (e.g. an AHDB Horticulture employee) is advising a client (e.g. a grower) who has presented a scenario concerning soil compaction, wanting to know what management options there are to address this concern.

Phase One

A query is made where the specific site / field factors are entered. This might include the location of the field in question and the soil type, together with other intrinsic factors. By entering the location, Cranfield University's Land Information System, LandIS can be consulted. This will reveal a range of related information about the site (Figure 28).

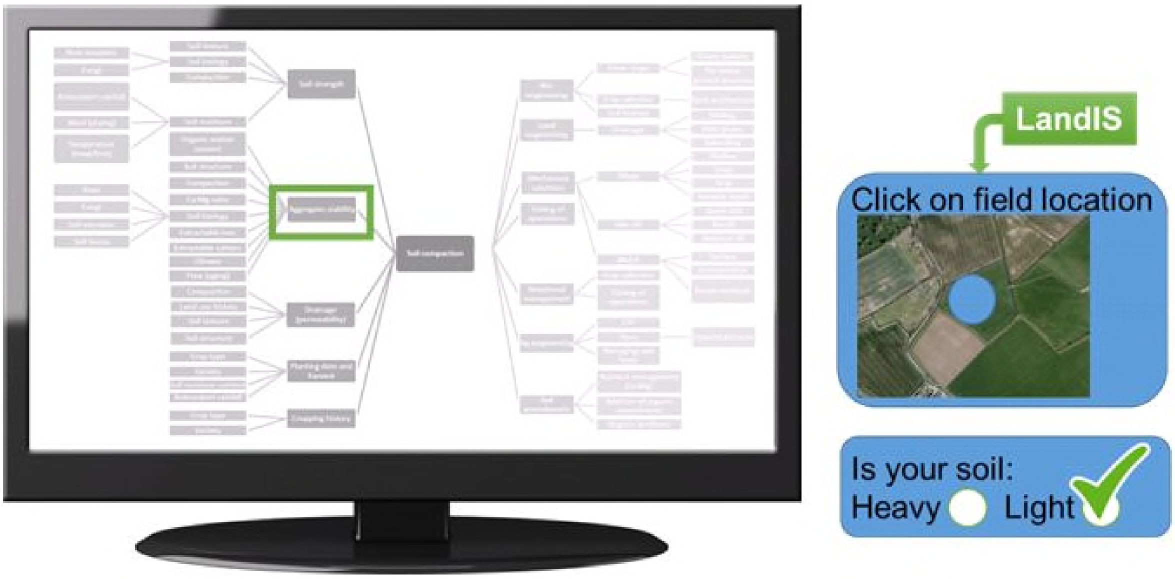

Figure 28. Use Case - Phase One – identifying the intrinsic conditions of the field(s)

Phase Two

A context for the query is then entered, related to the soil issues being presented – in this case soil compaction (Figure 29).

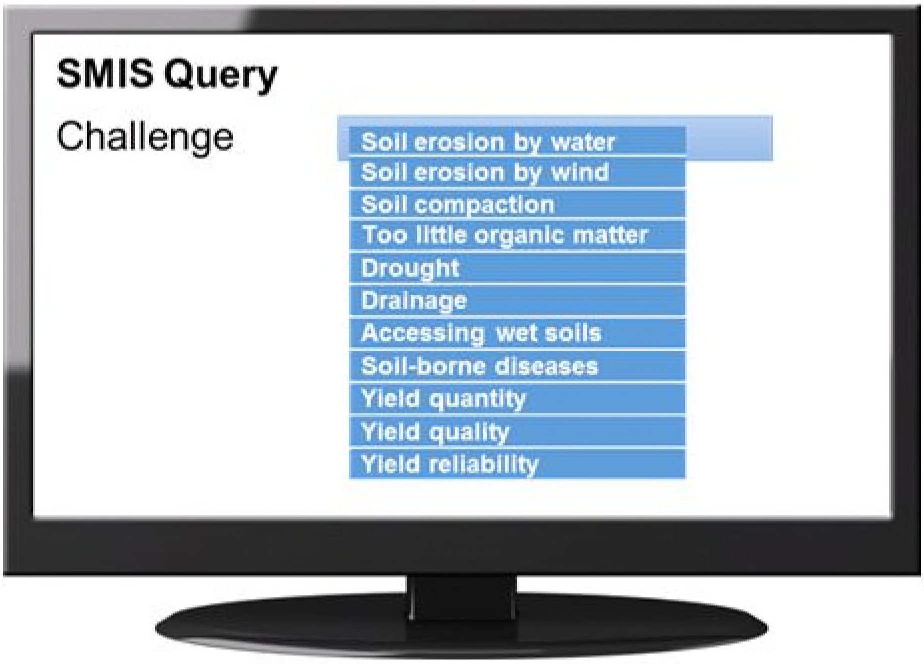

Figure 29. Use Case - Phase Two – selecting the soil management issue to be addressed.

Phase Three

All the linked sources of information that have been collated, integrated and stored in SMIS are interrogated. A report is then generated, presenting the results of querying the SMIS. This will include similar incidences of the soil management issue for similar field conditions, the methods that were used by others to mitigate the issue and the likely success of using these practices, given the site specific conditions under consideration (Figure 30).

A similar decision guide has been developed for fungicide planning for the control of Sclerotinia – the user has to answer several questions on weather, crop parameters, rotation, disease observations etc.to ascertain periodic Sclerotinia risk (% of crop affected) - https://cereals.ahdb.org.uk/hgca/afd/test.html. There is also an on-line potato cyst nematode (PCN) Pallida calculator (Figure 31), which allows the user to move around the various input tabs and so demonstrate "what if" scenarios (https://potatoes.ahdb.org.uk/online-toolbox/pcn-calculator#sthash.nT0V13zm.dpuf). This model can be updated with new information as it becomes available – the same principle applies with updating SMIS as and when new data becomes available.

It should be noted that the details of the three phases are addressed in Years 2 and 3 of the project (The eGuide construction will begin in August 2017; Appendix 1).

Figure 30. Use Case - Phase Three – generating a report of suitable soil management options for the field(s) in question.

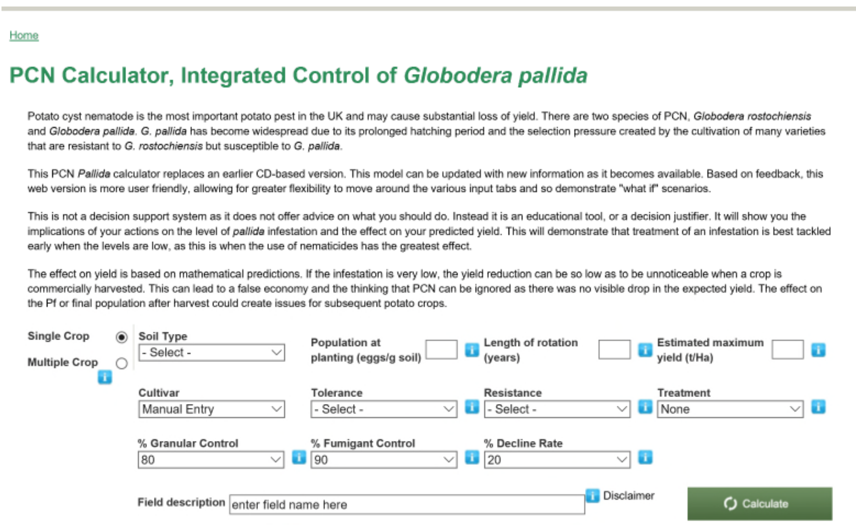

Figure 31. The on-line PCN Pallida calculator (https://potatoes.ahdb.org.uk/online-toolbox/pcn-calculator#sthash.nT0V13zm.dpuf)

.

8. Discussion

To date, a substantive data and information gathering exercise has been undertaken, focussing on the soil system and horticultural crop best management practices. As demonstrated in Section 3.1, this includes valuable information on the timing, type and frequency of a whole range of farming operations that are directly/indirectly related to addressing key soil management challenges.

Grower IP and often competitive advantage (perceived or otherwise) is tied into how [through experiential trial and error and their inherent knowledge and expertise] growers manage their farming system to optimise yield and yield quality. Obtaining grower data is built on trust and the ability to demonstrate tangible benefits. This level of trust has already been established with the development of a project date-pipeline. Pea Grower Groups and associated growers are a robust source of cross-rotational whole farm data as they are intensely aware of the need to manage their soils (with a focus on compaction and soil borne diseases) to optimise pea yields and yield quality. Critically demonstrated in the Worth Farms Case Study, these whole-farm datasets provide access to a comprehensive set of farming operations that when combined with Met Office data and LandIS soil series data can generate useful derived metrics of compaction and slaking risk that can be incorporated into the SMIS e-Guide.

It is important to note that the whole farm cross rotational datasets utilised in the study have the potential to provide benefits to levy payers across AHDB sectors namely Horticulture, Cereals & Oilseeds and Potatoes. In addition, non-AHDB levy boards and sectors include Maize Growers Association (MGA) and the British Beet Research Organization (BBRO).

Academic and practitioner literature has been collated from Rickson and Deeks et al. (2013), and this literature has been updated using project specific search criteria applied in Scopus, an academic paper search engine. The latest literature has been screened for relevance to horticultural soil management issues, in preparation for SMIS record generation based on the overarching SMIS meta-criteria; place, time, characteristics, land use, operations and outcomes. Relevant research projects from Defra, AHDB and BBSRC projects have been identified for data inclusion into SMIS.

The challenge of linking these different formats of data and information has been addressed by creating a switchboard that allows harmonisation of data going into SMIS database. As well as quantitative data, one particular challenge is how to express the more qualitative outputs of research outputs (e.g. CP107b) in the SMIS database. This will be addressed in Year 2 of the project.

Interrogation of the current data base using data mining techniques has generated outputs about best management practices to avoid soil degradation and enhance soil productivity. In this way, financial benefits to levy payers will result primarily from the better access to existing and state-of-the art knowledge relating to the optimisation of soil management on a soil type and crop specific basis but also critically on a cross-rotational whole farm basis. The ability of levy payers to minimise and incrementally reverse soil degradation will in turn have direct benefits to yield and yield quality. For example, minimising and reversing soil compaction will optimise the ability of soils to receive, retain and release water. In turn, this will;

- extend trafficability windows thus impacting on the timing of field operations [allowing growers to reduce compaction risk]
- decrease risk of soil borne disease associated with high/low soil moisture status impacting on yield and yield quality [optimising marketable yields]
- increase irrigation use efficiency [reducing quantify/frequency of irrigations]
- improve soil 'workability' [reduced fuel/operator costs]
- reduce erosion risk [GAEC Rule 5]
- enhance ability of growers to meet customer requirements in a timely manner [maintaining customer confidence and business relationships]

8.1. Next steps

Appendix 1 identifies the key activities and milestones for the project in Years 2 and 3. These will include:

- Continue to source data / information / knowledge as input to SMIS database. Over the next 6-months the existing cross-rotational whole farm Gatekeeper data from participating HMC Peas Ltd growers (Section 3.1) will be brought into SMIS. The datasets obtained from Caley Farms, Jack Buck Farms and Hay Farming will be screened and variables not obtained during the initial extraction of XML files will be gathered in January 2017. This data will be cleaned, transformed and combined with the Worth Farms data to generate a combined cross-rotational dataset of circa 3,000 ha, covering over eight cropping years (2009-2016).
- As just one example of soil management issues and their control, data from these growers will be combined with Met Office data and LandIS Soil Series data to generate suitable Machinery Work Days (MWD), Topsoil Slaking and Compaction susceptibility and Natural structural regeneration potential maps for the soils associated with these growers' landbanks.

- As SMIS outputs are further developed, these will be used to demonstrate system functionality and bring in data from addition substantive growers/grower groups. It is anticipated that initial data suppliers such as Richard Fitzpatrick of HMC Peas and Simon Day of Worth Farms can be used of advocates of SMIS thus widening the data pool. In addition, based on feedback from Becky Ward of PGRO focus in 2017 will be placed on obtaining grower group data from Ian Watson (Stemgold,), Andrew Lensen at Velcourt (Kings Lynn) and Stephen Francis at Fen Peas (Boston). Now that tangible outputs have been generated by linking grower, Met Office and LandIS data sets, one-to-one meetings will be held with agri-businesses associated with the CP107d Project Advisory Group to source additional data to feed into the project data-pipeline.
- Other sources of raw data include relevant AHDB Horticultural projects identified from the last 10 years. Further relevant research projects will be identified from other funding organisations including BBSRC, NERC.
- Continue to review and assess the academic and practitioner literature around soil management issues and their mitigation to ensure more sustainable horticultural production. More specifically, this means undertaking Phase 2 analysis (Section 6.2) on the peer reviewed literature collated in the Rickson and Deeks (2013) report and published between 2013-16.
- Issues of data integrity, reliability and accuracy will be addressed as new sources are incorporated into SMIS. This will include how to manage missing data – is it possible to use proxy data instead? For example, soil bulk density measurements are not always available (as an indicator of soil compaction). However, knowing the soil type, weather conditions and machinery used, the risk of soil compaction can be estimated.
- The large datasets within SMIS will require the use of complex data management techniques and advanced computational skills. We will continue to develop analytical methods and statistical modelling, drawing across the body of data assembled, allowing comparative assessment and benchmarking against available grower and case-study data.
- A rules-base for functional relationships between data members will be established based on expert opinion, established AHDB guidance documents and weight of evidence in the literature. This will form the basis of a suite of expert knowledge and hypothesis driven statistical analyses. However (and critically), as new whole farm cross-rotational datasets are brought into SMIS [and linked with other SMIS datasets e.g. to LandIS, AHDB Projects, climatic data etc] naïve statistical approaches can be used to tease out 'what we don't know'. These outputs will then be presented to the industrial panel and academic

experts to verify the validity of the observed causal relationships. These statistical outputs will in association with outputs derived from 'expert knowledge' and 'hypothesis driven' statistical approaches to feed into the e-Guide and form the basis of guidance to growers. As the SMIS system develops the robustness of these statistical outputs will increase.

- As the database within SMIS develops, it is envisaged that data mining techniques will provide useful insights to address AHDB Horticulture Panels' 2015-2018 priorities. Statistical interpretation of grower datasets within SMIS will provide a more scientific basis for guidance on a wider range of soil management issues (only soil compaction has been considered thus far). These issues include (and are not limited to):
    - the selection and role of cover crops (GAEC Rule 4) in a range of rotational contexts to address key soil borne diseases affecting the horticultural sector through 'bio-fumigation',
    - enhancement of water and nutrient use efficiency through promotion of Arbuscular Mycorrhizal Fungal associations with direct impacts of yield and yield quality as well as reduced reliance on a diminishing range of chemical solutions.
    - soil structural improvement via 'bio-drilling' thus reducing reliance on costly and often ineffectual mechanical options
    - Sclerotinia: Vining Peas: Development of improved control strategy as disease becoming more frequent
    - Root diseases, including Fusarium Foot Rot [Utilizing PGRO Foot Rot Index data]: (An increasing problem with no chemicals available). An evaluation of cultural methods of suppression is required and the use of mustard bio fumigant cover crops.
    - Crop Nutrition: More information needed on P and K requirements.

9. Conclusions

Sustainable soil management is needed to maintain / enhance crop yields (quantity and quality) whilst protecting environmental resources. We are information rich but knowledge poor with regard to soil management because of the complexity of how information and data are collected and reported. Sources of information and data related to soil management in horticulture are currently unstructured, uncentralised and difficult to find and/ or access. Collating, harmonising and querying this resource to provide better informed soil management advice requires tools that can harmonise the information and provide easy access to the underlying knowledge.

This report provides a summary of the first 12 months of work on the three year 'Horticultural Soil Management Information System' Project (CP107d), aiming to develop a Soil Management Information System (SMIS) which will provide users with a set of robust, empirically-based, best-practice guidelines, and the likely consequences of applying them.

To date, robust progress has been made on all Objectives of the research project (Figure 32).This report has elaborated the current state of play for each of the four key project phases, namely: (1) resource review; (2) data handling and validation; (3) analysis toolkit, and; (4) dissemination (development of the e-Guide). The principal areas of work in the past 12 months have included inputs to the requirement analysis, the systems analysis and the literature review. The preparation of case studies has commenced, with indicative output from this included in the report. Some early indication of the statistical inference modelling approaches under consideration is also included.

- Four keys sources of information pertaining to soil management have been assembled: (1) grower records systems such as gatekeeper; (2) AHDB-funded research project outputs; (3) published and grey literature; and (4) expert and practitioner knowledge.
- To draw together these disparate sources of information, and to provide an evidence base for future users of the SMIS and e-Guide, a structured meta layer has been assembled, informed using Fuzzy-logic Cognitive Mapping in the expert and practitioner workshops. This has identified a series of conjectures which the wider data can be used to support or reject. To assist the harmonisation of the different data sources, a switchboard design will provide the means to allow navigation within this body of knowledge.
- Data mining techniques have been used to query the information base to extract the key drivers of soil degradation and the effectiveness of mitigation measures to alleviate the problem.

- To demonstrate the feasibility of this approach, a case study concerning vining peas and soil compaction is presented. This has allowed consideration of how the causes and effects of particular soil management issues can be related to real-world data, as well as literature, expert opinion and research project outputs.
- There is a need to continue to source additional data from growers and researchers, as they become available.
- The foundations have been laid to receive data and to enable its categorisation and harmonisation.
- SMIS can provide a knowledge exchange pathway to promote the adoption of robust metrics of soil health [soil quality indicators] currently not routinely measured by growers. Going forwards this is critical to sustainable soil management in the UK.
- We conclude that the development of a soil management information system is a timely and effective means to provide AHDB and its levy payers with an evidence-based source of knowledge concerning soil management practices and consequences.

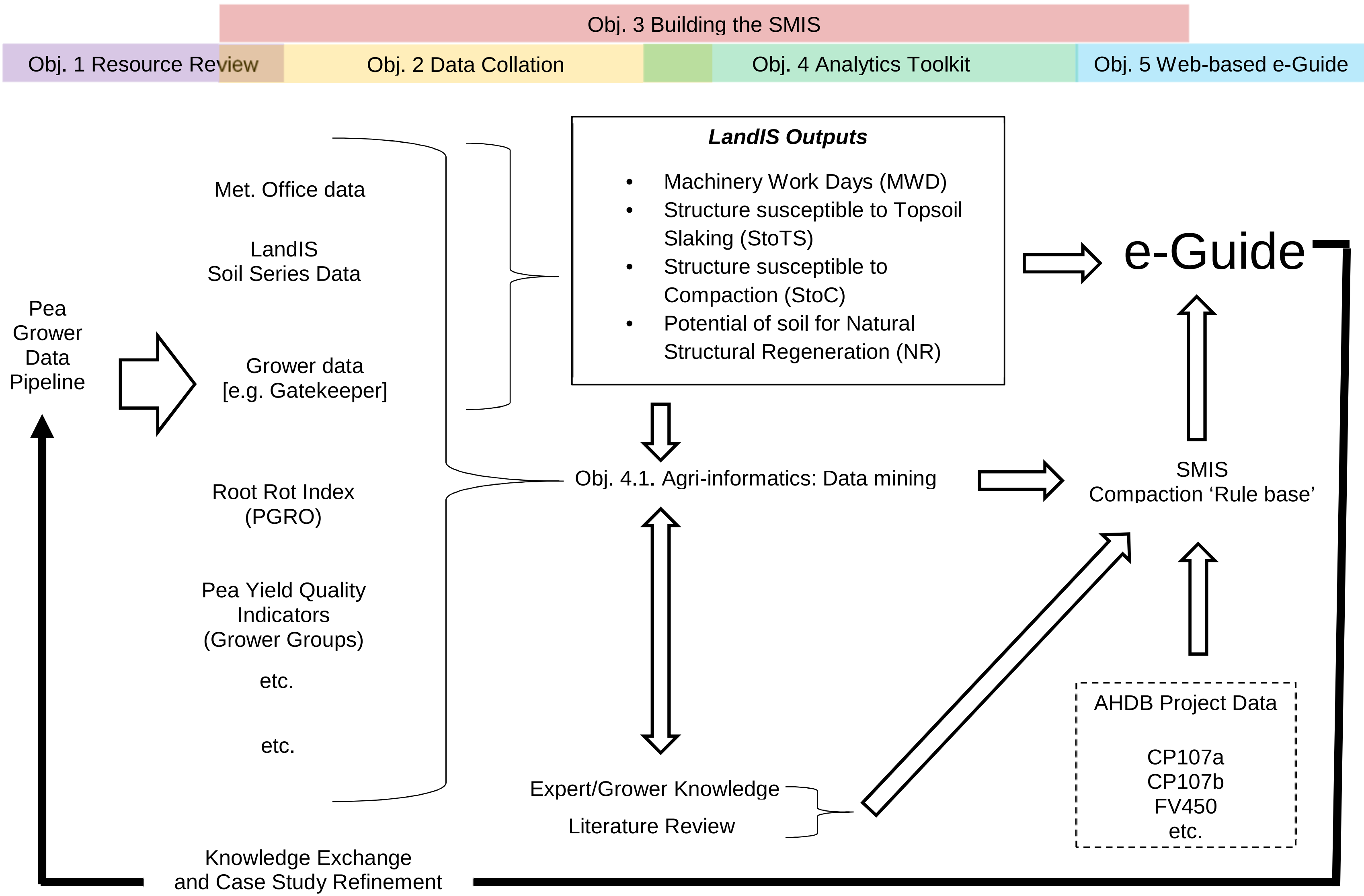

Figure 32. Using the case study of soil compaction in peas to demonstrate the stages of the project, including use of data, information and knowledge sources to build the SMIS rule base to inform the eGuide of best soil management practice.

10. Knowledge and Technology Transfer

This section describes the knowledge exchange activities carried out that relate to the SMIS project. These activities were carried out by members of the project team in conjunction with Dr Lynda Deeks, NERC Horticulture Knowledge Exchange Fellow.

PGRO Grower Groups

PGRO provides a direct knowledge exchange pathway to over 500 individual growers and 11 grower groups [plus Green Pea who provide to Birds Eye, who operate largely independently of PGRO]. This KE pathway was already generated four whole farm Gatekeeper data sets. To fully optimise the potential of agri-informatics to drive the development of case studies within SMIS, a data pipeline has been developed through Pea Grower Groups (as supported by PGRO).

AHDB Soil Research & Knowledge Exchange Workshop, (Rothamsted Research), 7th April 2016

The overall objective of the workshop was to provide an opportunity for AHDB levy payers, staff and project leads to learn more about AHDB's portfolio of soils R&KE. The aim was to foster collaboration and ultimately increase the impact of the research. At the event, Jane Rickson presented the aims and objectives of CP107d, and the methodology used to develop SMIS. All participants were asked to supply knowledge / data on soil management practices and their outcomes as input to SMIS.

British Herb Trade Association field event (Millets Farm, Oxfordshire, 25th August 2016

This was in response to a request by Grace Choto (AHDB) for AHDB CP 107 projects to help with knowledge exchange at the British Herb Trade Association (BHTA) field event. Specifically, Grace suggested that the CP 107d project members could dig a soil pit and explain profiles. Following discussions with Claire Donkin (British Herbs) and Tim Mudge (British Growers), Robert Simmons and Lynda Deeks set up an infield demonstration of the impact of rain splash on herb quality and also used a soil pit to engage the audience in a discussion on soil health related issues, including a discussion on earthworms, soil compaction and soil hydrology (infiltration and available water content).

To add continuity between the AHDB CP 107 projects we also co-ordinated with Martin Wood (CP 107b – GREATsoils) so that he could tie in some of their observational messages with what we were demonstrating and discussing.

The event also afforded an opportunity to discuss with a cross section of this sector their main concerns with regards to management of soil health and the type of data they were collecting themselves and what they felt they needed from knowledge exchange going forward.

Our activities at this event also received a mention in Horticulture Week, 16th September 2016, which covered the story of 'Herbs field day covers soil issues'.

British Carrot Growers Association - variety day and trade exhibition (West Knapton, Yorkshire, 6th October 2016)

Following on from the success of the BHTA field event, Robert Simmons and Lynda Deeks were invited by Tim Mudge to attend the British Carrot Growers Association (BCGA) variety day and trade exhibition, in order to engage growers in discussion about soil, soil management and soil health. The field demonstration, which involved opening up a cross section of soil across the ridge and furrow of the carrot crop was used to engage attendees in a discussion about the soil system and their own current management practices (where applicable).

The event open up an opportunity to discuss the type of data this sector were collecting and to understand their perceived soil related issues. We were able to spend time talking with Rodger Hobson (land owner of event site and chair of BCGA) and Andy Blunt (Farm Manager) getting their opinions on soil related issues.

Again, we co-ordinated with Martin Wood (AHDB CP 107b - GREATsoils) as we had found it was more effective for Martin to engage people regarding soil measurement options when they were already engage in a wider discussion about soil. The soil demonstration soil pit also attracted people into a discussion where as people were more hesitant to stop at a stand to engage in conversation.

Our activities at this event got a mention of twitter "Looking at soils for #carrots at the variety trials day"

GREATsoils workshop: Soil Health and Farm Viability (Lichfield, 22nd November 2016)

Lynda Deeks was invited by Martin Wood and Simone Osborn (Soil Association) to give a presentation on soil erosion management in relation to soil health at the GREATSoils workshop looking at Soil Health and Farm Viability. While this primarily supported the knowledge exchange of AHDB 107b it also offered an opportunity to discuss with another group of people the type of soil information they were collecting and how they engaged in knowledge exchange. The later ranged between groups who simply 'chatted via email' to slightly more structured forum groups. While this event only attracted a relatively modest group (*Ca*. 10) it enabled a more in-depth discussion with individuals.

Carrot Expert Group Meeting, (Croxton Park, Thetford, Norfolk, 24th November 2016)

Jane Rickson was invited by Tom Will (Vegetable consultancy Services), to present at the Carrot Expert Group Meeting on ‘Understanding the biotics and opportunities to improve soil health’. The group, which consisted of 12 agronomists representing UK, Denmark, Norway, Sweden, Germany, Holland and France, afforded a good opportunity to discuss different approaches and understanding in relation to soil health across a range of European countries with similar growing conditions to the UK.

Vegetable Consultants Association Annual Conference (Stilton, Peterborough, 29th November 2016)

Robert Simmons and Lynda Deeks were invited by Phillip Effingham (Greentech Consultancy Ltd and Chairman of the British Vegetable Consultants Association), to present to a group of 25 professional vegetable consultants, on ‘Challenges of Sustainable Soil Management: Metrics of soil health’. We used the interactive meeting to draw out knowledge from the group with regards indicators of soil health that they felt were commonly used. Discussion also turned to the potential power of the huge amount of data being collected by individuals and the need for a better way to utilise this resource more effectively. Also raised were issues related to the most effective way of collecting data that was informative, easy for an individual to compile but avoided individuals having to repeat data input (i.e. prevented duplication of information already being asked for elsewhere).

HMC Grower Group (February 2017)

The outcomes of the initial analyses undertaken on the Worth Farms data along with the LandIS derived maps for all the growers will be presented to the HMC Grower Group (Pers. Communication. Richard Fitzpatrick Grower Group Manager). It is expected that this will act as a catalyst for other growers to provide their Gatekeeper data.

AHDB Legume Panel (early 2017)

The outcomes of the initial analysis of Gatekeeper data (and ancillary supporting databases) will be presented to the AHDB Legume Panel in the form of a 2-page summary in early 2017 (Pers. Communication Becky Ward). It is anticipated that this will result in further data being obtained from additional grower groups. Further, in November 2017 and verbal presentation will be made to this panel (Pers. Communication Becky Ward).

11. Glossary

The project embodies a great many terms and concepts for which there needs to be a common reference and understanding. The following table provides working definitions for the common terms and concepts used in the development of SMIS. This will be updated as the project progresses.

Term	Definition
AHDB	Agriculture and Horticulture Development Board.
Arrangement	Any given item of data or information considered (e.g. an academic paper) may have one or more entry for each of the six descriptive criteria. Some items may have information for all six criteria, and some may have entries for only a few criteria.
Conjecture	An expert opinion as to both the inherent factors causing a given soil-based challenge, and/or the appropriate management solutions that can be deployed to address the same challenge. In the project, these derive from the FCM process.
Criteria	One of six metadata descriptions that can be completed to describe a given source of data. These comprise descriptions concerning place; time; characteristics; land use; operations; and outcomes. Typically this is recorded with an entry in a table for each item considered (e.g. an academic paper).
e-Guide	Knowledge-Based System for presenting options, outcomes and best practices for soil management with relation to horticultural practices. A key project delivery.
FCM	Fuzzy-logic Cognitive Mapping – a technique used to elicit expert opinion in a structured form, often used in workshop contexts. Used in the project to capture the rulebase of conjectures.
NLP	Natural Language Processing – a set of computational approaches to extract information out of a corpus of text.
Project	Project "Managing soil health for sustainable horticultural soils. CP 107d/3110107425", sponsored by AHDB under the CP 107a: Soils - Improved Sustainable Management for Horticultural Crops programme.
RDF	Resource Description Framework (https://www.w3.org/RDF/ and http://www.w3.org/TR/rdf11-concepts/), a means of recording semantic knowledge in computer-compatible form, using the concept of 'graphs', containing sets of subject-predicate-object triplets.

SMIS	Soil Management Information System. An information repository that contains a rule base, and supporting evidence from a range of sources. A key project delivery.
Soil Management Challenge	One of the soil management challenges identified by Rickson and Deeks (2013) that the SMIS sets out to address, identifying thematic areas of concern in the development and implementation of best practice guidelines for sustainable soil management. An example is 'soil compaction'.
W3C	The World Wide Web Consortium (https://www.w3.org), owners of the RDF schema.

12. References

Archer J.R. and Smith P.D. (1972). The relation between bulk density, available water capacity and air capacity of soils. Journal of Soil Science 23:475-480.

Askari M.S., Cui J. and Holden N.M. (2013). The visual evaluation of soil structure under arable management. Soil & Tillage Research 134:1-10.

Ball B.C., Campbell D.J., Douglas J.T., Henshall J.K. and O'Sullivan M.F. (1997). Soil structural quality, compaction and land management. European Journal of Soil Science 48:593-601.

Batey T. (2009). Soil compaction and soil management – a review. Soil Use and Management 25:335-345.

Berg, R.D. (1984). Straw residue to control furrow erosion on sloping, irrigated cropland. Journal of Soil & Water Conservation 39.58–60.

Bosma, C., Glenk, K. and Novo, P. (2017). How do individuals and groups perceive wetland functioning? Fuzzy cognitive mapping of wetland perceptions in Uganda. Land Use Policy 60:181-196.

Carron, M.P., Pierrat, M., Snoeck, D., Villenave, C., Ribeyre, F., Suhardi, Marichal, R. and Caliman, J.P. (2015). Temporal variability in soil quality after organic residue application in mature oil palm plantations. Soil Research 53:205–215.

Chamen, W.C.T. (2006). Controlled traffic farming: Literature review and appraisal of potential use in the U.K. Home-Grown Cereals Authority Research Review No. 59. Kenilworth, U.K.: Agriculture and Horticulture Development Board.

Chen, G., and Weil, R.R. (2011). Root growth and yield of maize as affected by soil compaction and cover crops. Soil and Tillage Research 117:17-27.

Christen, B., Kjeldsen, C., Dalgaard, T. and Martin-Ortega, J. (2015). Can fuzzy cognitive mapping help in agricultural policy design and communication? Land Use Policy 45:64-75.

Clark, L.J., Whalley, W.R. and Barraclough, P.B. (2003). How do roots penetrate strong soil? Plant and Soil 255: 93-104.

Cresswell, H.P. and Kirkegaard, J.A. (1995). Subsoil amelioration by plant roots - the process and the evidence. Australian Journal of Soil Research 33: 221-239.

Dadaser, F. and Özesmi, U. (2002). Stakeholder analysis for Sultan Marshes ecosystem: a fuzzy cognitive approach for conservation of ecosystems. In EPMR, Environmental Problems of the Mediterranean Region, International Conference, Near East University, Nicosia, North Cyprus, April 12–15, 2002.

Dorais, M. and Alsanius, B. (2015). Advances and trends in organic fruit and vegetable farming research. In Janick, J. (ed.) Horticultural Reviews. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 185–268.

Döring, J., Frisch, M., Tittmann, S., Stoll, M. and Kauer, R. (2015). Growth, Yield and Fruit Quality of Grapevines under Organic and Biodynamic Management. PloS one Public Library of Science 10(10).

Duiker, S.W. (2002) Diagnosing soil compaction using a penetrometer. Agronomy Facts 63. Penn State University Extension.

Evans, R. (1996). Soil erosion and its impacts in England and Wales. Friends of the Earth, London, UK 121pp.

Ge, T., Chen, X., Yuan, H., Li, B., Zhu, H., Peng, P., Li, K., Jones, D.L. and Wu, J. (2013). Microbial biomass, activity, and community structure in horticultural soils under conventional and organic management strategies. European Journal of Soil Biology 58:122–128.

Giordano, R., Passarella, G., Uricchio, V.F. and Vurro, M. (2005). Fuzzy cognitive maps for issue identification in a water resources conflict resolution system. Physics and Chemistry of the Earth 30:463-469.

Gonzalez Perez, P., Ye, J., Wang, S., Wang, X. and Huang, D. (2014). Analysis of the occurrence and activity of diazotrophic communities in organic and conventional horticultural soils. Applied Soil Ecology 79:37–48.

Gornall, J., Betts, R., Burke, E., Clark, R., Camp, J., Willett, K. and Wiltshire, A. (2010). Implications of climate change for agricultural productivity in the early twenty first century. Philosophical Transactions of the Royal Society B 365:2973-2989.

Graves, A., Morris, J., Deeks, L., Rickson, J., Kibblewhite, M., Harris, J. and Fairwell, T. (2011). The total costs of soils degradation in England and Wales. Final project report to Defra, SP1606. NSRI, Cranfield University, UK.

Groumpos, P.P. and Groumpos, V.P. (2016). Modeling Vineyards Using Fuzzy Cognitive Maps. In 24th Mediterranean Conference on Control and Automation (MED), Athens, Greece, June 21/24, 2016.

Grzesiak, S., Grzesiak, M.T., Hura, T., Marcińska, I. and Rzepka, A. (2013). Changes in root system structure, leaf water potential and gas exchange of maize and triticale seedlings affected by soil compaction. Environmental and Experimental Botany 88:2-10.

Hamza, M.A. and Anderson, W.K. (2005). Soil compaction in cropping systems: A review of the nature, causes and possible solutions. Soil and Tillage Research 82:121–145.

Hodgson, J.M. (1976). Soil Survey Field Handbook. Soil Survey Technical Monograph No.5, pp99. Harpenden, UK.

IPCC (2007) Climate Change 2007 synthesis report. IPCC, Geneva, Switzerland.

Jones, N. A., H. Ross, T. Lynam, P. Perez, and A. Leitch. 2011. Mental models: an interdisciplinary synthesis of theory and methods. Ecology and Society 16(1): 46. [online] URL: http://www.ecologyandsociety.org/vol16/iss1/art46/

Keay, C.A., Hallett, S.H., Farewell, T.S., Rayner, A.P. and Jones, R.J.A. (2009). Moving the National Soil Database for England and Wales (LandIS) towards INSPIRE Compliance. International Journal of Spatial Data Infrastructures Research 4:134-155.

Khadra, R., D'Agostino, D.R., Scardigno, A. and Lamaddalena, N. (2011). Down-scaling pan-European water scenarios to local visions in the Mediterranean: the Candelaro Basin case study in Italy. Journal of Water and Climate Change 2:180-188.

Kirkegaard, J., Christen, O., Krupinsky, J., and Layzell, D. (2008). Break crop benefits in temperate wheat production. Field Crops Research 107:185-195.

Liu, L., Sun, C., Liu, S., Chai, R., Huang, W., Liu, X., Tang, C. and Zhang, Y. (2015). Bioorganic Fertilizer Enhances Soil Suppressive Capacity against Bacterial Wilt of Tomato. PloS one Public Library of Science 10(4).

Mbele, M. and Masinde, M. (2016). Development of adaptive environmental management system: A participatory approach through fuzzy cognitive maps. In 2016 IST-Africa Conference, IST-Africa 2016.

Mueller L., Kay B.D., Hu C., Li Y., Schindler U., Behrendt A., Shepherd T.G. and Ball B.C. (2009). Visual assessment of soil structure: Evaluation of methodologies on sites in Canada, China and Germany. Part I: Comparing visual methods and linking them with soil physical data and grain yield of cereals. Soil & Tillage Research 103:178-187.

Natarajan, R., Subramanian, J. and Papageorgiou, E.I. (2016). Hybrid learning of fuzzy cognitive maps for sugarcane yield classification. Computers and Electronics in Agriculture 127:147–157.

Özesmi, U. and Özesmi, S. (2003). A participatory approach to ecosystem conservation: Fuzzy cognitive maps and stakeholder group analysis in Uluabat Lake, Turkey. Environmental Management 31:518-531.

Papageorgiou, E.I., Markinos, A.T. and Gemtos, T.A. (2011). Fuzzy cognitive map based approach for predicting yield in cotton crop production as a basis for decision support system in precision agriculture application. Applied Soft Computing Journal 11:3643-3657.

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. 64pp. NSRI, Cranfield University, UK.

Seymour, M., Kirkegaard, J.A., Peoples, M.B., White, P.F., and French, R.J. (2012). Break-crop benefits to wheat in Western Australia - insights from over three decades of research. Crop and Pasture Science 63:1-16.

Shen, Z., Wang, B., Lv, N., Sun, Y., Jiang, X., Li, R., Ruan, Y. and Shen, Q. (2015). Effect of the combination of bio-organic fertiliser with *Bacillus amyloliquefaciens* NJN-6 on the control of banana Fusarium wilt disease, crop production and banana rhizosphere culturable microflora. Biocontrol Science and Technology 25: 716–731.

Shepherd T.G. (2003). Assessing soil quality using Visual Soil Assessment. In: Currie, L.D. and Hanly, J.A. (Eds) Tools for nutrient and pollutant management: Applications to agriculture and environmental quality. Occasion Report No. 17. Fertilizer and Lime Research Centre, Massey University, Palmerston North. pp. 153-166.

Smolinska, U., Morra, M.J., Knudsen, G.R. and James, R.L. (2003). Isothiocyanates produced by Brassicaceae species as inhibitors of *Fusarium oxysporum*. Plant Disease 87:407–412.

Thomasson, A.J. (1982). Soil and climatic aspects of workability and trafficability. Proceedings of the 9th Conference of the International Soil Tillage Research Organization, Osijek, Yugoslavia, 1982. 551-7.

Tones, S., Collier, R. and Parker, B. (2004). Large narcissus fly spatial dynamics. Final project report to Defra, HH1747. ADAS, Exeter.

Van den Akker, J.J.H. and Canarache, A. (2001). Two European concerted actions on subsoil compaction. Landnutzung und Landentwicklung / Land Use and Development 42:15-22.

Van Vliet, M., Flörke, M., Varela-Ortega, C., Cakmak, E.H., Khadra, R., Esteve, P., D'Agostino, D., Dudu, H., Bärlund, I. and Kok, K. (2016). FCMs as a common base for linking participatory products and models. In: Gray, S., Paolisso, M., Jordan, R., Gray, S. (Eds) Environmental Modelling with Stakeholders, Springer.

Whalley W.R., Clark, L.J., Finch-Savage, W.E. and Cope, R.E. (2004). The impact of mechanical impedance on the emergence of carrot and onion seedlings. Plant Soil 265:315-323.

White, R.G. and Kirkegaard, J.A. (2010). The distribution and abundance of wheat roots in a dense, structured subsoil - implications for water uptake. Plant, Cell and Environment 33:133-148.

Appendix 1. Project deliverables and milestones.

Objective	Description	YEAR 1 2015/16				YEAR 2 2016/17				YEAR 3 2017/18				Milestone Date mm/dd/yyyy
		Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	
		Nov-Jan 2016	Feb-April 2016	May-July 2016	Aug-Oct 2016	Nov-Jan 2017	Feb-April 2017	May-July 2017	Aug-Oct 2017	Nov-Jan 2018	Feb-April 2018	May-July 2018	Aug-Oct 2018	
1	**Resource review**													
1.1	Requirements analysis													31/10/2016
1.2	Systems analysis													31/01/2018
2	**Data collation**													
2.1	Literature review													30/04/2017
2.2	Case study assembly													30/04/2018
3	**Building the SMIS**													
3.1	Systems construction													31/01/2018
3.2	Data loading and testing													30/04/2018
4	**Analytics toolkit**													
4.1	Statistical inference modelling													28/07/2018
4.2	Benchmark analysis													28/07/2018
5	**Implement a web-based SMIS e-Guide portal**													
5.1	e-Guide construction													28/07/2018
5.2	System testing and deployment													31/10/2018
5.3	Documentation and user guide													31/10/2018

Appendix 2. Requirements analysis

Intrinsic site factors / properties		Extrinsic factors / inputs		Outcomes / impacts (for better or worse!)		Consequences
Field Locations	Field grid references	Surface tillage	Direct drilling	Productivity	yield quantity	costs reduction
	Field shapefile		strip tillage		yield quality	increased £ margins
	Farm map		etc		yield reliability	soil quality
Field Information	Field size			Pests and diseases		water quality
	Cropping history			Weeds and volunteers		
	Rented/owned	Subsoiling		Water management	Drainage	
	Organic/conventional, etc.	Organic amendments	composts,		Water conservation	
Cropping Data	Crop, variety, planting date, seeding rate, programmed yield, actual yield, harvest date, etc.		mulches		waterlogging	
	Variety		green manures	Soil nutrient status		
	planting date	Buffer strips	in field	Soil degradation	Soil erosion	
	seeding rate		field edge		soil compaction	
	harvest date	Cover cropping			lack of organic matter	
Soil Analyses	N	Companion cropping	Product/type		loss of biodiversity	
	P	Fertiliser applications	date of application (both in-organic and organic), etc.		soil contamination	

	K			application rate,		Access on wet soils			
	Mg								
	Trace elements		use of alternative growing media						
	PTEs								
	CEC								
	organic matter								
	texture								
	Rainfall								
	min temperature								
On-farm weather station data	max temperature								
	sunlight hours, etc.								
Field Operations	Field operations (non soil management)								
Irrigation	Type								
	method								
	volume,								
	irrigated Y/N,								
Spray Operations	Product/type								
	application rate								
	date of application,								

Appendix 3. Capturing expert knowledge

Expert feedback on proposed SMIS

<u>Respondent 1</u>

Overall comment – analyses are thorough but a bit academic and do not always relate to the world as we find it.

Slide 1: Decline in Productivity

- Do you mean sclerotia or sclerotinia, if you mean to list all soil borne diseases that are legion…

Slide 2: Soil compaction

Axel/Axle

- Most compaction is caused by forced operations; usually harvesting too wet, peas in July, rigorous sales programmes, potatoes in late October, celeriac in late November. The management response is to have adequate harvesting resources for unusual times (easier said than done).

Slide 3: Soil erosion by water

- Timing is important in erosion management
- There is a conflict of winter cover crops and frost mould creation for spring crops.
- No knowledge of conditioners

Slide 4: Soil erosion by wind

- Timing is also important here.

Slide 5: Soil-borne diseases

- Alternative hosts for soil borne diseases; weeds, cover cops, bio fumigant crops (do mustard crops host sclerotinia?).
- Trap crops similarly
- Free living nematodes are a growing problem
- Irrigation management is important as it can spread disease

Most important soil management issues:

1. Seed bed quality –the most important moment in the life of a crop.
    a. Cloddy/poor establishment
    b. Too fine/capping

c. Too wet/compaction
d. Late ploughing/no frost mould. Capping and frost mould relate particularly to silt land.

All problems will vary on different soil types.

2. Remedial action and what action –how to measure compaction?
3. Soil borne diseases
4. Rotation

Respondent 2

Overall comment: You seem to have covered everything very well.

Most important soil management issues:

1. Soil compaction or 'soil slump' due to lack of organic matter and lack of soil structure on our light and medium silt soils. This can lead to soil water management problems where we can get ponding and water logging during heavy rainfall events. This is showing up (this year) where the crop has either died altogether where the pea roots have been sitting in water for more than 24 hours, or showing up as areas of pea foot rot where soils have remained wet and cold for periods of time.
2. Loss of soil biodiversity
3. Loss of nutrient availability

Respondent 3

Slide 1: Decline in Productivity

- Additions to:
    - Erosion: sheet, gully.
    - Soil water management: too dry
    - Acidity: pH too high, pH too low
    - Weeds: herbicide resistance (offered in place of black grass)
    - Loss of soil biodiversity: earthworms, mycorrhizal fungi
    - Nutrient management: nutrient lock-up, too high a nutrient e.g. Mg

Slide 2: Soil compaction

- Bioengineering addition: companion crops

Slide 3: Soil erosion by water

- Note for field characteristics: length and angle of slope

- Addition to management issues: row direction in relation to slope
- Addition to mechanical solutions: tied ridging etc.

Slide 4: Soil erosion by wind

- Additions to inherent factors affecting soil-borne disease: drought, flooding, carried on tractor tyres/machinery, crop stress from any cause.

Most important for business:

Sold kg's per Ha

Most important soil management issues:

1. Soil structure
2. Available nutrients
3. Available water

Management issues that work for you:

Soil compaction –cover and companion crops –the most exciting option for improving soils.

Erosion by water –long term we need to increase organic matter content. Short term: vegetative strips, cover and companion crops, row direction in relation to slope.

Erosion by wind –row direction in relation to predominant wind.

Soil-borne disease –avoid crop stress. Vigorous crops less likely to succumb.

<u>Respondent 4</u>

Slide 1: Decline in Productivity

- Factors missing: water holding capacity, loss of carbon, many more weeds –suggest common weeds as a catch all (https://cereals.ahdb.org.uk/media/433546/g61-managing-weeds-in-arable-rotations-a-guide.pdf) suggest similar for pests and diseases –much diversity. Add worm numbers to loss of soil biodiversity.

Most important soil management issues:

1. Compaction
2. Soil structure
3. Erosion by water and wind
4. Pest/weed/disease pressure
5. Soil biology
6. Nutrient availability

Management issues that do/don't work for you:

All could be appropriate dependent on farm context.

Respondent 5

Main challenge: getting on the land early enough in the spring and its ability to let us finish off cropping into the Autumn. The farm is run in partnership with an arable/potato grower. Land is rotated for Herb and Salad production on average every 2 years. For land rotation, soil type and access to irrigation water from the underground mains are the most important considerations.

Most important soil management issues:

1. Drainage –ditch management and timeliness of cultivations
2. Soil friability –ease of working it into a seedbed. Our cultivations are all 'powered' and so the fewer times we need to move the soil to create a seedbed the better.
3. Compaction –subsoiling is critical to reducing pans and wheelings. We also set out the beds using GPS to try and match the wheelings up with the previous crops.
4. Erosion –the nature of how we grow on raised beds and irrigation mean we do suffer from water run off, if we are unable to suitably cross slopes as opposed to going up and down them.
5. Organic matter –rotation with the arable crops also for straw to be incorporated. Our own trash is incorporated and we will sow cover crop of Mustard to aid organic matter
6. Soil exhaustion –coriander is a problem when we grow too tight in the rotation.

Management issues that do/don't work for you:

1. Cover crops –the land otherwise would be bare during October to April
2. Land drainage –essential for the land to be free draining and manmade drains to function allowing us to be able to cultivate weekly from March to September.
3. Tillage –min-till is not an option for us
4. CTF –we adopt this system to reduce compaction. Unfortunately, by working on a bed system we then add to the compaction through the use of narrow tyres.

Respondent 6

Overall comment: it is a very comprehensive analysis of soil issues. However, behaviour is driven by a practical need –to get into the field- therefore best practice is often sacrificed. It is hoped that this project does not lose this prospective.

Most important soil management issue:

Asparagus: Compaction that leads to water management issues that can also relate to organic matter loss and disease loading.

Cherries: soil water dynamics. When soils go from being extremely dry to extremely wet the roots suck up the water in such a way that fruit cracking occurs. It is very important in perennial crops to achieve an even, uninterrupted growth across the year/years.

Soil-borne disease: generally weak pathogens so can only access the plants when they are in a weakened state. Phytophora is an aggressive pathogen so is able to weaken plants making them susceptible to other diseases. Aggressive pathogens tend to be rarer. Fusarium is always present in varying strains.

Management issues that do/don't work for you:

Compaction –cover crops (more work needed on this), sub-soiling, layout, field choice, being careful with ridging (minimal soil movement –smaller ridges). Adding organic matter in rows would help but there is no perceivable cost benefit to this so few if any growers take it up (it is done in Canada/USA).

Erosion –choice of site, min soil disturbance, slope management and cover cropping (more work needed).

Soil-borne disease –as you can't get to the disease it is all about maintaining plant health to combat invasion, use of fungicide if available (only 1 available in asparagus) and ground drains.

<u>Respondent 7</u>

Overall comment: "You've been very comprehensive in your assessments of the factors affecting productivity and soil management. I would be minded to keep things simple."

Slide 2: Soil compaction

- Made more explicit that one contributing factor to compaction is timing of operation.

Slide 3: Soil erosion by water

- Top 5 water erosion factors: bare soil, fetch, reduced infiltration rate, slope and rainfall intensity.
- Change-rainfall volume and intensity separated.

Slide 5: Soil-borne diseases

- Addition of crop rotations as a management solution

Note- improving soil resilience will help manage compaction, and erosion.

Most important soil management issues:

1. Nutrient management (including pH)
2. Water management (including increasing plant available water capacity)
3. Soil structure (to allow roots in)

<u>Respondent 8</u>

Slide 1: Decline in Productivity

- Soil management mainly influences water and nutrients (considering building blocks of yield – water, nutrients and solar energy.
- Factor effects on productivity decline depends on the thresholds and the level of severity. For some factors the scale of the issues would need to be severe before you would see a significant impact (e.g. loss of biodiversity, soil erosion).
- Greenhouse gas emissions are very unlikely to affect productivity
- Re-wording suggestion (1): 'soil water management' to 'soil drainage' –an important influence on degree of prolonged waterlogging, timeliness of operations and, at the other end of the scale, droughtiness.
- Re-wording suggestion (2): 'Loss of/inadequate soil organic matter/carbon' to 'Soil organic matter decline'.
- Nutrient management – the key issue is nutrient supply, i.e. that all nutrients are supplied in adequate amounts, rather than any particular nutrient being at a high level in terms of soil nutrient reserves.
- The most important factors in terms of scale of impact on productivity:
    - Nutrient supply and acidity; and in no particular order
    - Pests, soil borne disease, weeds, seedbed quality, compaction and drainage – all can be significant, but magnitude of impact depends on severity/thresholds.

Slide 2: Soil compaction

- The key factor is the weight on the soil (machinery and livestock) when it's wet.
- The other factors affect resistance and resilience.
- Compared to the importance of timing of field operations (and weight of machinery etc.) and stocking (and stocking rate), some of these factors have much lower significance (i.e. extractable iron, extractable cations etc.)

- Compaction avoidance requires:
    - Land engineering
    - Timing of operations
    - Rotational management
    - Ag-engineering
    - Soil amendments
    - Cover crops and companion crops
- Compaction alleviation requires:
    - Correct identification of soil compaction issue – depth of compaction etc.
    - Mechanical solutions
    - Cover crops (where compaction is not severe)

Slides 3 and 4: Soil erosion by water and wind

- Key factor is soil cover
- For wind erosion, importance of fetch and shelter
- Key management factor is soil cover; by vegetation, crop residues or other material to protect the soil surface from direct raindrop impact etc.
- Adding organic matter can improve soil resilience to erosion and can act as a mulch. Inorganic fertiliser will help reduce soil erosion when it encourages early season vegetative growth.

Slide 5: Soil-borne disease

- Crop rotation is a key factor
- Minimising the potential for cross contamination through movement of soil (e.g. on machinery), crop residues or via other amendments (i.e. compost/soils/others materials or wastes added to land).

Appendix 4. Literature review

Academic literature

Detailed below is a list of peer-reviewed literature identified as relevant to horticultural crop best management practices (2013 - 2016) and their associated soil management issue.

	Relevant management issue											
Reference	**Acidity**	**Biodiversity (in soil)**	**Carbon**	**Compaction**	**Drainage**	**Erosion by water**	**Erosion by wind**	**Nutrient supply**	**Organic matter decline**	**Pests**	**Soil-borne disease**	**Weeds**
Alexander, P. D. and Nevison, I. M. (2015) 'The long-term effects of repeated application of the same organic material to soil in a horticultural context', Acta Horticulturae, 1076, pp. 143–150.	x				x			x	x			
Achmon, Y., Harrold, D. R., Claypool, J. T., Stapleton, J. J., VanderGheynst, J. S. and Simmons, C. W. (2016) 'Assessment of tomato and wine processing solid wastes as soil amendments for biosolarization', Waste Management. Elsevier Ltd, 48, pp. 156–164.										x		
Asad-Uz-Zaman, M., Bhuiyan, M. R., Khan, M. A. I., Alam Bhuiyan, M. K. and Latif, M. A. (2015) 'Integrated options for the management of black root rot of strawberry caused by Rhizoctonia solani Kuhn', Comptes Rendus Biologies. Elsevier Masson SAS, 338(2), pp. 112–120.										x	x	
Basta, N. T., Busalacchi, D. M., Hundal, L. S., Kumar, K., Dick, R. P., Lanno, R. P., Carlson, J., Cox, A. E. and Granato, T. C. (2016) 'Restoring Ecosystem Function in Degraded Urban Soil Using Biosolids, Biosolids Blend, and Compost', Journal of Environment Quality. ASA/CSSA/SSSA, 45(1), p. 74.								x				
Beniston, J. W., Lal, R. and Mercer, K. L. (2016) 'Assessing and Managing Soil Quality for Urban Agriculture in a Degraded Vacant Lot Soil', Land Degradation & Development. John Wiley and Sons Ltd, 27(4), pp. 996–1006.				x					x			

Reference	Acidity	Biodiversity (in soil)	Carbon	Compaction	Drainage	Erosion by water	Erosion by wind	Nutrient supply	Organic matter decline	Pests	Soil-borne disease	Weeds
	Relevant management issue											
Beslic, Z., Pantelic, M., Dabic, D., Todic, S., Natic, M. and Tesic, Z. (2015) 'Effect of vineyard floor management on water regime, growth response, yield and fruit quality in Cabernet Sauvignon', Scientia Horticulturae. Elsevier, 197, pp. 650–656.					x							
Biala, J. and Milgate, M. (2014) 'Grower expectations and experiences with the use of organic mulches and soil amendments in the horticultural industry in Queensland, Australia', Acta Horticulturae, 1018, pp. 473–480.								x	x			x
Campiglia, E., Radicetti, E. and Mancinelli, R. (2015) 'Cover crops and mulches influence weed management and weed flora composition in strip-tilled tomato (Solanum lycopersicum)', Weed Research. Edited by C. Kempenaar. Blackwell Publishing Ltd, 55(4), pp. 416–425.												x
Carron, M. P., Pierrat, M., Snoeck, D., Villenave, C., Ribeyre, F., Marichal, R. and Caliman, J. P. (2015) 'Temporal variability in soil quality after organic residue application in mature oil palm plantations', Soil Research. CSIRO, 53(2), pp. 205–215.	x	x										
Castro, E., Mañas, P. and De Las Heras, J. (2013) 'Effects of wastewater irrigation in soil properties and horticultural crop (LACTUCA SATIVA L.)', Journal of Plant Nutrition, 36(11), pp. 1659–1677.								x				
Congreves, K. A. and Van Eerd, L. L. (2015) 'Nitrogen cycling and management in intensive horticultural systems', Nutrient Cycling in Agroecosystems. Kluwer Academic Publishers, 102(3), pp. 299–318.			x									
De Corato, U., Viola, E., Arcieri, G., Valerio, V. and Zimbardi, F. (2016) 'Use of composted agro-energy co-products and agricultural residues against soil-borne pathogens in horticultural soil-less systems', Scientia Horticulturae. Elsevier, 210, pp. 166–179.											x	

Reference	Relevant management issue											
	Acidity	Biodiversity (in soil)	Carbon	Compaction	Drainage	Erosion by water	Erosion by wind	Nutrient supply	Organic matter decline	Pests	Soil-borne disease	Weeds
de Oliveira, S. P., de Lacerda, N. B., Blum, S. C., Escobar, M. E. O. and de Oliveira, T. S. (2015) 'Organic Carbon and Nitrogen Stocks in Soils of Northeastern Brazil Converted to Irrigated Agriculture', Land Degradation & Development. John Wiley and Sons Ltd, 26(1), pp. 9–21.			x									
Di Lorenzo, R., Pisciotta, A., Santamaria, P. and Scariot, V. (2013) 'From soil to soil-less in horticulture: Quality and typicity', Italian Journal of Agronomy, 8(4), pp. 255–260.											x	
Dorais, M. and Alsanius, B. (2015) 'Advances and trends in organic fruit and vegetable farming research', in Janick, J. (ed.) Horticultural Reviews. Hoboken, New Jersey: John Wiley & Sons, Inc., pp. 185–268.		x										x
Döring, J., Frisch, M., Tittmann, S., Stoll, M. and Kauer, R. (2015) 'Growth, Yield and Fruit Quality of Grapevines under Organic and Biodynamic Management.', PloS one. Public Library of Science, 10(10), p. e0138445.		x	x			x	x	x	x		x	
Forget, V., Brochier, V., Vidal-Beaudet, L. and Poitrenaud, M. (2014) 'Compressibility and hydraulic conductivity of reconstituted soil-compost mixtures for urban horticulture', Acta Horticulturae, 1018, pp. 593–600.				x								
Ganeshamurthy, A. N. (2013) 'Strategies for soil carbon sequestration through horticultural crops', in Climate-resilient horticulture: adaptation and mitigation strategies, pp. 2016–2017.			x									
Ge, T., Chen, X., Yuan, H., Li, B., Zhu, H., Peng, P., Li, K., Jones, D. L. and Wu, J. (2013) 'Microbial biomass, activity, and community structure in horticultural soils under conventional and organic management strategies', European Journal of Soil Biology. Elsevier Masson SAS, 58, pp. 122–128.		x	x									

Reference	Relevant management issue: Acidity	Biodiversity (in soil)	Carbon	Compaction	Drainage	Erosion by water	Erosion by wind	Nutrient supply	Organic matter decline	Pests	Soil-borne disease	Weeds
Gonzalez Perez, P., Ye, J., Wang, S., Wang, X. and Huang, D. (2014) 'Analysis of the occurrence and activity of diazotrophic communities in organic and conventional horticultural soils', Applied Soil Ecology. Elsevier B.V., 79, pp. 37–48.		x								x	x	
Grassi, F., Mastrorilli, M., Mininni, C., Parente, A., Santino, A., Scarcella, M. and Santamaria, P. (2014) 'Posidonia residues can be used as organic mulch and soil amendment for lettuce and tomato production', Agronomy for Sustainable Development. Springer-Verlag France, 35(2), pp. 679–689.												x
Gravel, V., Dorais, M., Dey, D. and Vandenberg, G. (2015) 'Fish effluents promote root growth and suppress fungal diseases in tomato transplants', Canadian Journal of Plant Science. Agricultural Institute of Canada, 95(2), pp. 427–436.											x	
Guzzon, R., Gugole, S., Zanzotti, R., Malacarne, M., Larcher, R., von Wallbrunn, C. and Mescalchin, E. (2016) 'Evaluation of the oenological suitability of grapes grown using biodynamic agriculture: the case of a bad vintage.', Journal of applied microbiology. Blackwell Publishing Ltd, 120(2), pp. 355–65.										x		
Herencia, J. F. (2014) 'Enzymatic activities under different cover crop management in a Mediterranean olive orchard', Biological Agriculture & Horticulture. Taylor and Francis Ltd., 31(1), pp. 45–52.			x									
Jay, C. N., Fitzgerald, J. D., Hipps, N. A. and Atkinson, C. J. (2015) 'Why short-term biochar application has no yield benefits: evidence from three field-grown crops', Soil Use and Management. Edited by M. Goss. Blackwell Publishing Ltd, 31(2), pp. 241–250.								x				
Keesstra, S., Pereira, P., Novara, A., Brevik, E. C., Azorin-Molina, C., Parras-Alcántara, L., Jordán, A. and Cerdà, A. (2016) 'Effects of soil management techniques on soil water erosion in apricot orchards.', The Science of the total environment. Elsevier, 551–552, pp. 357–66.						x			x			

Reference	Acidity	Biodiversity (in soil)	Carbon	Compaction	Drainage	Erosion by water	Erosion by wind	Nutrient supply	Organic matter decline	Pests	Soil-borne disease	Weeds
	Relevant management issue											
Li, H. C., Gao, X. D., Zhao, X. N., Wu, P. T., Li, L. S., Ling, Q. and Sun, W. H. (2016) 'Integrating a mini catchment with mulching for soil water management in a sloping jujube orchard on the semiarid Loess Plateau of China', Solid Earth. Copernicus GmbH, 7(1), pp. 167–175.					x							
Liu, L., Sun, C., Liu, S., Chai, R., Huang, W., Liu, X., Tang, C. and Zhang, Y. (2015) 'Bioorganic Fertilizer Enhances Soil Suppressive Capacity against Bacterial Wilt of Tomato', PLOS ONE. Edited by G. Zhang. Public Library of Science, 10(4), p. e0121304.	x	x									x	
Manjunath, M., Kanchan, A., Ranjan, K., Venkatachalam, S., Prasanna, R., Ramakrishnan, B., Hossain, F., Nain, L., Shivay, Y. S., Rai, A. B. and Singh, B. (2016) 'Beneficial cyanobacteria and eubacteria synergistically enhance bioavailability of soil nutrients and yield of okra', Heliyon. Elsevier Ltd, 2(2), p. e00066. doi: 10.1016/j.heliyon.2016.e00066.			x					x				
Miller, S. A. and Smucker, A. J. M. (2015) 'A new soil water retention technology for irrigated highly permeable soils', in Joint ASABE/IA Irrigation Symposium 2015: Emerging Technologies for Sustainable Irrigation. American Society of Agricultural and Biological Engineers, pp. 726–730.					x							
Pattison, A., Kukulies, T., Forsyth, L. and Geense, P. (2014) 'Can soil nematode community structure be used to indicate soil carbon dynamics in horticultural systems?', Acta Horticulturae, 1018, pp. 425–434.			x									
Phonglosa, A., Bhattacharyya, K., Ray, K., Mandal, J., Pari, A., Banerjee, H. and Chattopadhyay, A. (2015) 'Integrated nutrient management for okra in an inceptisol of eastern India and yield modeling through artificial neural network', Scientia Horticulturae. Elsevier, 187, pp. 1–9.								x				
Pradhan, A., Dash, A. K., Mohanty, S. S. and Das, S. (2015) 'Potential use of fly ash in floriculture: A case study on the photosynthetic pigments content and vegetative growth of	x											

Reference	Relevant management issue Acidity	Biodiversity (in soil)	Carbon	Compaction	Drainage	Erosion by water	Erosion by wind	Nutrient supply	Organic matter decline	Pests	Soil-borne disease	Weeds
Tagetes erecta (MARIGOLD)', Ecology, Environment and Conservation. Enviro Media, 21, p. AS369-AS376.												
Ruzzi, M. and Aroca, R. (2015) 'Plant growth-promoting rhizobacteria act as biostimulants in horticulture', Scientia Horticulturae. Elsevier, 196, pp. 124–134.								x				
Sahoo, D. C., Madhu, M., Muralidharan, P. and Sikka, A. K. (2015) 'Land management practices for resource conservation under vegetable cultivation in Nilgiris hills ecosystem', Journal of Environmental Biology. Triveni Enterprises, 36(4), pp. 1039–1044.					x			x				
Shen, Z., Wang, B., Lv, N., Sun, Y., Jiang, X., Li, R., Ruan, Y. and Shen, Q. (2015) 'Effect of the combination of bio-organic fertiliser with Bacillus amyloliquefaciens NJN-6 on the control of banana Fusarium wilt disease, crop production and banana rhizosphere culturable microflora', Biocontrol Science and Technology. Taylor and Francis Ltd., 25(6), pp. 716–731.		x									x	
Wang, H., Wang, C., Zhao, X. and Wang, F. (2015) 'Mulching increases water-use efficiency of peach production on the rainfed semiarid Loess Plateau of China', Agricultural Water Management. Elsevier, 154, pp. 20–28.					x							
Wang, J., Zhao, Y. and Ruan, Y. (2015) 'Effects of Bio-organic Fertilizers Produced by Four Bacillus amyloliquefaciens Strains on Banana Fusarium Wilt Disease', Compost Science & Utilization. Taylor and Francis Inc., 23(3), pp. 185–198.											x	
Wang, P., Wang, Y. and Wu, Q. S. (2016) 'Effects of soil tillage and planting grass on arbuscular mycorrhizal fungal propagules and soil properties in citrus orchards in southeast China', Soil and Tillage Research. Elsevier, 155, pp. 54–61.			x									

	Relevant management issue											
Reference	**Acidity**	**Biodiversity (in soil)**	**Carbon**	**Compaction**	**Drainage**	**Erosion by water**	**Erosion by wind**	**Nutrient supply**	**Organic matter decline**	**Pests**	**Soil-borne disease**	**Weeds**
Wang, X.-L., Ye, J., Perez, P. G., tang, D.-M. and Huang, D.-F. (2013) 'The impact of organic farming on the soluble organic nitrogen pool in horticultural soil under open field and greenhouse conditions: a case study', Soil Science and Plant Nutrition, 59(2), pp. 237–248.	x		x					x				
Yang, L., Zhao, F., Chang, Q., Li, T. and Li, F. (2015) 'Effects of vermicomposts on tomato yield and quality and soil fertility in greenhouse under different soil water regimes', Agricultural Water Management. Elsevier, 160, pp. 98–105.									x			
Zaidi, A., Ahmad, E., Khan, M. S., Saif, S. and Rizvi, A. (2015) 'Role of plant growth promoting rhizobacteria in sustainable production of vegetables: Current perspective', Scientia Horticulturae. Elsevier, 193, pp. 231–239.								x				

Appendix 5. Research reports

AHDB Horticultural research projects and publications from the last 10 years identified as relevant to SMIS and a source of supporting data.

AHDB Horticulture soil related projects	Management issue	Sector
CP 107b - Growing Resilient Efficient And Thriving Soils (GREAT) Soils	All issues	CS
PE 010 - Improvement of soil health by manipulation of microbial community characteristics	Decline in organic matter/nutrient supply/aciditiy	PE
BOF 076a - Understanding the physiological disorders in daffodil - BOF 076 project extension to study the three-year-down crop	Drainage	BOF
TF 179 - Pear: The effect of soil moisture on fruit storage quality	Drainage	TF
FV 299 - An investigation into the adoption of green manures in both organic and conventional rotations to aid nitrogen management and maintain soil structure	Nitrogen supply/pests	FV
FV 380 - Identifying critical soil P in vining pea crops	Nutrient supply	FV
FV 299a - Extension of FV 299 - Investigation into the adoption of green manures in both organic and conventional rotations to aid nitrogen management and soil structure	Nutrient supply	FV
FV 428 - Vining peas: The effect of soil phosphate levels on rhizobial population	Nutrient supply	FV
FV 299a - Extension of FV 299 - Investigation into the adoption of green manures in both organic and conventional rotations to aid nitrogen management and soil structure	Nutrient supply/Drainage	FV
Green manures – effects on soil nutrient management and soil physical and biological properties	Nutrient supply/organic matter decline	CS
FV 377a - Leeks: improving risk assessment for free-living nematodes	Pests	FV
FV 447 - Carrots & Parsnips - Developing a strategy to control Free Living Nematodes	Pests	FV
CP 115 - Enhancing the soil food web to control soil dwelling pests of field vegetables (Teagasc Walsh Fellowship)	Pests	FV
SF 122 - Using soil nematode threshold levels to reduce direct feeding damage on roots and interactions with verticillium wilt of strawberry and raspberry	Pests	SF
BOF 050a - Narcissus: overcoming the problem of ‘soil sickness’ with particular reference to the Isles of Scilly	Pests/soil-borne disease	BOF
Soil disinfestation options for cut flower growers	Soil-borne disease	BOF

AHDB Horticulture soil related projects	Management issue	Sector
BOF 050 - Narcissus: overcoming the problem of soil sickness with particular reference to the Isles of Scilly	Soil-borne disease	BOF
BOF 069 - Narcissus: suppression of Fusarium basal rot using composts amended with specific biocontrol agents	Soil-borne disease	BOF
BOF 045 - Bulbs and cut flowers: development of a combined dazomet and metam-sodium treatment as an alternative to methyl bromide for soil sterilisation.	Soil-borne disease	BOF
BOF 039 - Narcissus: examination of the links between soil nitrogen and basal rot	Soil-borne disease	BOF
CP 128 - Development and delivery of knowledge transfer activities on current best practice for Oomycete root rot detection and control	Soil-borne disease	CS
FV 349 - Brassicas: Further Development of "in field" tests for resting spores of clubroot and the development of clubroot control based on detection	Soil-borne disease	FV
FV 358 - Onion: Pot experiment to examine the suppression of Fusarium basal rot using compost colonised with Trichoderma viride	Soil-borne disease	FV
FV 405 - Carrots: Control of carrot cavity spot through the use of pre-crop green manures/biofumigation	Soil-borne disease	FV
CP 046 - Carrot cavity spot: the effects of non-carrot crops on levels of relevant Pythium spp in the soil (HDC Studentship)	Soil-borne disease	FV
FV 449 - Onions: Investigation into the control of White Rot in bulb and salad onion crops	Soil-borne disease	FV
FV 446 - Leeks: White tip control (Phytophthora porri)	Soil-borne disease	FV
FV 448 - Carrot: An early warning system for risk of cavity spot in crops	Soil-borne disease	FV
CP 157 - Aerial oomycetes: A review of management and control options available for the UK horticulture industry	Soil-borne disease	FV/CS
HNS 196a - Identification of factors which influence infection and control of the newly emerged Peronospora causing downy mildew on aquilegia	Soil-borne disease	HNS
PE 017 - Nutrient management for disease control in tomato	Soil-borne disease	PE
PC 213a - Protected cut flowers: evaluation of two steaming methods for disinfesting soil of Fusarium oxysporum and Sclerotinia sclerotiorum (extension to PC 213)	Soil-borne disease	PO
Biofumigant Crops as Replacements for Methyl Bromide Soil Sterilisation in Sustainable Strawberry Production	Soil-borne disease	SF
CP 006 - Integrated use of soil disinfection and microbial organic amendments for the control of soil borne diseases and weeds in sustainable crop production (HortLINK)	Soil-borne disease / weeds	CS
FV 361 - Reducing the impact of sclerotinia disease on arable rotations, vegetable crops and land use	Soil-borne disease	FV
FV 266 - Mechanical weed control for integrated and organic salad brassica production	Weeds	FV

Appendix 6. Text mining prototyping

The following scripts, written in Python and R, have been used to explore the potential for text mining and natural language processing of the literature gathered. An initial development phase involved seeking means to aid the targeting of the substantial corpus of academic and practitioner literature gathered. That text mining approach followed a workflow designed to highlight potential matching of literature against target themes.

Typical Workflow:

1. Literature gathered - ‘pdf’ files of all materials (corpus)
2. Run makefile (see ‘Panel 1’), run clean command to prepare environment
3. Run makefile, convert PDF files to ASCII text format (using ‘tika’)
4. Run makefile, generate word-clouds, ngrams and frequencies
5. Run makefile, generate PDF report output summarising each document

Panel 1: Text mining ‘makefile’

The material below is a ‘makefile’ used on a Linux computer, used to direct the workflow of the text mining.

```
TEMP = tmp
SOURCE = sourcedocs
TAGGED = taggeddocs

install:
	bin/install.R
	pip install nltk
	pip install textstat
	bin/install.py
clean:
	rm $(TEMP)/*
	rm $(TAGGED)/*
	rm index.*
	rm Rplots.pdf
	rm logfile.txt
tika:
	java -jar lib/tika-app-1.11.jar --text -i $(SOURCE) --outputDir $(TEMP)
nlp:
	ls $(TEMP)/*.txt > filelist.txt
	java    -cp    lib/stanford-corenlp-3.5.2.jar:lib/stanford-corenlp-3.5.2-
models.jar:lib/xom.jar:lib/joda-time.jar:lib/jollyday.jar:lib/ejml-0.23.jar     -
Xmx2g     edu.stanford.nlp.pipeline.StanfordCoreNLP     -replaceExtension     -
outputDirectory $(TAGGED) -filelist filelist.txt
	rm filelist.txt
ner:
	java -mx400m -cp lib/stanford-ner.jar edu.stanford.nlp.ie.crf.CRFClassifier
-loadClassifier  classifiers/english.all.3class.distsim.crf.ser.gz  -outputFormat
tabbedEntities -textFiles $(TEMP)/*.txt >> $(TAGGED).tsv
ner2:
	java  -mx400m  -cp  lib/stanford-ner.jar  bin/NER.java  [serializedClassifier
[$(TEMP)/*.txt]]
ner3:
	java            -mx200m            -cp            lib/stanford-ner.jar
edu.stanford.nlp.parser.lexparser.LexicalizedParser      -retainTmpSubcategories
-originalDependencies -outputFormat "penn,typedDependencies" -outputFormatOptions
"basicDependencies" englishPCFG.ser.gz $(TEMP)/*.txt
```

```
wordclouds:
	bin/wordcloud.R $(TEMP)/*.txt
freqs:
	bin/freq.R $(TEMP)/*.txt
correlations:
	bin/correlation.R $(TEMP)/*.txt
styles:
	bin/style.py $(TEMP)/*.txt
summaries:
	bin/summarise.py $(TEMP)/*.txt
ngrams:
	bin/ngrams.py $(TEMP)/*.txt
doc:
	bin/figures.py $(TEMP)/*.pdf > index.tex
	pdflatex index.tex
	rm index.tex
```

Panel 2: Example Text Mining script

The material below is activated by the ‘make correlations’ command above. This is one of the scripts, by example, used to extract key data from the literature corpus.

```
#!/usr/bin/env rscript
library(tm)
library(Rgraphviz)

MakeCorrelation <- function (filename) {
    # Log output to logfile
       log <- file("logfile.txt")
       sink(log, append=TRUE)
       sink(log, append=TRUE, type="message")

    print(filename)
    text <- readLines(filename)
    docs <- Corpus(VectorSource(text))

    ## Replace problematic characters.
    toSpace <- content_transformer(function (x , pattern ) gsub(pattern, " ", x))
    docs <- tm_map(docs, toSpace, "/")
    docs <- tm_map(docs, toSpace, "@")
    docs <- tm_map(docs, toSpace, "\\|")

    ## Clean up the text.
    docs <- tm_map(docs, content_transformer(tolower))
    docs <- tm_map(docs, removeNumbers)
    docs <- tm_map(docs, removePunctuation)

    ## Remove English stop words.
    docs <- tm_map(docs, removeWords, stopwords("english"))

    ## Remove any other arbitrary stop words.
    docs <- tm_map(docs, removeWords, c("department", "email"))

    docs <- tm_map(docs, stripWhitespace)

    ## Expand TLAs
    toString <- content_transformer(function(x, from, to) gsub(from, to, x))
    docs <- tm_map(docs, stemDocument, language = "english")
    tdm <- TermDocumentMatrix(docs)
    print(tdm)
    terms <- findFreqTerms(tdm, lowfreq = 5)
```

```
    if (length(terms) > 20) {
        terms <- sample(terms, 20)
        pdf(sprintf("%s.correlation.pdf", filename))
        plot(tdm, terms = terms, corThreshold = 0.3)
        dev.off()
    }
    # Restore output to console
    sink()
    sink(type="message")
}

args <- commandArgs(trailingOnly = TRUE)
for (arg in args) {
    MakeCorrelation(arg)
}
```

Panel 3: Provisional text analysis results

As Panel 1 exemplifies, a range of text analysis scripts were developed, one such being presented in Panel 2. Below is representative output relating to one example report considered, the 2004 ADAS final report, Tones *et al.*, to Defra for project ‘HH1747’**,** ‘Large narcissus fly spatial dynamics’.

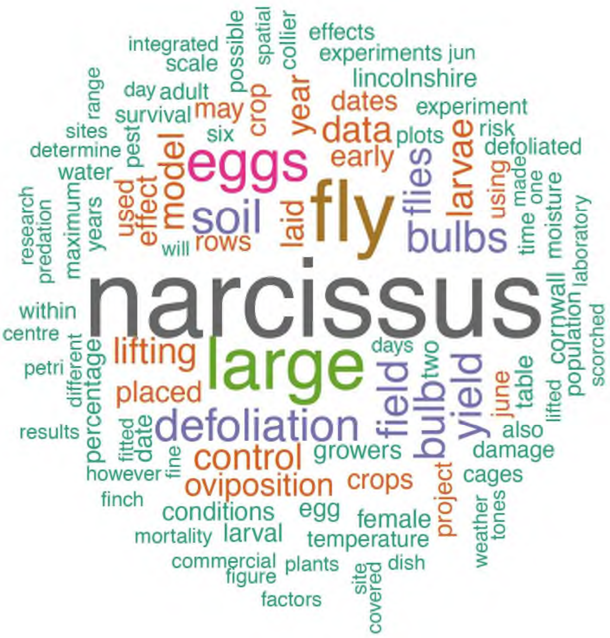

Figure a. Wordcloud extracted from report Tones *et al.* (2004)

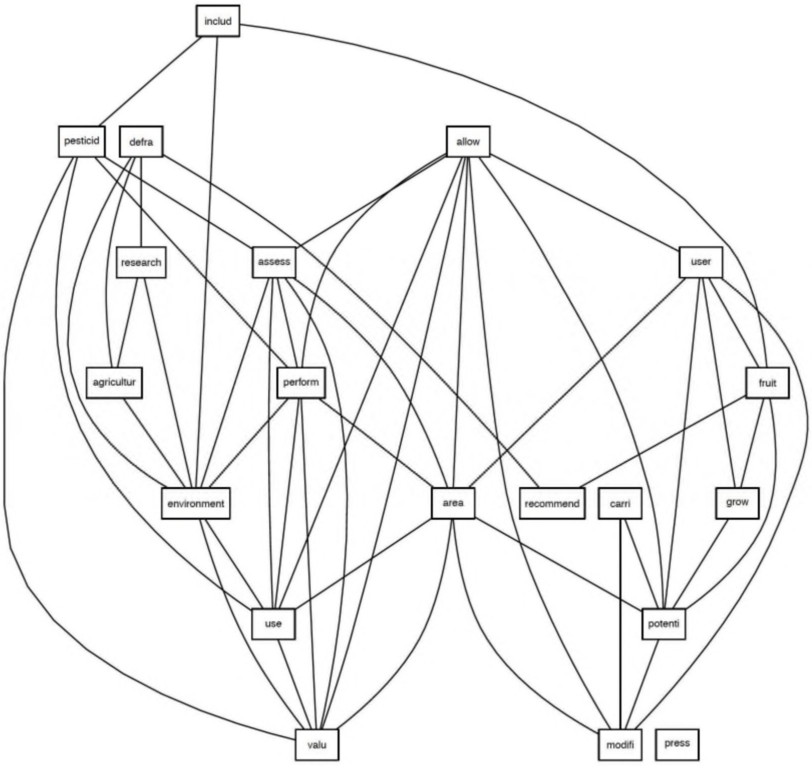

Figure b. Word correlations extracted from report Tones *et al.* (2004)

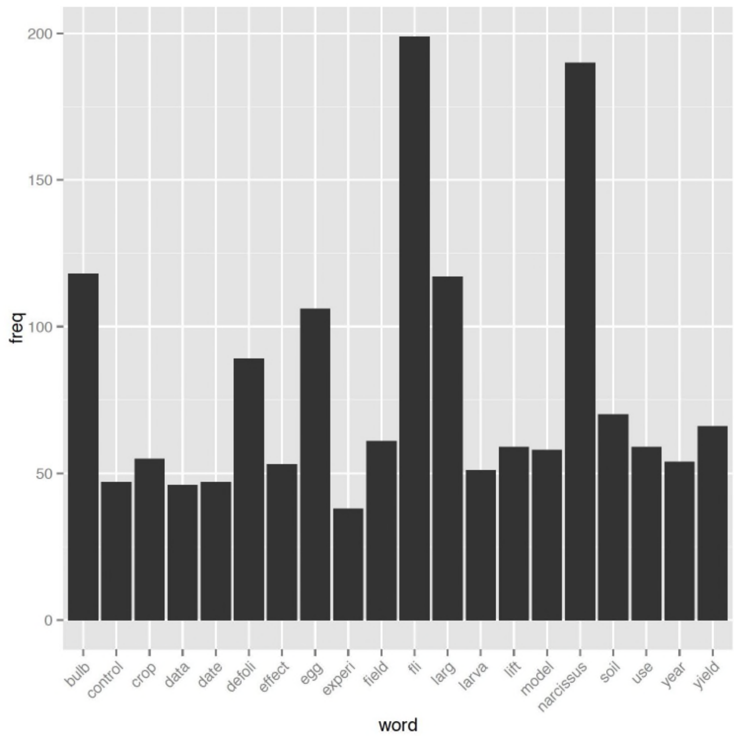

Figure c. Word frequencies extracted from report Tones *et al.* (2004)

Appendix 7. Initial database prototype and example queries

The following SQL code is used to create the database structure used in the initial database prototype. The database was built to specifically encode knowledge and relationships from academic papers, but the general concept is one that can be adapted to a variety of datasources.

```
CREATE TABLE Subjects (
      ID Counter,
      SubjectName Text(255),
      SubjectCategoryID Long
)
;

CREATE TABLE SubjectCategories (
      ID Counter,
      SubjectCategoryName Text(255)
)
;

CREATE TABLE PaperSubjects (
      ID Counter,
      PaperID Long,
      SubjectID Long
)
;

CREATE TABLE Papers (
      ID Counter,
      PaperTitle Text,
      PaperDate DateTime,
      Author Text,
      Location Text,
      Field Size Text
)
;

CREATE TABLE PaperOutputs (
      ID Counter,
      PaperID Long,
      OutputID Long,
      OutputValue Text(255)
)
;

CREATE TABLE PaperInputs (
```

```
        ID Counter,
        PaperID Long,
        InputID Long,
        InputValue Text(255)
)
;

CREATE TABLE Outputs (
        ID Counter,
        OutputName Text(255),
        Description Text(255),
        ParentCategory Long
)
;

CREATE TABLE OutputCategories (
        ID Counter,
        CategoryName Text(255)
)
;

CREATE TABLE Inputs (
        ID Counter,
        InputName Text(255),
        Description Text(255),
        ParentCategory Long
)
;

CREATE TABLE InputCategories (
        ID Counter,
        CategoryName Text(255)
)
;

CREATE INDEX SubjectCategoryID
ON Subjects (SubjectCategoryID ASC)
;

CREATE INDEX PaperID
ON PaperSubjects (PaperID ASC)
;

CREATE INDEX SubjectID
ON PaperSubjects (SubjectID ASC)
;

CREATE INDEX OutputID
ON PaperOutputs (OutputID ASC)
;
```

```
CREATE INDEX InputID
ON PaperInputs (InputID ASC)
;

CREATE INDEX PaperID
ON PaperInputs (PaperID ASC)
;

ALTER TABLE Subjects ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE SubjectCategories ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE PaperSubjects ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE Papers ADD CONSTRAINT PrimaryKey
PRIMARY KEY (ID)

;


ALTER TABLE PaperOutputs ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE PaperInputs ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE Outputs ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE OutputCategories ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE Inputs ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
;

ALTER TABLE InputCategories ADD CONSTRAINT PrimaryKey
      PRIMARY KEY (ID)
```

```
;


ALTER TABLE Subjects ADD CONSTRAINT SubjectCategoriesSubjects
      FOREIGN KEY (SubjectCategoryID) REFERENCES SubjectCategories (ID)
;

ALTER TABLE PaperSubjects ADD CONSTRAINT PapersPaperSubjects
      FOREIGN KEY (PaperID) REFERENCES Papers (ID)
;

ALTER TABLE PaperSubjects ADD CONSTRAINT SubjectsPaperSubjects
      FOREIGN KEY (SubjectID) REFERENCES Subjects (ID)
;

ALTER TABLE PaperOutputs ADD CONSTRAINT OutputsPaperOutputs
      FOREIGN KEY (OutputID) REFERENCES Outputs (ID)
;

ALTER TABLE PaperOutputs ADD CONSTRAINT PapersPaperOutputs
      FOREIGN KEY (PaperID) REFERENCES Papers (ID)
;

ALTER TABLE PaperInputs ADD CONSTRAINT InputsPaperInputs
      FOREIGN KEY (InputID) REFERENCES Inputs (ID)
;

ALTER TABLE PaperInputs ADD CONSTRAINT PapersPaperInputs
      FOREIGN KEY (PaperID) REFERENCES Papers (ID)
;

ALTER TABLE Outputs ADD CONSTRAINT OutputCategoriesOutputs
      FOREIGN KEY (ParentCategory) REFERENCES OutputCategories (ID)
;

ALTER TABLE Inputs ADD CONSTRAINT InputCategoriesInputs
      FOREIGN KEY (ParentCategory) REFERENCES InputCategories (ID)
;
```

The following PHP/HTML code is used to demonstrate the use of some example queries to retrieve items from the above database that meet certain metadata criteria.

examplequeries.php:

```
<!doctype html>
<html>
```



```
<head>
<meta charset="utf-8">
<title>Some example queries</title>
<link href="css/main.css" rel="stylesheet" type="text/css">
</head>

<body>
<h2>Some example queries</h2>
List all papers that tagged with <em>output</em> <strong>Yield
quality</strong>
<table border=1>
      <tr>
      <th>ID</th>
      <th>Title</th>
      <th>Author</th>
      <th>More info</th>
      </tr>
<?php

      $db = new SQLite3('phpLiteAdmin/ahdb.db');

      $stmt = $db->prepare("SELECT P.PaperTitle, P.ID, P.Author FROM
Papers P, PaperOutputs PO WHERE P.ID = PO.PaperID AND PO.OutputID = 2");
      $result = $stmt->execute();
      while ($row = $result->fetchArray()) {
      $count += 1;
      if (($count % 2) == 0)
      {
            $rowColour = "#D0FFD0";
      }
      else
      {
            $rowColour = "#FFFFD0";
      }
      echo "<tr bgcolor=\"" . $rowColour . "\">";
      echo "<td>" . $row["ID"] . "</td>\n";
      echo "<td>" . $row["PaperTitle"] . "</td>\n";
      echo "<td>" . $row["Author"] . "</td>\n";
      echo "<td><a href=\"paper.php?id=" . $row["ID"] .
"\">View</a></td>\n";
      echo "</tr>";
  }
?>
</table>

<hr />

List all <em>inputs</em> linked with <em>output</em> <strong>Yield
quality</strong>
<table border=1>
```



```
        <tr>
        <th>Input</th>
        <th>Frequency</th>
        </tr>
<?php
        $stmt2 = $db->prepare("SELECT I.InputName , COUNT(O.OutputName) AS
OutputCount
        FROM Outputs O, PaperOutputs PO, Papers P, PaperInputs PI, Inputs
I
        WHERE O.ID = PO.OutputID AND P.ID = PO.PaperID AND PI.PaperID =
P.ID AND I.ID = PI.InputID AND PO.OutputID = 2 GROUP BY I.InputName
        ORDER BY OutputCount DESC
");
        $result2 = $stmt2->execute();
        while ($row = $result2->fetchArray()) {
        $count += 1;
        if (($count % 2) == 0)
        {
                $rowColour = "#D0FFD0";
        }
        else
        {
                $rowColour = "#FFFFD0";
        }
        echo "<tr bgcolor=\"" . $rowColour . "\">";
        echo "<td>" . $row["InputName"] . "</td>\n";
        echo "<td>" . $row["OutputCount"] . "</td>\n";
        echo "</tr>";
    }
?>
</table>

<hr />

List all <em>outputs</em> linked with <em>input</em>
<strong>Rainfall</strong>
<table border=1>
        <tr>
        <th>Output</th>
        <th>Frequency</th>
        </tr>
<?php
        $stmt3 = $db->prepare("SELECT  O.OutputName, COUNT(I.InputName) AS
InputCount
FROM Outputs O, PaperOutputs PO, Papers P, PaperInputs PI, Inputs I
WHERE O.ID = PO.OutputID AND P.ID = PO.PaperID AND PI.PaperID = P.ID AND
I.ID = PI.InputID
AND PI.InputID = 1
GROUP BY O.OutputName
ORDER BY InputCount DESC
");
```

```
      $result3 = $stmt3->execute();
      while ($row = $result3->fetchArray()) {
      $count += 1;
      if (($count % 2) == 0)
      {
            $rowColour = "#D0FFD0";
      }
      else
        {
            $rowColour = "#FFFFD0";
      }
      echo "<tr bgcolor=\"" . $rowColour . "\">";
      echo "<td>" . $row["OutputName"] . "</td>\n";
      echo "<td>" . $row["InputCount"] . "</td>\n";
      echo "</tr>";
  }
?>
</table>

</body>
</html>
```

paper.php:

```
<!doctype html>
<html>
<head>
<meta charset="UTF-8">
<title>AHDB Draft Database</title>

<link href="css/main.css" rel="stylesheet" type="text/css">
</head>

<body>

  <a href="papers.php">&lt;&lt; Back to list of papers</a>
  <br />
  <strong>Paper details:</strong><br />
  <?php
  $urlID = htmlspecialchars($_GET["id"]);
  $db = new SQLite3('phpLiteAdmin/ahdb.db');

  echo "<table border=1>";
  $results = $db->query('SELECT * FROM Papers WHERE ID = '.$urlID);
  while ($row = $results->fetchArray()) {
      echo "<tr><td>ID</td><td>" . $row["ID"] . "</td></tr>\n";
      echo "<tr><td>Title</td><td>" . $row["PaperTitle"] .
"</td></tr>\n";
```

```
      echo "<tr><td>Author</td><td>" . $row["Author"] . "</td></tr>\n";
      echo "<tr><td>Date</td><td>" . $row["PaperDate"] . "</td></tr>\n";
  }
  echo "</table>";

  ?>
  <br/><strong>Tagged against the following subjects:</strong><br />
  <?php
  //Now query for the Subjects this paper is tagged under...

  $sqlQuery =   "SELECT ".
               "sc.SubjectCategoryName, ".
                "s.SubjectName ".
                "FROM Papers p ".
                "LEFT JOIN PaperSubjects ps ON p.ID = ps.PaperID ".
                "LEFT JOIN Subjects s ".
                "ON ps.SubjectID = s.ID ".
                 "LEFT JOIN SubjectCategories sc ".
                "ON s.SubjectCategoryID = sc.ID ".
                "WHERE p.ID = ". $urlID . " ".
                "ORDER BY sc.SubjectCategoryName ";

      echo "<table border=1>";
      echo "<tr><th>Category</th><th>Subject</th></tr>";
      $results = $db->query($sqlQuery);
      while ($row = $results->fetchArray()) {
      echo "<tr>";
      echo
"<td>".$row["SubjectCategoryName"]."</td><td>".$row["SubjectName"]."</td
>\n";
      echo "</tr>";
      }
      echo "</table>";
  ?>
  <br /><strong>Tagged with the following <em>inputs</em>:</strong><br
/>
  <?php
  //Now query for the Inputs this paper is tagged under...

  $sqlQuery =   "SELECT ".
                "ic.CategoryName, ".
                "i.InputName, ".
                "pi.InputValue as Notes ".
                "FROM Papers p ".
                "LEFT JOIN PaperInputs pi ".
                "ON p.ID = pi.PaperID ".
                "LEFT JOIN Inputs i ".
                "ON pi.InputID = i.ID ".
                "LEFT JOIN InputCategories ic ".
                "ON i.ParentCategory = ic.ID ".
                "WHERE p.ID = ". $urlID . " ".
```



```
                "ORDER BY ic.CategoryName";

      echo "<table border=1>";
      echo "<tr><th>Category</th><th>Input</th><th>Notes</th></tr>";
      $results = $db->query($sqlQuery);
      while ($row = $results->fetchArray()) {
      echo "<tr>";
      echo
"<td>".$row["CategoryName"]."</td><td>".$row["InputName"]."</td><td>".$r
ow["Notes"]."</td>\n";
      echo "</tr>";
      }
      echo "</table>";
  ?>

  <br /><strong>Tagged with the following <em>outputs</em>:</strong><br
/>
  <?php
  //Now query for the Inputs this paper is tagged under...

  $sqlQuery =   "SELECT ".
                "oc.CategoryName, ".
                "o.OutputName, ".
                "po.OutputValue as Notes ".
                "FROM Papers p ".
                "LEFT JOIN PaperOutputs po ".
                "ON p.ID = po.PaperID ".
                "LEFT JOIN Outputs o ".
                "ON po.OutputID = o.ID ".
                "LEFT JOIN OutputCategories oc ".
                "ON o.ParentCategory = oc.ID ".
                "WHERE p.ID = ". $urlID . " ".
                "ORDER BY oc.CategoryName";

      echo "<table border=1>";
      echo "<tr><th>Category</th><th>Output</th><th>Notes</th></tr>";
      $results = $db->query($sqlQuery);
      while ($row = $results->fetchArray()) {
      echo "<tr>";
      echo
"<td>".$row["CategoryName"]."</td><td>".$row["OutputName"]."</td><td>".$
row["Notes"]."</td>\n";
      echo "</tr>";
      }
      echo "</table>";
  ?>

</table>
<br />
<a href="papers.php">&lt;&lt; Back to list of papers</a>
```



```
</body>
</html>
```

newpaper.php:

```
<!doctype html>
<html>
<head>
<meta charset="utf-8">
<title>New Paper</title>
<link href="css/main.css" rel="stylesheet" type="text/css">
</head>

<body>
<form id="form1" name="form1" method="PUT" action="newpaperaction.php">

<table>
<tr>
<td><label for="papertitle">Paper title:</label></td>
<td><input type="text" name="papertitle" id="papertitle"></td>
<td></td>
</tr>

<tr>
<td><label for="author">Author:</label></td>
<td><input type="text" name="author" id="author"></td>
<td></td>
</tr>

<tr>
<td><label for="paperdate">Paper date:</label></td>
<td><input type="date" name="paperdate" id="paperdate"></td>
<td></td>
</tr>

<tr>
<td><label for="location">Location:</label></td>
<td><input type="text" name="location" id="location"></td>
<td></td>
</tr>


<tr>
<td><label for="subjects">Paper subjects:</label></td>
<td><select name="subjects">
<?php
  $db = new SQLite3('phpLiteAdmin/ahdb.db');
```



```
  $results = $db->query('SELECT * FROM Subjects');
  while ($row = $results->fetchArray()) {
      echo "<option value=\"" . $row["ID"] . "\">" . $row["SubjectName"]
. "</option>\n";
  }
?>
</select></td>
<td></td>
</tr>



<tr>
<td><label for="inputs">Paper inputs:</label></td>
<td><select name="inputs">
<?php
  $db = new SQLite3('phpLiteAdmin/ahdb.db');
  $results = $db->query('SELECT * FROM Inputs');
  while ($row = $results->fetchArray()) {
      echo "<option value=\"" . $row["ID"] . "\">" . $row["InputName"] .
"</option>\n";
  }
?>
</select></td>
<td><label for="inputnotes">Notes:</label> <input type="text"
name="inputnotes" id="inputnotes"></td>
</tr>



<tr>
<td><label for="outputs">Paper outputs:</label></td>
<td><select name="outputs">
<?php
  $db = new SQLite3('phpLiteAdmin/ahdb.db');
  $results = $db->query('SELECT * FROM Outputs');
  while ($row = $results->fetchArray()) {
      echo "<option value=\"" . $row["ID"] . "\">" . $row["OutputName"]
. "</option>\n";
  }
?>
</select>  </td>
<td><label for="outputnotes">Notes:</label> <input type="text"
name="outputnotes" id="outputnotes"></td>
</tr>

<tr>
<td colspan="3">
<input type="submit" name="submit" value="Submit">
</td>
```

```
</tr>

</table>
</form>

</body>
</html>
```

newpaperaction.php:

```
<!doctype html>
<html>
<head>
<meta charset="utf-8">
<title>Submitting new paper</title>
<link href="css/main.css" rel="stylesheet" type="text/css">
</head>

<body>

<?php

      $db = new SQLite3('phpLiteAdmin/ahdb.db');

      $stmt = $db->prepare("SELECT (MAX(ID)+1) AS MAXID FROM Papers");
      $result = $stmt->execute();
      $row = $result->fetchArray();
      $MAXID = $row["MAXID"];

      $stmt = $db->prepare("INSERT INTO Papers (ID,
PaperTitle,PaperDate,Author,Location) VALUES
(:id,:title,'now()',:author,:location)");
      $stmt->bindValue(':id', $MAXID);
      $stmt->bindValue(':title', $_GET["papertitle"]);
      $stmt->bindValue(':author', $_GET["author"]);
      $stmt->bindValue(':location', $_GET["location"]);
      //$result = $db->exec($sql);
      $result = $stmt->execute();
      if ($result) {
      echo ("INSERT 1 succesful<br />\n");
      }
      else {
      echo ("INSERT 1 FAILED!<br />\n");
      }
```

```
        $stmt2 = $db->prepare("INSERT INTO PaperInputs
(ID,PaperID,InputID,InputValue) VALUES ((SELECT (MAX(ID)+1) FROM
PaperInputs),:paperid,:intputid,:inputvalue)");
        $stmt2->bindValue(':paperid', $MAXID);
        $stmt2->bindValue(':intputid', $_GET["inputs"]);
        $stmt2->bindValue(':inputvalue', $_GET["inputnotes"]);
        $result2 = $stmt2->execute();
        if ($result2) {
        echo ("INSERT 2 succesful<br />\n");
        }
        else {
        echo ("INSERT 2 FAILED!<br />\n");
        }

        $stmt3 = $db->prepare("INSERT INTO PaperOutputs
(ID,PaperID,OutputID,OutputValue) VALUES ((SELECT (MAX(ID)+1) FROM
PaperOutputs),:paperid,:outputid,:outputvalue)");
        $stmt3->bindValue(':paperid', $MAXID);
     $stmt3->bindValue(':outputid', $_GET["outputs"]);
        $stmt3->bindValue(':outputvalue',  $_GET["outputnotes"]);
        $result3 = $stmt3->execute();
        if ($result3) {
        echo ("INSERT 3 succesful<br />\n");
        }
        else {
        echo ("INSERT 3 FAILED!<br />\n");
        }

        $stmt4 = $db->prepare("INSERT INTO PaperSubjects
(ID,PaperID,SubjectID) VALUES ((SELECT (MAX(ID)+1) FROM
PaperSubjects),:paperid,:subjectid)");
        $stmt4->bindValue(':paperid', $MAXID);
        $stmt4->bindValue(':subjectid', $_GET["subjects"]);
        $result4 = $stmt4->execute();
        if ($result4) {
        echo ("INSERT 4 succesful<br />\n");
        }
        else {
        echo ("INSERT 4 FAILED!<br />\n");
        }


?>


</body>
</html>
```

Appendix 8. Stakeholder Enquiry

The following request was prepared for use with our stakeholder partners, explanation followed by an enquiry form.

Providing your soil management data to SMIS

Overview

Thanks you for assisting the Cranfield University team with the AHDB-funded research ‘Development of a Horticultural Soil Management Information System’ (or SMIS). This document provides a summary overview as to how we can take in datasets from your project relating to soil management, providing guidance as to how stakeholders and interested parties to the SMIS project, who are data holders, can align their own data with the SMIS data repository.

What are we trying to do?

We are seeking to bring together a range of datasets and information, and descriptive ‘metadata’ for each, relating to the specific effects of soil management practices on horticultural crop productivity and environmental protection. Ultimately, we will use all of this together to produce an ‘*e-Guide toolkit*’, able to provide AHDB users with a set of robust, empirically-based, best-practice guidelines, and the likely consequences of applying them.

What do we need?

We will be asking for two things from you. Firstly, we would like to receive the **datasets** or key information itself, hopefully in computer format. This may mean sending us a DVD/CD, or using our electronic ‘dropoff’ tool to send us large files (see https://dropoff.cranfield.ac.uk/ for this useful tool). Equally, data may be accessible already via the web or a ‘web service’ which you can tell us about. Secondly, we would like to ask for any **related documentation** concerning this data, and in particular for you to help create a **descriptive statement** for the dataset, helping us to understand its content.

All of the data items we take in need to have this common description, and it follows six simple categories. This will help us be able to draw together and compare data from all the different sources we are using. There are six ‘criteria’ we are using for this description, thus:

Place
Time
Characteristics

Landuse
Operations
Outcomes

May we therefore ask you to help us by characterising your data in these terms. Where you are sharing a datasets with us, it will help to have an idea of which tables and columns relate to each of these. An appendix contains a form you can complete.

Where will the data reside?

The AHDB e-Guide toolkit we are developing needs to be able to draw from a range of data and information sources. Data can be supplied to the SMIS team in any electronic format. We would ask you provide us some relevant instructions as to how the data are arranged. If the data are made available online, e.g. as a web service, then that is also helpful. In the latter case, we would ask for details as to how to access the data in that electronic form (with passwords if that is required).

Will SMIS reveal personal information?

It is our intention that the e- Guide toolkit will incorporate the broad spectrum of data from stakeholder sources. However, in doing this, some personal and/or sensitive data may be provided. It is **NOT** the intention of SMIS to make such information available and any confidential data characteristics will be anonymised. We ask you help us by indicating clearly any such concerns in data being made available so that we can ensure its security and protection. If there are any concerns, please contact us to discuss.

Any further questions?

For any questions relating to the project and the overall objectives, please contact:

(Jane) R.J. Rickson, MSc, PhD, FIAgrE, CEnv, FHEA, MCIWEM, MI Soil Sci.
Professor of Soil Erosion and Conservation
Cranfield Soil and AgriFood Institute (Building 52A)
School of Water, Energy and Environment
Cranfield University
Cranfield
Bedfordshire MK43 0AL
United Kingdom
Telephone: +44 (0) 1234 750111 (extension 2705)
Fax: +44 (0) 1234 752970
email: j.rickson@cranfield.ac.uk
https://www.cranfield.ac.uk/people/professor-jane-rickson-770215

For any technical queries, please contact

Dr Stephen Hallett, SFHEA
Principal Research Fellow in Environmental Informatics
School of Water, Energy and Environment
Building B52a First Floor, Front office. Cranfield University
Cranfield, Bedfordshire
MK43 0AL, UK
Telephone +44 (0) 1234 750111 Switchboard (extension 2750)
Mob +44 (0) 786 7500697
email: s.hallett@cranfield.ac.uk
https://www.cranfield.ac.uk/people/dr-stephen-hallett-786115

Data Source Form

Please can you complete the following three sections relating to the data source you are sharing with us:

About you	***The person responsible for this data source***
Name	
Affiliation	
Address	
Telehone	
Email	

About the data	
Name of data source	
Purpose of data	
Place	
Time	
Characteristics	
Landuse	
Operations	
Outcomes	

Intellectual Property usage issues statement	
Personal information issues statement	

Accessing the data	
How will the data be made available	
What format is the data held in electronically	
Any technical notes to aid us accessing the data	

Appendix 9. Data, Information and Knowledge

This report, and the SMIS project as a whole, makes frequent reference to the concepts of 'data', 'information' and 'knowledge'. For the purposes of this project, we follow the definitions of Anderson, R. (1991) Information and Systems. Journal of Applied Systems Analysis, 18. 57-60., thus:

Data

That body of facts or figures, which have been gathered systematically for one or more specific purposes

Information

Those data which have been processed into a form that is meaningful to a recipient and is of perceived value in current or prospective decision making

Knowledge

Information that is transformed to encapsulate understanding that can be recorded

>> **For example** one may collect and hold raw meteorological station record keeping data. In turn, this data, acquire a significance, having relevancy and a purpose when used to generate pressure isobars, or temperature isotherm information, and from this information, with the capacity to transform information into a valuable outcome or result, thus knowledge may be derived as the weather forecast for tomorrow.

We refer to 'databases', to 'information systems' and to 'knowledge management systems'. SMIS is an information system holding various forms of purposeful data. The ultimate e-Guide represents a knowledge management system. Metadata is a specialised form of data, used to produce a description of another body of data, information or knowledge. In this project, we use metadata to refer to the simple recording of the six criteria used to describe sources used.

Appendix 10. Towards a Common Data Model

Quantitative data on field and land operations destined for the SMIS arises from a number of sources, as outlined in Section 2 of this report. Where data is available, at the current time, and in digital form, the specific formats have been found to vary between sources. Likewise, where records exist only in paper format, there is a need to propose a coherent and consistent digital data model to act as a template. Furthermore, looking to the future, where new data is required to be entered into the SMIS, then, again, a digital template will be required. This appendix presents such a proposed template, both in schematic, and in computer-compatible Structured Query Language (SQL) form.

The purpose of this information is to permit the presentation of a set of common field-specific grower data attributes that serve the following purposes:

1. They allow a harmonisation of critical data between respective data sources associated with different growers – whose data may be in a formal package (e.g. Gatekeeper^{tm} – in all its varieties, Muddy boots^{tm} etc; MS Excel^{tm} spreadsheets; handwritten records etc).
2. They allow the basis for an on-screen data capture form in the web application for grower clients to use who do not currently have additional software tools – thus a data entry form.
3. They provide a basis for harmonised data analysis between farms and fields that would otherwise be impossible given heterogeneity in the source data formats.

Common Data Model - Schematic form

The data tables below can never be fully inclusive of all the permutations and combinations of data from the various sources. However, they do represent a coherent and useful set of data attributes for describing agronomic practices on the farm at the field scale. It is important to note that the approach taken for this harmonised data designs in data anonymity for growers for any overarching analysis – whilst retaining name and location for analysis and cross-referencing purposes. Themes for the data to be recorded are as follows:

Table 1. Data Themes

Theme	Content
Grower	Information about the user/client of Soil for Life
Field	Information about the field site, the location, the field name, site characteristics such as area, soils etc.
Field Crop Results	Information about the crops grown, year on year, and the yield characteristics, in a given field site

Field Operations	Information about any interventions at the field site level, year by year – operations and additives to the field: fungicides; growth regulators; herbicides; molluscicides; Organic manure; Lime; Insecticides; Seeds and planting; Trace elements; Adjuvants; Desiccants, soil management etc.
Field Pests Diseases	Information about any specific pests and diseases recorded at the field site, year by year – PCN etc.

Further to this, a series of 'look up tables' (LUT) will be required.

Data table: Grower

Field name	Data Type	Example
Grower_Id (PK)	Integer	1
Grower_Name	Text	John Bull
Grower_Address_1	Text	Old Manor Farm
Grower_Address_2	Text	West Riding
Grower_Address_3	Text	Shopshire
Grower_Town	Text	Gloucester
Grower_Postcode	Text	PL12 5JU
Grower_Telephone	Text	01234 567890
Grower_Mobile	Text	07891 23456789
Grower_Email	Text	j.bull@manorfarm.co.uk
Grower_Preference_Newsletter	Boolean	Y

Data table: Field

Field name	Data Type	Example
Field_Id (PK)	Integer	1
Field_FieldName	Text	001 Bridge
Site_FieldCode	Text	TF33316053
Field_AreaHaOfficial	Float	18.5

Field_AreaHaWorking	Float	12
Field_EastingOSGB	Integer	*533600 (calculated)*
Field_NorthingOSGB	Integer	*331530 (calculated)*
Field_SoilTextureDominant	Text	Light sand
Field_SoilSulphurDeficiencyDominant	Boolean	No
Field_KReleasingClay	Boolean	No
Field_DrainageDominant	Text	Well-drained
Field_SlopeDominant	Text	Moderate
Field_AspectDominant	Text	South
Field_InNVZ	Boolean	Yes

Data table: FieldCropGrowerResults

Pivot table joining Field, Grower and Crop, with outcomes

Field name	**Data Type**	**Example**
FieldCropResults_Id (PK)	Integer	1
Field_Id (FK)	Integer	1
Grower_Id (FK)	Integer	1
Crop_Id (FK)	Integer	1
FieldCropResults_Year	Date	2009
FieldCropResults_AreaHa	Float	35
FieldCropResults_SowingDate	Date	2009-03-15
FieldCropResults_HarvestDate	Date	2009-11-04
FieldCropResults_Yield	Float	7.89
FieldCropResults_Residue	Text	
FieldCropResults_Undersown	Boolean	N

Data table: FieldOperations

Field name	Data Type	Example
FieldOperations_Id (PK)	Integer	1
Product_Id (FK)	Integer	2
FieldOperations_Date	Date	2009-10-05
FieldOperations_Quantity	Float	183.890
FieldOperations_Units	Text	L

Data table: FieldPestsDiseases

Field name	Data Type	Example
FieldPestsDiseases_Id (PK)	Integer	1
PestDiseases_Id (FK)	Text	2
FieldPestsDiseases_Date	Date	2009-10-05
FieldPestsDiseases_Quantity	Float	183.890
FieldPestsDiseases_Units	Text	L

Look up Tables

Look up tables will be required to hold the following:

Data table: Crop

Field name	Data Type	Example
Crop_Id (PK)	Integer	1
Crop_Name	Text	Spring Barley
Crop_Group	Text	Cereals
Crop_Variety	Text	RGT Planet
Crop_Destination	Text	Malting

Data table: Analysis

Field name	Data Type	Example
Analysis_SiteId (PK - compound)	Integer	1
Analysis_Date (PK - compound)	Text	2009-10-05
Analysis_K	Text	114
Analysis_Mg		65
Analysis_P		26
Analysis_pH	Integer	4
Analysis_SOM	Float	34

Data table: Product

Field name	Data Type	Example
Product_Id (PK)	Integer	1
Product_Heading	Text	Molluscicide
Product_CommercialName	Text	Adigor
Product_ManufacturersApplicationRate	Float	66
Product_ManufacturersApplicationRateUnits	Text	Mg/l
Product_ManufacturersApplicationDate	Date	October

Data table: PestsDisease

Field name	Data Type	Example
PestsDisease_Id (PK)	Integer	1
PestsDisease_Type	Text	Molluscs
PestsDisease_Heading	Text	Slug

Common Data Model - SQL form

Once a schematic design has been decided upon, The next stage is to develop a digital representation in physical form of the logical design itself.This section below presents in structured query language (SQL). A design for the structure as outlined.This design must be considered a working document at the current time, but is useful as an indication of the design approach.

```
USE smis
;


IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[Grower]') AND  OBJECTPROPERTY(id, 'IsUserTable') =
1)
DROP TABLE [Clients_Unified].[Grower]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[Field]') AND  OBJECTPROPERTY(id, 'IsUserTable') =
1)
DROP TABLE [Clients_Unified].[Field]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[FieldCropGrowerResults]') AND  OBJECTPROPERTY(id,
'IsUserTable') = 1)
DROP TABLE [Clients_Unified].[FieldCropGrowerResults]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[FieldOperations]') AND  OBJECTPROPERTY(id,
'IsUserTable') = 1)
DROP TABLE [Clients_Unified].[FieldOperations]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[FieldPestsDiseases]') AND  OBJECTPROPERTY(id,
'IsUserTable') = 1)
DROP TABLE [Clients_Unified].[FieldPestsDiseases]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[Crop]') AND  OBJECTPROPERTY(id, 'IsUserTable') = 1)
DROP TABLE [Clients_Unified].[Crop]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[Analysis]') AND  OBJECTPROPERTY(id, 'IsUserTable')
= 1)
DROP TABLE [Clients_Unified].[Analysis]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[Product]') AND  OBJECTPROPERTY(id, 'IsUserTable') =
1)
DROP TABLE [Clients_Unified].[Product]
;

IF EXISTS (SELECT * FROM dbo.sysobjects WHERE id =
object_id('[Clients_Unified].[PestsDiseases]') AND  OBJECTPROPERTY(id,
```

```
'IsUserTable') = 1)
DROP TABLE [Clients_Unified].[PestsDiseases]
;


CREATE TABLE [Clients_Unified].[Grower] (
	[Grower_Id] varchar(50) NOT NULL,
	[Grower_Name] varchar(50),
	[Grower_Address_1] varchar(50),
	[Grower_Address_2] varchar(50),
	[Grower_Address_3] varchar(50),
	[Grower_Town] varchar(50),
	[Grower_Postcode] varchar(50),
	[Grower_Telephone] varchar(50),
	[Grower_Mobile] varchar(50),
	[Grower_Email] varchar(50),
	[Grower_Preference_Newsletter] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[Field] (
	[Field_Id] varchar(50) NOT NULL,
	[Field_FieldName] varchar(50) NOT NULL,
	[Site_FieldCode] varchar(50),
	[Field_AreaHaOfficial] varchar(50),
	[Field_AreaHaWorking] varchar(50),
	[Field_Easting_OSGB] varchar(50),
	[Field_Northing_OSGB] varchar(50),
	[Field_SoilTextureDominant] varchar(50),
	[Field_SoilSulphurDeficiencyDominant] varchar(50),
	[Field_KReleasingClay] varchar(50),
	[Field_DrainageDominant] varchar(50),
	[Field_SlopeDominant] varchar(50),
	[Field_AspectDominant] varchar(50),
	[Field_InNVZ] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[FieldCropGrowerResults] (
	[FieldCropGrowerResults_Id] varchar(50) NOT NULL,
	[Field_Id] varchar(50),
	[Grower_Id] varchar(50),
	[Crop_Id] varchar(50),
	[FieldCropResults_Year] varchar(50),
	[FieldCropResults_AreaHa] varchar(50),
	[FieldCropResults_SowingDate] varchar(50),
	[FieldCropResults_HarvestDate] varchar(50),
	[FieldCropResults_Yield] varchar(50),
	[FieldCropResults_Residue] varchar(50),
	[FieldCropResults_Undersown] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[FieldOperations] (
	[FieldOperations_Id] varchar(50) NOT NULL,
	[Product_Id] varchar(50),
	[FieldOperations_Date] varchar(50),
	[FieldOperations_Quantity] varchar(50),
	[FieldOperations_Units] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[FieldPestsDiseases] (
	[FieldPestsDiseases_Id] varchar(50) NOT NULL,
```



```
        [PestsDiseases_Id] varchar(50),
        [FieldPestsDiseases_Date] varchar(50),
        [FieldPestsDiseases_Quantity] varchar(50),
        [FieldPestsDiseases_Units] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[Crop] (
        [Crop_Id] varchar(50) NOT NULL,
        [Crop_Name] varchar(50),
        [Crop_Group] varchar(50),
        [Crop_Variety] varchar(50),
        [Crop_Destination] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[Analysis] (
        [Analysis_FieldId] int NOT NULL,
        [Analysis_Date] varchar(50) NOT NULL,
        [Analysis_K] varchar(50),
        [Analysis_Mg] varchar(50),
        [Analysis_P] varchar(50),
        [Analysis_pH] varchar(50),
        [Analysis_SOM] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[Product] (
        [Product_Id] varchar(50) NOT NULL,
        [Product_Heading] varchar(50),
        [Product_CommercialName] varchar(50),
        [Product_ManufacturersApplicationRate] varchar(50),
        [Product_ManufacturersApplicationRateUnits] varchar(50),
        [Product_ManufacturersApplicationDate] varchar(50)
)
;

CREATE TABLE [Clients_Unified].[PestsDiseases] (
        [PestsDiseases_Id] varchar(50) NOT NULL,
        [PestsDiseases_Type] varchar(50),
        [PestsDiseases_Heading] varchar(50)
)
;


ALTER TABLE [Clients_Unified].[Grower] ADD CONSTRAINT [PK_Grower]
        PRIMARY KEY CLUSTERED ([Grower_Id])
;

ALTER TABLE [Clients_Unified].[Field] ADD CONSTRAINT [PK_Field]
        PRIMARY KEY CLUSTERED ([Field_Id])
;

ALTER TABLE [Clients_Unified].[FieldCropGrowerResults] ADD CONSTRAINT
[PK_FieldCropGrowerResults]
        PRIMARY KEY CLUSTERED ([FieldCropGrowerResults_Id])
;

ALTER TABLE [Clients_Unified].[FieldOperations] ADD CONSTRAINT
[PK_FieldOperations]
        PRIMARY KEY CLUSTERED ([FieldOperations_Id])
;
```



```
ALTER TABLE [Clients_Unified].[FieldPestsDiseases] ADD CONSTRAINT
[PK_FieldPestsDiseases]
	PRIMARY KEY CLUSTERED ([FieldPestsDiseases_Id])
;

ALTER TABLE [Clients_Unified].[Crop] ADD CONSTRAINT [PK_Crop]
	PRIMARY KEY CLUSTERED ([Crop_Id])
;

ALTER TABLE [Clients_Unified].[Analysis] ADD CONSTRAINT [PK_Analysis]
	PRIMARY KEY CLUSTERED ([Analysis_FieldId], [Analysis_Date])
;

ALTER TABLE [Clients_Unified].[Product] ADD CONSTRAINT [PK_Product]
	PRIMARY KEY CLUSTERED ([Product_Id])
;

ALTER TABLE [Clients_Unified].[PestsDiseases] ADD CONSTRAINT [PK_PestsDiseases]
	PRIMARY KEY CLUSTERED ([PestsDiseases_Id])
;
```

