PROJECT NUMBER 289706

Start date of the project: 01/12/2011, duration: 48 months

Assessing and Monitoring The Impacts of Genetically Modified Plants on Agro-ecosystems

Project acronym: AMIGA **Funding Scheme:** Large Cooperative Project

Report: Indicators of Agro-Ecosystem Function

Deliverable 3.3 Report on a definitive set of indicators

Author: Dr**.** Geoffrey SQUIRE - geoff.squire@hutton.ac.uk

Organisation names of lead contractors for this deliverable: James Hutton Institute

Dundee DD2 5DA, UK

February 2014 **Dissemination Level: Public**

Project leader: Dr. Salvatore ARPAIA
ENEA - Centro Ricerche Trisaia
Tel:+39-0835-974714
E-mail: Salvatore.arpaia@enea.it
Project website address: www.amigaproject.eu

SUMMARY

1. Indicators are widely used in genetically modified (GM) plants risk-assessment to define experimental sites, compare plants and crops and assess the effects of GM cultivation on the environment. Most approaches to date begin by characterising the GM plant then measuring its influence on other organisms and ecological processes.

2. In contrast, the rationale is presented here for a system-centred approach to assessing long term effects of new cropping systems. The main required outputs of the agricultural ecosystem are first defined in terms of ecosystem services. The ecological processes that deliver the outputs are identified, then the life forms (crops, weeds, invertebrates) that mediate the processes and finally the interventions in the form of new crops and field management (Chapter 1). The purpose of this study is to define a comprehensive set of biophysical and economic indicators needed for this system-centred approach.

3. The indicators (or measurables) used so far in major field studies of GM cropping in Europe are summarised, with particular attention given to ecological impacts of GMHT and Bt cropping and to geneflow and persistence (Chapter 2). The groups of indicators used in these studies are comprehensive and detailed in themselves but are mostly restricted to organisms and processes that are affected directly by the GM crop and its management.

4. In order to devise a more comprehensive set of indicators, approaches are examined in two major (non-GM) field studies that aimed to establish a system-centred approach to sustainable cropping (Chapter 3). Indicators defining the supporting functions of soil, the stores and fluxes of energy, carbon, nitrogen and other major plant nutrients, and the economics of cropping are considered to be essential for a system-centred approach. Several issues of scale were identified: the need to anticipate effects of increasing complexity as a new crop is commercialised; the scales at which ecosystem services can be both influenced and achieved; and the means to reference across scales, for example when examining the representativeness of field experiments to receiving environments.

5. Nine groups of indicators that would form a comprehensive set, fit for purpose, are then summarised under the headings: 1) crop and management, 2) field structure, 3) energy and matter cycling, 4) soil biophysical status, 5) soil microbial and faunal status, 6) wild plants and food webs, 7) pests and integrated pest management, 8) economics and 9) regional and national census data on inputs and outputs. Several of these (e.g. 5, 6, 7, 8) are being refined and tested in other AMIGA workpackages and will be updated as the project evolves.

6. This comprehensive set of indicators will be used in a system-centred approach to define existing long term trends in production systems, to assess regions, field sites and alternative interventions and finally to compare the likely long term impacts of GM cropping with other recent and current human-induced change in agriculture.

Acknowledgements

All partners in **AMIGA** are here acknowledged for their constructive discussions on indicators of ecosystem processes, particularly:

Astrid Näther and Christophe Tebbe, VTI Germany (WP4), who provided the material for Table 4.4 and associated text;

Julian Park and Alice Mauchline, University of Reading, UK (WP10) who suggested the economic indicators in Table 4.8 and provided text on economic indicators and the Agri-Environmental Footprint Index;

Salvatore Arpaia (project coordinator) for concepts on invertebrate functional groups, food webs and focal species;

Cathy Hawes, JHI UK, for constructing an original representation of scaling in agroecosystems on which the simplified diagram in Fig. 3.1 was based.

FP6 SIGMEA project

Partners in the SIGMEA project http://www6.inra.fr/sigmea which reported in 2007, developed and tested much of the methodology in geneflow, persistence and coexistence in European conditions (Table 2.2). Some initial ideas on a system-centred rather than GM-centred approach arose from discussion in SIGMEA WP2 with Marco Debeljak at the Josef Stefan Institute, Slovenia.

CONTENTS

1. INTRODUCTION, CONCEPT, SCOPE

The state of an ecological or economic system has to be quantified if it is to be understood and managed. Agroecological systems are highly complex, comprising stores and fluxes of energy and matter and living things that interact with mineral substances to produce soil, food webs and useful offtake. Not everything in such a complex system can be quantified – some of the inner workings might be unclear and some of the methods of characterisation might simply need more time and effort than can be justified or made available.

Therefore, indicators are needed that will enable researchers to observe a system and characterise its evolution and dynamics. Indicators, as used here, are measurable or modelled entities with values attached to them (e.g. primary production, t ha^{-1}; population density, m^{-2}; nitrogen uptake g m^{-2}).

1.1 Aims and purpose of indicators within AMIGA

In AMIGA WP3 on *Long term effects*, indicators are needed to:

- form the quantitative evidence-base for characterising arable production systems;
- allow comparison of managed ecosystems across different geographical regions or between different farming approaches;
- assess the impact of change on these systems, such as the introduction of new technologies and crop management practices.

To serve such purposes, indicators should be widely understood, be generic, i.e. not specific to particular regions or crops, and be measurable with proportionate amounts of time, effort and skill. They should be able to point to sensitivities in a system and also its responses to external and internal change.

This task on indicators (AMIGA WP3.5) is part of a sequence of work that leads towards answering the major question of the final task in WP3 which is 'Assessment of the degree of long term change due to introduction of GM cropping'. Since it impossible to be able to define long term effects in a cropping system from measurements in a few consecutive years, attention is directed to several questions that reverse the typical direction of query. So, rather than ask about the effect in the long-term of a GM crop and its management, the study will address the following questions.

- What sorts of impact would an innovation have to exert in a region or cropping system a change similar to major trends, for example those that have occurred during intensification in the past 30 years?
- What sorts of impact would push the agroecosystem in a negative (or positive) direction and more specifically towards or beyond limits of concern (safe limits for ecological processes)?
- How might the potential long term effects of GM cropping compare, in terms of size and direction, with those identified in questions (1) and (2)?

The centre of the argument is not therefore on GMOs but on the systems into which they would be introduced. By concentrating on the system, any substantial and potentially damaging long term effects of a new technology should be distinguishable from minor effects that are unlikely to rise above the general background trends and noise of agriculture.

1.2 Long term effects

Historically, most innovations in agriculture occurred without attempts to understand their long-term consequences. Even in recent times, the increased intensification of agriculture in Europe, through mineral fertiliser, pesticides and new tillage, released major limits to production such that yields increased steadily for half a century; but other ecological effects were not anticipated, with the result that soil and food webs are degrading and yields of the main staple crops in much of Europe have levelled.

Long term effects of a change in agricultural land use are of two broad types, differentiated by whether they can be measured in contained experimental studies (BEETLE 2007; EFSA 2010a; EFSA Scientific Colloquium, 2007).

The first type of effect takes a long time to rise above the background trends and noise of agriculture. It is nevertheless feasible to quantify this type of effect by experiments in growth room, glasshouse and field plot. For example, a suppression of the arable seedbank and its local food web might not be detectable until after several years of applying a consistent herbicide treatment such as that associated with a GM herbicide tolerant (GMHT) crop. However, the effect would undoubtedly be measurable in a small field plot.

The second type of long term effect remains unrealised until the GM crop is exposed to the inevitable complexity of the agricultural environment following commercialisation and widespread growing of the crop. For example, the development of the *Brassica* complex of rapeseed crops (*Brassica napus* and *B. rapa*), volunteers, ferals, wild relatives and their hybrids (Squire et al. 2013), having the potential for local evolution of populations with new properties, could not have been discovered by experimentation prior to the great expansion of rapeseed in Europe from the 1970s. This second type of effect is very difficult to predict and is generally not anticipated. It is this second type that is the main focus of attention in AMIGA WP3 on *Long term effects*.

There is no attempt to put any limit on what is meant by 'long term'. The actual duration, whether measured in hours or millennia, depends on the process and context. For example, soil in Europe is still changing as a result of the immigration of cereal farming from west Asia after the last ice sheets retreated north several thousand years ago.

1.3 Multifunctionality – the chain from ecosystem services to interventions

An underlying assumption in the study is that the systems into which GM crops might be introduced are explicitly multifunctional: they deliver more than simply an agricultural product. The systems have to sustain their fabric, regulate the movement of materials through them and generate a landscape that people live and work in. Increasingly, multifunctionality is understood and expressed through the concept of 'ecosystem services' (ES) when considering the utility of land to humans (Millenium Ecosystem Assessment 2005; UK National Ecosystem Assessment 2011; Firbank et al 2013; Smith et al. 2013). Ecosystem services are usually categorised into four types.

Supporting

- soil condition – physical, chemical, biological
- food webs – plant, invertebrates, microbes
- energy and matter cycling – solar, fossil, carbon, nitrogen, phosphorus

Provisioning

- agricultural – the balance among crops, livestock, forestry
- agronomic and economic – food, feed and financial gain
- markets, food security, imports, exports

Regulating

- hydrological process – flood prevention, water storage
- pollution – reduction, control
- pests and disease – prevention and control of invasions, epidemics

Cultural

- landscape and wildlife, iconic biodiversity
- clean air and water
- a living from the land
- regional foods and other products that have significance to people and societies beyond the provision of material food and income from farming

Ecosystem services are not linked directly to an intervention such as a GM crop. The crop or its management must first affect an ecological process and it is this process that determines whether the relevant ecosystem service is satisfied. The crucial question here is the direction in which the links between the intervention and the ecosystem services are examined.

1.4 The flow of ideas and information in assessment and design

The work in AMIGA WP3 examines the direction of querying. Should the analysis -

- begin with the GMO and consider its effects on ecological processes – this is here named a GM-centred approach; or
- begin with the system and consider which interventions are needed to satisfy its outputs – a system-centred approach?

Most investigations, including those underlying much environmental risk assessment to date, primarily operate through the GM-centred approach. They might deal in ecological processes, but the questioning is firmly based in characterising the GM0 and asking about its effect on other things. In EFSA 2010a, these things include persistence and invasiveness, target and non-target organisms, biogeochemical cycles and associated effects on management. In each instance, the risk assessment considers what the GM crop does, compared to its non-GM equivalent. In contrast, the work here will examine the proposition that the system-centred approach may be a better way to examine the questions on long term effects posed above in 1.1.

The first schematic below (Fig. 1.1) shows the flow from 'the change ', such as a new crop or practice, through 'effect on functional life forms' to 'effect on ecological processes' to 'impacts on ecosystem outputs or services'. The term 'functional life forms' is used here to emphasise that the organisms in question undertake a function that alters an ecological process.

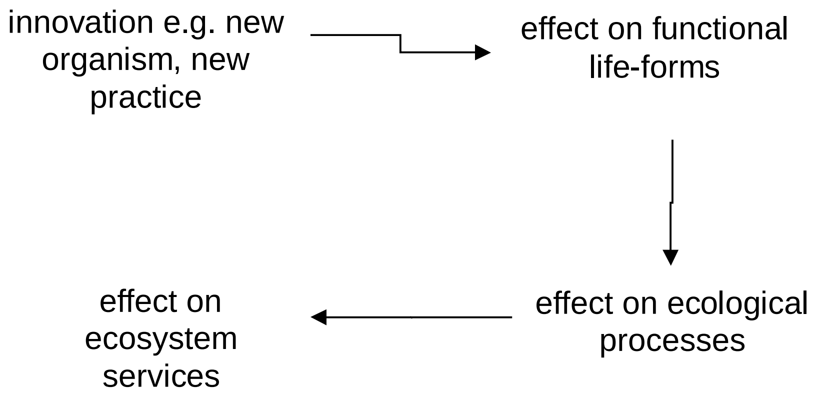

Fig. 1.1 Diagram to show the direction of enquiry in an innovation-led ecological assessment that examines the effect of an innovation on organisms (life forms), ecological processes and ecosystem services.

The innovation is usually a change in a crop and its associated management. One or other or both has some influence on life forms which might include the crop plants, weeds, animal trophic groups and soil microorganisms. These life forms mediate ecological processes such as primary production, nitrification or decomposition, which in turn determine whether an ecosystem service is satisfied or not. For example, a herbicide tolerant variety and its management are observed to reduce weeds (functional forms) which might alleviate competition with the crop for resource (process) which results in a greater food production and profit (service); or else the decline in weeds might reduce the activity of the food web (process) and thereby limit the number of flowers that get pollinated (service). This type of design is largely reactive.

An alternative approach first defines the needs and outputs of a production system in terms of ecosystem services (or similar high-level descriptors) and proceeds in the reverse direction. The ecosystem services would be first defined and target limits set for each service. The ecological processes that give rise to the services are identified, and then the combination of life forms (e.g. crops and weeds) and intervention (e.g. agronomy) that are thought best to deliver the services. This direction is proactive, more akin to ecological design than risk assessment (Fig. 1.2).

1.5 The comparator

When assessing an environmental risk, and specially by the process from innovation to ecosystem service as in Fig. 1.1, the entity under consideration is compared against something whose role and impact are already known or partly known. This approach, of comparing an innovation against a comparator that is *substantially equivalent* and has a known history of safe usage is also the basis of risk assessment of GM products for use in food or feed (EFSA 2010a). For example, maize flour has been used for millennia and its nutritional effects are well known and appreciated; if a maize variety produced by recombinant technology is equivalent in food quality to a conventional maize, then the new type may be judged to be safe.

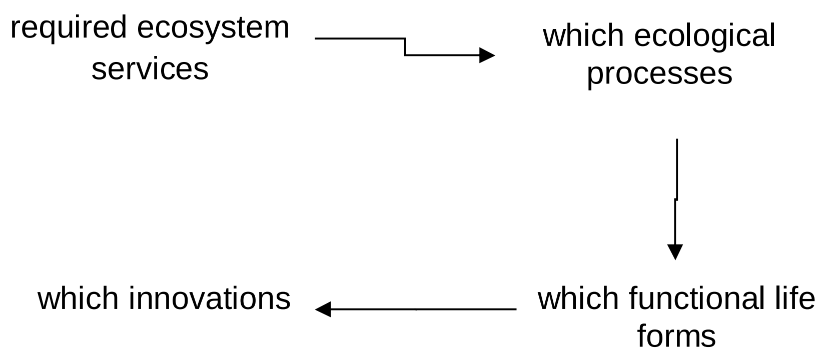

Fig. 1.2 Diagram to show the connections in ecological design working from ecosystem services to life forms and interventions (see text for explanation). Arrows indicated the direction in which questions are asked. After Squire et al. (in preparation).

However, a similar procedure (comparator of substantial equivalence) might not be the most effective way to assess a new crop or field practice. The comparator might still be something that is familiar and has a long history, but it might not be ecologically safe. For example, modern, high-intensity farming has been around for decades, but is having deleterious effects on a number of ecological indicators (Royal Society 2009; Stoate et al. 2001; Marshall et al. 2003). If therefore a high-intensity cropping sequence is used as a comparator for a new variety of cereal or potato, for example, and the innovation shown to be no different, the outcome could be the perpetuation of an unsafe practice.

This question of the comparator is one that figures highly in the current assessment of long term effects. When the chain is worked from innovation to services (Fig. 1), at least three systems need to be compared – the current practice, the current practice with the innovation, and an ideal state that is ecologically safe and sustainable. So far in environmental risk assessment, this third state is not usually considered. This third state might not exist, at least outside an experimental farm, in which case it can only be defined theoretically.

If the chain is worked from service to innovation (Fig. 1.2), the same three systems are compared but the ideal one is the benchmark and starting point. The analysis estimates what needs to be done to move the current system to the ideal state, and this includes identifying those innovations that would assist the process. In this approach, the GM crop or innovation may be almost incidental – if it assists or facilitates the movement, it is acceptable, but if not, it is rejected.

1.6 Scale and interaction

The schemes in Fig. 1.1 and Fig. 1.2 are subject to aspects of scale, or rather interactions across scale, that are different in quality.

The method in Fig. 1.1 (innovation to service) is primarily conducted first within the production unit and then perhaps extended to larger scales. A GM crop would pass through tests at laboratory, plot and semi-field scales and finally released for commercial use. Certain endpoints might be resolvable within the production unit during the early stages. For example, impacts on seedbank species and their associated food webs, or on decomposition or nitrogen transformations in soil. But further impacts cannot be assessed at the initial scales because they depend on levels of complexity after commercial release that cannot be reproduced at that scale.

In contrast, the method in 1.2 (service to innovation) explicitly begins at a range of temporal and spatial scales and narrows down to determine which functions at which scales need to be manipulated to achieve the aims. The distinction arises because some ecosystem services can be satisfied at scales of the production unit whereas others can only be achieved when many units and other parcels of land operate together. For example, a provisioning service such as the production of 'high quality grain to sell for milling' might be satisfied within a unit such as a field or farm. The national statistics on output of grain for milling are derived by adding up all the individual fields (or a subsample of fields) that produce such grain. But a regulating service, such as control of the spread of a fungal disease, might only be satisfied by concerted 'area-wide' action across a catchment or even a continent. Similarly, the provision of high-purity seed for distribution to farmers may need some understanding and application of separation distances between seed crops and production crops to avoid cross pollination. The service-to-innovation route has to begin by considering which scales are important.

More specific examples to demonstrate scale will be discussed later in this document at 3.3.

1.7 Sensitivities and limits of concern

In order to be able to approach many of the questions raised above, it is necessary to define 'limits of concern' or 'safe ecological ranges' in which ecological systems and processes can operate sustainably, that is without suffering long term malfunction or degradation. This is so whether the concerns are about GM crops or any other forms of human-induced change.

The concept of limits of concern is emerging in risk assessment studies within AMIGA and differentiates those effects of (say) a GM crop that simply alters a variable and those effects that move or keep a variable in an ecologically damaged or damaging state. The concept has not yet been widely applied in environmental risk assessment for GM crops but will be explored in AMIGA Task 3.5 using limits definable from existing data and expert knowledge.

The concept is illustrated in Fig. 1.3. The jagged line shows the progression of an entity – e.g. an ecosystem service, an ecological process, a population of organisms - through time or multivariate space. While the process remains within range A, it can operate indefinitely without harm. If it goes outside range A but remains within range B it still operates, but sub-optimally. Outside B, the process deteriorates to collapse. The process in the diagram is seen to move outside range A on several occasions, and where an * is shown, work is needed to bring the process back within range A. In the agricultural context work might include, soil cultivation, changing the cropping pattern, introducing a new crop variety or altering the fertiliser regime.

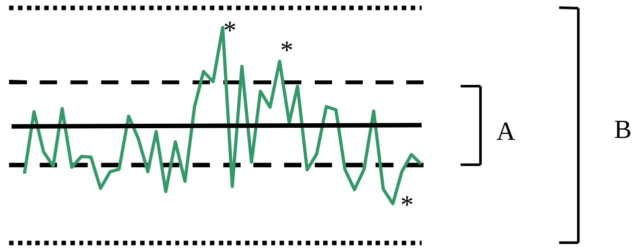

Fig. 1.3 Diagram to illustrate the concept of safe limits. After Squire et al. (in preparation).

The difficulty with applying this approach is that of quantifying as precisely as possible, in terms of indicators, the limits of A and B. The information necessary might simply not be available. Moreover, the limits set for one production ecosystem may not be the same as those for another. Therefore, the quantification of A and B may sometimes have to be done though expert knowledge and opinion, which might be sufficient only to set the direction of an effect: for example, if an ecological process is in range B, then an improved design might only be able to move it towards A without knowing precisely the limits of A.

1.8 Indicators in current Environmental Risk Assessment (ERA)

While new concepts and new methods will be developed in AMIGA, indicators are already widely used in ERA. To give an example – if, following the direction in Fig. 1, the intervention under scrutiny is a herbicide tolerant crop, then the indicator for the intervention might be herbicide usage (quantified by active ingredients, formulations, timing and mode of application of a chemical pesticide, mass per unit area, toxicity index). Indicators for life forms might be phenotypic and compositional characteristics associated with a crop species or variety or a weed population defined by abundance of functional types. The processes affected would be competition for resource and primary production, quantified by rates of plant growth and nutritional quality. The services might be provision of food or other material, quantified by the contribution of the mass or profit of the crop to the needs of the farming enterprise (units: t, £).

The categories of risk assessment in EFSA 2010a are each quantified through measurable indicators. The categories in current use, including persistence and invasiveness, interactions with non-target organisms, biogeochemical processes, are not themselves ecosystem services or ecological processes, but they are categories in which several combinations of the service-process link could be defined. So under 'biogeochemical processes', one regulating service could be 'reducing GHG emissions' and the process 'release of nitrogen gases in the soil by microbial activity'. Under non-target effects, a service could be 'supporting the generation and maintenance of soil' and the process 'decomposition and incorporation of plant litter'.

As argued earlier, the specific purpose of indicators in AMIGA WP3 is to provide quantification that would enable us to answer the broad questions posed in section 1.1 above, particularly about the background trends and dynamics against which GM impacts might be compared. Therefore, the line of argument in the body of this report and subsequent Tasks is to examine the links between services, processes, life forms and interventions in a more comprehensive and inclusive manner than is currently the norm in ERA in Europe.

1.9 Outline of the study

- Examples of the methodology and the indicators used in the chain - intervention-life forms-process-service – and the limitations of this direction of enquiry are presented in Chapter 2.
- The need for a broader set of indicators that would be required in a more comprehensive approach taking the direction in Fig. 1.2 was assessed by experience in two experimental studies (2012, 2013) and is summarised in Chapter 3.
- The more comprehensive set of indicators that was suggested in Chapter 3 is presented in Chapter 4.

2. EXAMPLES OF A GM-CENTRED APPROACH - INNOVATION TO ECOSYSTEM SERVICES

Most existing, including the most well documented, approaches to ERA have been GMO-centred, asking what is affected by the GM crop and whether the effect is substantive enough to merit adoption, prevention or monitoring. The use of indicators (measurables) in the GM-centred approach is here illustrated by reference to case studies of GM herbicide tolerance and insect resistance.

2.1 Genetically modified herbicide tolerant (GMHT) crops in Europe – impact on weeds and food webs

The first example is based on field experimentation around the proposed introduction of GMHT (food-quality) crops. The genetic modifications in question produced tolerance to a particular broad spectrum herbicide. Two questions arose with respect to the introduction of GMHT in Europe.

- Would the state of farmland biodiversity and food webs be adversely affected by the new crop and associated practice?
- Would coexistence of GM and non-GM in the same environment be feasible and would it lead to downstream effects that were economically and environmentally adverse?

Normally, the topic of the second question - coexistence - is seen as an economic issue, rather than one of ERA, but in the broader scope is considered here, because the act of implementing coexistence schemes would inevitably affect a range of ecological processes at various scales. Coexistence has been considered for a range of crops and is examined below at 2.2.

Indicators in the UK's Farm Scale Evaluations

The Farm Scale Evaluations in the UK remains one of the largest field experiments in arable land (Firbank et al. 2003; Squire et al. 2003). The concept was to introduce a new form of weed management, tried in other parts of the world, in which crop varieties that were little affected by a herbicide would be grown with that herbicide being used as the main form of chemical weed control. The herbicide would control the weeds, but do little harm to the crop. Three crops – winter oilseed rape, spring oilseed rape and maize would be accompanied by the herbicide glufosinate ammonium, while beet would be accompanied by glyphosate. The herbicides were 'broad spectrum' in that they killed or seriously set back most species of weed.

This form of management was intended as an alternative to the existing forms. It was not expected to cause any major rise in yield since weeds were already well controlled. In general, and particularly in the case of maize, the broad spectrum herbicides were less toxic than the ones in current usage. There were potential benefits therefore.

The question revolved around the dual role of weeds – their suppression of the crop by competition for resource, and their support of the farmland food web. As indicated, GMHT was unlikely to revolutionise the former but it might compromise the food-web role, which was already under threat from decades of intense crop management. The crucial difference therefore between GMHT and conventional management was that the GMHT could attack the weeds and food web later into the year when all other forms of control could not be implemented because they would reduce yield.

The processes of the weeds and food web are complex, involving the prior survival of seedbank and invertebrates, the uptake of resource by the weeds and crop, and the transfer of that resource to invertebrates through consumption of living or dead matter. Since not everything to do with these processes could be measured, a suite of indicators was developed and applied over three years on commercial fields. Ranges of the type in Fig. 1.3 were not set in advance, largely because safe ranges, defined in terms of population numbers, etc., were not known. Differences of around two-fold, which is the sort of signal that was considered important within statistically noisy, farmland food webs, were taken as signifying a large effect, and replication sufficient to detect such as effect was estimated by statistical power analysis (Perry et al. 2003).

Examples of the indicators measured and the reasons for measuring them are given in Table 2.1. Most indicators had been used in previous work and were only tested and refined for the circumstances of the trial.

Table 2.1. Examples of main indicators used in the Farm Scale Evaluations of GMHT crops and subsequent analyses. From Champion et al. 2003; Heard et al 2003a; Hawes et al. 2003; and other papers in the same issue; Bohan et al. 2005; Hawes et al. 2009; Squire et al. 2005.

Indicator	**measurement / units**	**method / source of data**	**purpose**
Site			
field area, margins and boundaries	various: ha, m	field survey backed by GIS	background and context, identification of untypical situations
texture - % sand, silt, clay, pH	qualitative texture (e.g. sandy loam), % clay	soil sampling and standard lab processing	as above, covariate in analysis
soil organic carbon, soil total nitrogen	% by mass or g m^{-2}	lab analysis of soil samples	as above, comparison with regional soils
Agronomic			
cropping history (8 years)	name of crop, season	reports from farmers	background and context, surrogate for farming intensity
timings of sowing, tillage, herbicide application, harvest, etc.	day in year	reports from farmers, field observations	interpretation of weed, crop and invertebrate data
pest incidence (usually by commercial agronomists)	(variable)	field walks	to define the timing and type of pest control in conventional and GM treatments
fertiliser, also strategy for application	kg ha^{-1}	reports from farmers	context and background, identification of untypical situations
pesticide (usually determined by commercial agronomists)	active ingredients, formulations, mode of application	reports from farmers	interpretation of impacts of weeds and invertebrates; definition of the comparators
tillage	machinery, depth	reports from farmers	interpretation of impacts of weeds and invertebrates; definition of the comparators
Seedbank and vegetation			
seedbank, by species	(number of individuals) m^{-2}	emergence from soil samples	check baseline conditions, relate to previous cropping intensity, assess carry over effects to subsequent

			years
plant population (crops and weeds)	(number) m^{-2} for each species or group	counts in a quadrat	impact of weed management during the season and subsequently
weed mass	g m^{-2}	sampling before harvest and drying of weed vegetation	impact of weed management on plants and food web
crop and weed ground cover	% of ground covered	regular visual estimates	compare crop and weed canopy expansion (later converted to intercepted solar radiation as a crop-weed comparator)
seed rain, by species	(number) m^{-2} for each species	seed rain traps under the canopy	impact of weed management at end of season
vegetation of field margin	(various)	counts in designated strips	effect of herbicide drift, background and context for sites
Invertebrate trophic groups			
plant-living invertebrates	(number) m^{-2} for each species of group	vortice suction sampling	size and diversity of populations in relation to resource and habitat
invertebrates moving on the soil	number for each species or group	pitfall trapping	size and diversity of populations in relation to resource and habitat
bees and butterflies	number for each species or group	observation in the crop	size and diversity of populations in relation to resource and habitat

Weed samples were further classified into functional groups such as monocot (grass) and dicot weeds, and finer groupings related to life history traits such as annuality and determinacy. Invertebrate samples were classified into herbivores, detritivores, predators and parasitoids, and more finely by life history traits (Hawes et al. 2008). These functional group characteristics gave rise to firm conclusions on the impact of the GMHT management on some of the underlying processes (Firbank et al. 2003; Perry et al. 2004; Bohan et al. 2005; Hawes et al. 2009). The populations were further categorised to construct a food web linking plants to the various invertebrate functional groups (Hawes et al. 2009).

Extension of the data to higher scales and interactions

Most of the entities measured in Table 2.1 are indicators of a state or process occurring in a field. The indicators attempted to define happenings at the spatial scale – the field and field margins - at which the new technology would be introduced. However, some processes at that scale were not measured, largely due to cost. Perhaps the most important of these was some measure of crop mass or yield and an associated comparison of the economics of the two treatments.

Several indicators were later used in studies that extrapolated the results to higher spatial and temporal scales and for other purposes than to assess food webs. The extensive data on oilseed rape as a crop, on seed drop, presence of volunteer weeds and volunteer seedbank (all from Table 2.1) provided information on the oilseed rape life cycle that was relevant to Table 2.2 on GM coexistence. Notably, the comparison of seedbank, baseline soil conditions and previous crops helped define the conditions under which volunteer oilseed rape was present in fields (Debeljak et al. 2008), while the data on all aspects of the life cycle of the plant was used to structure and populate models of dynamics and persistence (Begg et al. 2006).

A different type of subsequent analysis showed the potential of the data to examine impacts at a higher spatial scale. In this, the information on weed populations at each site were up-scaled using methods of 'species accumulation' which showed that difference in terms of number of species, which were small when assessed as means between treatments, were magnified when number of species was accumulated through all sites (Squire et al. 2009). The general lesson was that any treatment that suppresses weeds indiscriminately can have large effects on the rarer species in a regional pool. In a separate analysis, information on crop sequence and cropping intensity was used to demonstrate that seedbanks were an indicator of previous management (Bohan et al. 2012).

The interpretation of ground cover measurements (Table 2.1) in terms of intercepted solar radiation is a further example of how indicators aimed a one scale can be drawn on to test concepts beyond that of the original experiment. In this instance, the cumulative solar energy intercepted by the crops can be compared between GM and conventional treatments as a surrogate for crop dry matter, and then with standards in the literature. Energy intercepted by the weeds was shown to be an important discriminant of weed mass over a 100-fold range.

2.2 GM oilseed rape, maize and beet in Europe – gene movement and coexistence

The work on the movement of genetic material among crops, volunteer weeds, ferals and wild relatives had both ecological and economic aims, though much of the funding came to be directed in pursuit of the economic (impurity in crops). The ecological ones included the movement out of fields of GM traits into ferals and wild relatives and subsequent impacts on biodiversity and food webs.

The processes underlying coexistence depend on the persistence, spread and mixing of genetic material in both fields and the landscape. The questions had to be answered by reference to processes at the landscape scale, but needed information and indicators at all scales from the micro-patch in a field. An arbitrary but useful target of 0.9% GM in a non-GM crop had been set in Europe, which can be taken by analogy as a limit of the type in Fig. 1.3. If GM-content is below 0.9%, then a crop, system or harvest is in the range B; if above 0.9%, then it is outside range B. In practice, coexistence managers might identify a range A in which there is no possibility of the crop failing. Therefore, aiming for a situation between A and B would need extra measures to be taken during growth or harvest.

At the beginning of studies on geneflow in relation to GM, in the late 1980s and early 1990s, the spatial units investigated were those of the plot and field, where the persistence of volunteers or short range cross- pollination was measured. But the spatial and the temporal scales soon increased to those of many fields and even several countries as the agents of dispersal began to be understood. So even as early as 1993-1995, landscape scale studies of feral oilseed rape were in place and by 2000-2003, the configurations of sources and sinks in the landscape were integral to understanding gene movement. The main purpose of the work was still driven by the innovation - it was still GM-centred.

Examples of indicators used in many studies, such as those in the EU SIGMEA Project, are given in Table 2.2. Many more methods and approaches are available, for example to estimate pollen and seed viability (e.g. Kjellson et al. 1997), but in a substantial proportion of work related to coexistence, the things measured were limited to some aspect of cross pollination, male (pollen) sterility and distance between source and sink.

Table 2.2. Examples of indicators used in local and regional studies of geneflow and coexistence of GM and non-GM crops, where target plants are typically crops, volunteers, ferals and wild relatives. Sources: den Nijs et al. (2004); Kjellson et al. (1997); Messean et al. (2009); Squire et al. (2013); SIGMEA (2007).

Indicator type	**units / measures**	**method**	**purpose**
Site (variable between studies)			
soil (variable between studies)	variable: texture, etc.	standard laboratory techniques	background conditions, interpretation of seedbank abundance
conditions at soil surface (variables between studies)	(temperature, radiation, water content)	met sites and local sensors	background conditions, interpretation of seedbank and secondary dormancy
weather	temperature, windspeed and direction, saturation deficit (humidity)	met site or mobile weather station	background for all processes, especially pollen and insect movement
Agronomy			
cropping history	name of crops by year	farmers' records	previous source of volunteers
sown seed purity (not routine)	%, gene copies	advanced knowledge from seed supplier, seed testing	check on potential for impurity
timing and type of tillage	day in year, depth, type of machine, weather at the time	farmers' records	interpretation of secondary dormancy in the seedbank
period between harvest or seed drop and next cultivation	day	farmers' records, observation	strong influence on emergence of dropped seed
fertiliser, pesticide	see Table 1	see Table 1	interpretation of crop growth and volunteer life cycle
Plant life cycle			
seedbank and seed rain	(number of individuals) m^{-2}	emergence from soil samples; field traps for seed rain	establish a baseline, assess carry over effects to subsequent years, define, main steps in life cycle
persistence of buried seed	fraction, decline over time; m^{-2}	decomposition from seed buried in bags at various depths	estimation of seedbank decline
abundance of plant populations	(number) m^{-2} for each category	counts in a unit field area	impact of conditions on survival, fitness and reproductive success
plant architecture (not usual)	dimensions (m), branching pattern, root-shoot surface,	measurements on standing or sampled plants	estimate of competitive ability, habitat for higher trophic groups
mass and allocation to reproductive parts of plants	g m^{-2}; ratios	sampling, drying and weighing of vegetation	impact of conditions on fitness, competitive and reproductive success
canopy cover	% ground cover for crops, volunteers, etc.	visual observations, solarimeters	estimate of cumulative intercepted solar energy
seed drop (seed rain)	(number) m^{-2}	catch traps on soil surface	return to seedbank, predation rate of seed
Cross pollination and geneflow			
pollen donors and receptors	species, population size, location	demographic recording, mapping	identifying all sources and sinks for geneflow and introgression

flowering	% plants in flower, number of flowers m^{-2}	field counts or sampling	assess overlap and therefore potential for cross pollination
male sterility	% plants	observation of floral structures, seed company specifications	low male (pollen) sterility encourages high cross pollination
pollen density, deposition	(number) m^{-3}, m^{-2}	various pollen traps, open (e.g. slide) or mechanical (e.g. Burkard)	estimating potential for pollination at distance
cross pollination	% outcrossing, or % trait in seed	various: DNA-based methods; seed colour change; phenotypic tests (e.g. herbicide tolerance)	estimate movement of a trait from donor and recipient in flower
male sterile 'trap' plants	% plants, % potential seed set	location of male sterile plants in the landscape, collection of plants and seed	maximum rate of crossing (useful at long distance when cross pollination to fully fertile plants is low)
purity (genetic) in all life forms	% GM by individuals, genes, copies	range of laboratory methods including quantitative PCR	assess outcrossing, the result of all exchanges, assess status of sown seed
purity (biochemical)	% composition, e.g. erucic acid	various laboratory methods	assess outcrossing, the result of all exchanges, assess status of sown seed
introgression	%, gene copies	estimate of genome size, DNA methods including qPCR	assessment of whether genetic traits have been established in a recipient population
Pollinating insects			
pollinators	number per unit time on standard transect, flower visiting rate, etc.	observation, trapping	relate cross pollination to pollinators
pollen load and composition	mg, fraction of different types of pollen	insect trapping followed by compositional analysis	origin of pollen in the landscape
habitat for pollinators	type of vegetation, location of honey bee hives	field survey	assess distance-relations, landscape scale factors
Landscape characterisation			
distance between sources and sinks	m, km	direct measurement, remote sensing	main discriminant of cross pollination
crop areas, landscape mosaics	various: ha, km	field survey, remote sensing,	regional estimation of outcrossing; interpretation of movement of ferals and volunteers

2.3 GM insect-resistant crops

GM insect resistant crops have received wider study that GMHT cropping. In Europe, the most realistic field studies are arguably those in Spain where *Bacillus thuringiensis* expressing (Bt)-maize is grown commercially (SIGMEA 2007). However, data valuable for risk assessment have been obtained in a range of experimental trials and scenarios in Europe, through experimental projects, notably ECOGEN (2008), SIGMEA (2007) and BEETLE (2007).

The approaches to Bt crops have been broadly similar in each of the main studies. Since Bt toxins are likely to affect organisms that eat the crops directly, eat the organisms that eat the crops or come into contact with the toxins after release to the soil, the main aim is to define those functional groups and species that are most likely to be affected by the particular type of Bt toxin in the plant. The Bt crop does not usually come with any coupled management, unlike a GMHT system. Experience in other countries suggests that assessing the impact of introducing Bt should include factors other than the crop itself, including any structural alterations to the field and its surrounds in establishing refugia to slow down the build-up of resistance, and any change in pesticide usage.

Among the most comprehensive of Bt risk assessments have been those published by CABI in a series of crop- and country-specific case studies, for example that for Bt cotton in Brazil (Hilbeck et al. 2006). Each study considered broad aspects of the cropping system and environment in which the GM crops would be introduced, then dealt in detail with potential impacts on, for example, non-target organisms (NTOs), effects on plant populations through geneflow, the build-up of resistance in target insect populations and effects on soil processes. Great emphasis was placed on defining potential non-target species and selecting ones that were most appropriate for study.

A notable development in Bt studies has been the incorporation of impacts on organisms and processes in the soil (ECOGEN 2008; chapter by Mendonca Hagler et al. in Hilbeck et al. 2006) who consider the routes by which Bt plant tissue might affect soil, the important ecological processes in the systems under study (e.g. plant residue decomposition, nitrification, nitrogen fixation, phosphorus mobilisation, water movement) and the organisms that mediate these processes. Nothing approaching such a 'systems' approach had been attempted for GMHT cropping. Similarly, the economic implications of GM cropping were considered in relation to Bt (ECOGEN 2008).

Given the very substantial literature and experience on indicators for Bt risk assessment, Table 2.3 lists generic headings for indicators used in field studies. References cited in the text provide, or lead to, detailed information on indicators in each generic group.

Table 2.3. Generic types of indicator used in the assessment of insect-resistant cropping. Sources: BEETLE (2006), ECOGEN (2008), EFSA (2010b), Hilbeck et al. (2006), SIGMEA (2007).

Indicator type	**measures / units**	**method**	**purpose / interpretation**
landscape structure	areas, distances, aggregation	mapping	understanding of pest incidence, epidemiology
field structure	areas, dimensions,	mapping	planning and assessment of refugia, effectiveness of IPM strategies
pesticide usage	timing, type, active ingredients, quantity, as in Table 2.1	farm/plot records	defining the comparators, pesticide reduction due to implementing Bt crops
crop performance	phenology, canopy expansion, dry matter accumulation,	field sampling for crop growth and productivity	relative performance against comparator
crop loss	% dry matter loss, spoiling	field sampling	relative performance against comparator
pest incidence	population growth	field sampling	relative performance of Bt and all other measures against comparator
pest resistance	% resistant types	sampling and ex situ testing by toxicological or	assess the development of genetic resistance to the Bt crop in the target pest; efficacy or refugia and

		molecular assays	alternative control
other pest build up	population size, damage, crop loss	field sampling	assess opportunistic population growth and impact of alternative pest following control of target pest
identification of non-target functional groups or species	herbivores, parasitoids, etc.	literature search, local corroboration	narrowing sampling schemes to manageable level
population size of all relevant target and non-target groups	number per unit area, per plant	field sampling by a range of methods	asses impacts of Bt crop and management on trophic processes
soil processes and organisms	variable: population size and composition; e.g. rates of comminution	visual and molecular quantification of soil organisms; process studies	assess effect of Bt crop roots and residues on
geneflow and introgression	see Table 2.2	similar to those in Table 2.2	similar to those in Table 2.2

2.4 Some limitations of the GMO-centred approach

The assessments of both GMHT and Bt cropping illustrate that the use of indicators in experiments can define the chain from intervention to ecosystem services. The transmitted effects of GMHT oilseed rape are shown by the diagram in Fig. 2.1. It begins at the left hand side with the GMHT crop and move rightwards through life forms and ecological processes to ecosystem services. The diagram includes the following indicators, mainly at the field scale but also moving rightwards at the landscape scale.

- The main intervention is the GMHT crop and the change in herbicide profile it brings with it, defined by type of active ingredient and timing of applications; the introduction of GMHT had no influence on field structure or on any other agronomic operations.
- Functional forms include the crop plant, weeds, invertebrate functional groups and the volunteers and ferals that the crop generates over time.
- Ecological processes including energy capture, primary production and feeding interactions, dispersal, hybridisation and geneflow; they would have occurred at the field scale initially, then over landscapes when the crop became widespread.
- Finally to the right are some ecosystem services under categories of supporting (S), provisioning (P) and regulating (R).

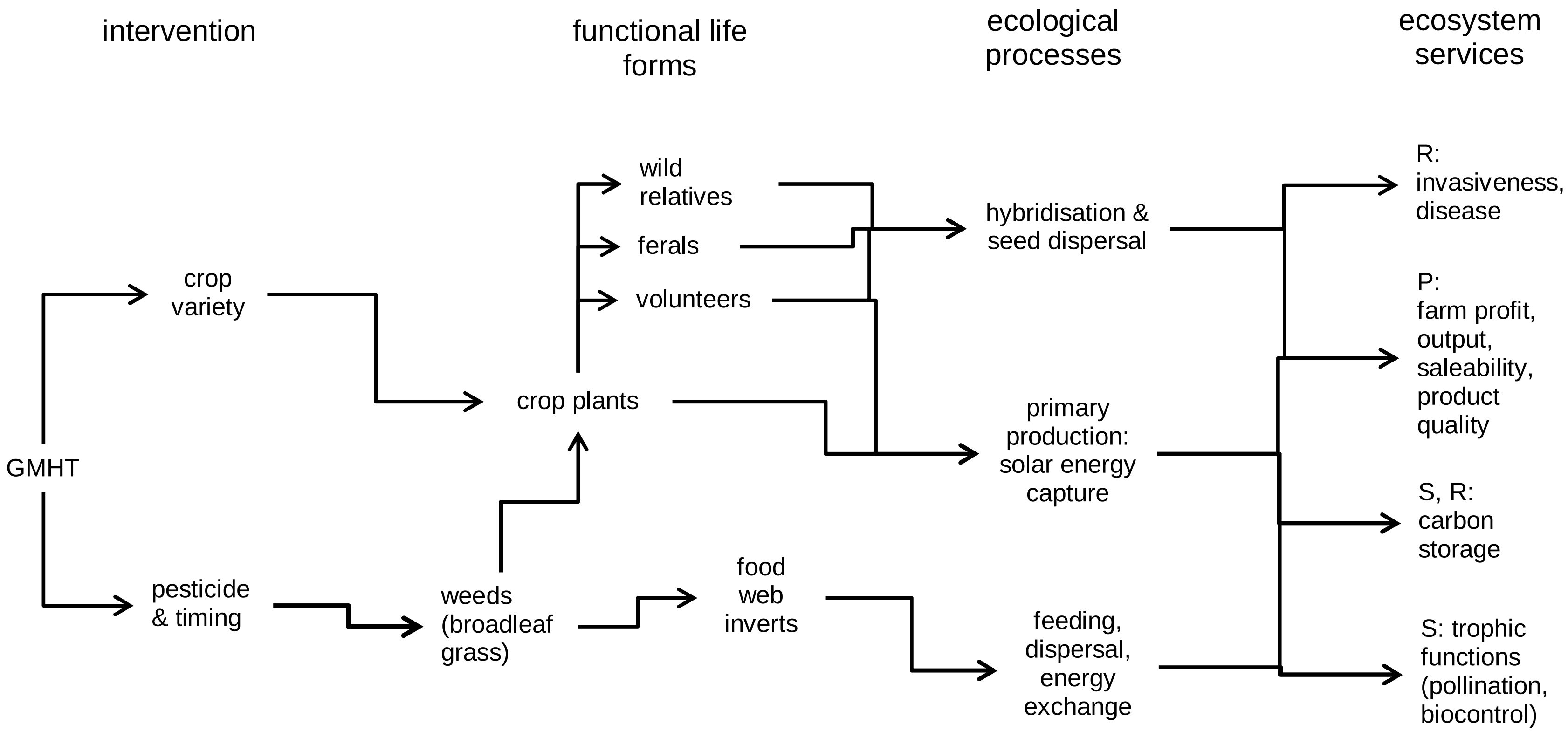

Fig. 2.1 Diagram to show the transmission of effect of GMHT cropping from the original intervention to functional life forms, ecological processes and ecosystem services.

In terms of quantitative changes, GMHT cropping would have had small impacts when considering the ecosystem as a whole. The effects on crop growth were minor and those on weeds were typically 1.5- to 2-fold (which is large for this specific target) and were transmitted variously to the functional groups. Because oilseed rape is grown one every few years, its effects towards the right of Fig. 2.1 would depend on context: if, for example, several other crops could support trophic groups, then its effects via the weeds would be diluted; if it was in a sequence with high intensity crops that discouraged the weed-based food web, its impact would be greater. Because weeds constituted little more than 1-2% of crop mass, and because crop mass was hardly affected by the treatment, the overall impact of GNHT on energy flow and primary production was very small, as was its likely effect on output and profit.

It was the secondary effects, related to complexity, that were more important. The dispersal of seed around the landscape, together with cross-pollination, soon spread the trait to many fields, even those that had not grown GM. In consequence (and as reported by SIGMEA 2007) coexistence of GM and non-GM crops would be very difficult and probably unprofitable. So the effects on provisioning would most likely be negative. In the late 1980s and early 1990s when GM oilseed rape was being considered for introduction, neither the propensity for volunteers and ferals nor long range geneflow were appreciated. They were long term effects which were not anticipated by small-scale experimentation.

The balance of ecosystem services in nearly all GMHT and most Bt studies was not considered in full at the outset. In GMHT work, only a limited range of supporting services were examined (plants and food webs), the contribution to provisioning services was not fully taken into account (e.g. possibly higher yield, better economics) while regulating and cultural services were mostly ignored.

Moreover, the limits of concern or safe ecological ranges were not set out in advance, largely because they were not known and the processes themselves were not fully understood. In the context of Fig. 1.3, and using its terminology, it was thought that intensification had driven food webs well outside A and very near the limits of B; and the concern was that GMHT oilseed rape would further diminish the food webs (and send them beyond B). The size of a substantial impact on food webs in the Farm Scale Evaluations was taken to be about 2-fold but there was no hard evidence as to how far this constituted an ecologically damaging effect. The FSEs were not alone in this – they arguably looked at potential effects more comprehensively than in any previous study.

One of the reasons why the assessments of HT and Bt 'worked', in that they gave clear results and recommendations, often that impact was small or negligible, was that the traits were unlikely to have much effect on any of the main supporting services, such as those depending on the biogeochemical cycles. Neither intervention would alter nitrogen additions, for example. For GMHT in Europe, the intervention would not induce a widespread and permanent effect on tillage in the typical wet, heavy soils. So the experiments and analyses could focus on relatively simple trophic relations. If GMHT oilseed rape had altered aspects of the nitrogen or phosphorus cycles, say, or caused a major change in tillage, then the investigations would have had to be much more comprehensive, and the impacts on functional biodiversity would arguably have been small in comparison to those on biogeochemical cycles and soil structure.

Nevertheless, the approach and conclusions in the SIGMEA project, which included both Bt and HT, and the examples of country- and crop-specific work on Bt cited above, showed a movement towards a systems approach.

3. INDICATORS FOR A SYSTEM-CENTRED APPROACH IN REGIONAL SURVEY AND FIELD EXPERIMENT

The information and indicators needed for working the chain from ecosystem services to innovations (Fig. 1.2) were examined in two field studies, one being a large scale field survey and the other a six-crop, split-field experiment. The aim was to identify those sets of generic indicators that would be suitable for use in a system-centred approach to GM crops.

The adequacy of indicators was assessed by their ability to detect differences caused by types of agricultural practice. No GM crops were used, but the treatments differed in the intensity of field management. During planning of the experiments, it was considered whether tried and tested indicators were available or whether further work would be needed to modify current indicators to ensure they were robust and reliable.

3.1 Field study A. Regional survey of sustainable agricultural practices

The study involved a large-scale survey of more than 100 fields in a high yielding, maritime agricultural region (Hawes et al. 2010). The aim was to identify cropping patterns and management inputs that could deliver the following ecosystem services (ES):

- soil of a quality that was not deficient in any major biophysical attributes and was optimum for plant growth (supporting ES);
- limited use of fossil energy resources, for example in terms of reduced carbon footprint due to using less nitrogen fertiliser and fuel (regulating ES)
- yield that was high, of good quality and profitable (provisioning ES).

Ecological processes were identified that would deliver these ES, and life forms and interventions identified that would mediate the processes. Fields were sampled across a range of soils and climates but no attempt was made to ask farmers to modify what they would normally do.

All three ES above were chosen because they could be satisfied within the production unit. The main centre of attention was therefore the field. Indicators were chosen that were able to differentiate between types of practice: e.g. commercial best practice versus organic; high intensity winter cereal and potato farming versus low intensity spring cereal farming. A factor-of-two difference was expected in attributes such as fertiliser and pesticide use, yield and carbon footprint, while associated differences in soil attributes were uncertain.

The main indicator sets that would enable types of field to be distinguished were as follows.

- soil chemical and biophysical properties (carbon concentration, bulk density, water holding capacity, penetration resistance)
- crop yield (farmers' estimates)
- agronomic practices
- in-field seedbank, in field and marginal vegetation, food webs
- economic aspects of crop production
- regional/national production and inputs (census data on crop area, yield, fertiliser, pesticide)

Certain other groups were not measured because of constraints in time and funding or were estimated from the data collected:

- soil microbial and microfaunal activity (not measured or estimated)
- greenhouse gas emissions and pollution to water (emission not measured, but estimated; pollution to water not measured or estimated)

- carbon and nitrogen pools and fluxes (not measured but estimated from solar energy capture, N & C concentrations, fertiliser inputs).

Results so far (Hawes et al. 2010; Valentine et al. 2012) point to two main findings relevant to the present work. First, that fields could be differentiated according to both farming preference (commercial best practice, organic, integrated) and intensity of management (winter crops versus spring crops) by at least some indicators that were measured or calculated in all the groups above. None were redundant. Second, that several of the cropping sequences fell far short of satisfying more than one of the desired, high-level ES; but that some sequences and strategies could be identified that satisfied at least two, if not all three, of the ES. (The definition of ideal and actual states that provided or might provide multiple ES is not considered here, but will be used along with other data in WP3 Task 3.6.)

One set of indicators – those for regional and national production - proved essential in making the link between regional trends and the current state of fields that had taken different trajectories of intensification over the previous 30 years. Among the other indicators, those for soil biophysical status (Valentine et al. 2012) proved particularly valuable for both baseline site characterisation and differentiating between long-term effects of intensity of management.

3.2 Field study B. Long-term experiment on current and sustainable systems

The second study was based at a 40-hectare field platform - the Centre for Sustainable Cropping at the James Hutton Institute, UK http://www.hutton.ac.uk/about/facilities/centre-sustainable-cropping - consisting of six crops in rotation, subject to two forms of management – current best practice and a 'sustainable' management arranged in a split-field design. The experiment aimed to examine the links in the direction from ecosystem services to innovation. The sustainable management treatment was designed to satisfy multiple ES as listed at the beginning of 3.1 above. Rates of ecological processes were defined and appropriate forms of management put in place.

AMIGA does not fund the infrastructure and measurements at the platform, but will use data collected there for analysis and modelling. The platform does not include GM crops but the aim is to develop a general working methodology that can be applied to innovations of any type. The procedure for designing systems offering multiple ES will be considered under WP3 Task W3.6.

Effects between crop types and management were again likely to differ by a factor of two or more. Effects between treatments would be slower to emerge but could be a factor of 1.5 to 2 cumulative over several years. Indicators sets were then designed to differentiate between these crops and management options. The indicators are similar to those listed under the broad surveys in 3.1 but detailed measurements rather than estimates were implemented in almost every case. The list now reads:

- soil chemical and biophysical properties (carbon concentration, bulk density, water holding capacity, penetration resistance)
- soil microbial and microfaunal activity (samples archived for later analysis)
- crop dry matter, yield, cover, estimated intercepted radiation
- agronomic inputs
- crop quality and composition (varies with crop)
- pest (weed, animal pest, disease) pressure, epidemiology, control
- in-field seedbank, in field and marginal vegetation, food webs
- carbon and nitrogen, measured in soil, plants, invertebrates to estimate pools and fluxes, including N fixation

- greenhouse gas emissions (measured), leachate (surface runoff required but not measured due to constraints)
- economic aspects of crop production
- regional/national production and inputs (for context)

In total, the above categories of indicator are expected to be usable both to design multi-functional systems, to predict potential long-term effects and in many cases to detect differences between treatments. They are all targeted at the sub-field and field scales, as are most current indicators of GM performance. However, the ideal system cannot be designed only at these scales since GM impacts may occur over time at much larger scales.

3.3. Considerations of scale

The incorporation of scale is essential to many of the arguments around long-term effects is and discussed here through the schematic in Fig. 3.1. Any number of scales could be incorporated in this diagram, but here the domain is condensed into five: patch, field, farm, landscape (or region) and national (or global). While the processes at each scale could be examined and quantified independently, the connections between and through scales need to be understood in order to address the main questions in WP3.

Three arrows cross the scales. The one labelled 'markets, policy' includes factors such as the influence of the Common Agriculture Policy, yields and markets in major external producing regions and regional or national protection goals (the latter described in the AMIGA report on Task 2.1, see reference list). Two arrows lead out of the inner box to the widest scale: one dealing with outputs and services and one dealing with losses such as pollution, soil erosion and waste. These two arrows could be classed as one when dealing in some ecosystem services because gains and losses are often part of the same process: for example, gain in yield of a cereal is inevitably accompanied by loss of material as greenhouse gases.

Scale is discussed briefly below in terms of a) the increase in complexity defined as number of possible interactions, b) the scale at which ecosystem services can be satisfied, and c) the scale at which interventions may be introduced. And there is also the need for indicators at several scales to enable cross-referencing: for example, to check that conditions at an experimental site are representative.

Complexity may increase with scale

An important type of long term effect was recognised (1.2) that would not be revealed in experiments conducted under tight constraints. This type only appears when an innovation such as a new crop is widely grown over many years and results from increasing interactions between the crop and its various environments. Examples are given in Table 3.1 for GMHT oilseed rape in Europe. The negative impact of GMHT on the weed flora and food webs (section 2.1, Table 2.1) occurs initially at the field scale, but if the crop were to be grown widely, it would begin to erode the regional species pool because consistent effects in many fields would remove the rarer plant species. Then the difficulties caused by the connectivity among elements of the *Brassica* complex of oilseed rape crops, volunteer weeds, feral plants and wild relatives led to the conclusion that coexistence of GM and non-GM oilseed rape would be almost impossible in northern Europe (Messean et al. 2009; Squire et al. 2013). The results of initial, small scale, field experiments on cross pollination and the survival of ferals were poor predictors of the subsequent interactions in the complex.

Instances were also predicted of knock on effects to other ecosystem services. Over time, the reduction in the weed flora and food webs, though initially affecting supporting services, would negatively affect crop production through reduced biocontrol for example; while some possible management changes to assist in the control of volunteers such as reducing broadleaf crops, though intended to maintain provisioning services, would have knock-on effects to supporting services because the broadleaf crops harbour greater biodiversity than cereals. Interactions of the type summarised in Table 3.1 mean that the performance of a new crop would need to be assessed at a range of scales to appreciate its full impact.

Table 3.1. Interactions envisaged following the commercialisation of GMHT oilseed rape in Europe. Sources: SIGMEA (2007).

	weeds and food webs	**GM coexistence**
primary effects:		
innovation	change in herbicide type and timing	a GM field growing near non-GM field
life forms	weeds and functional invertebrate groups	GM and non-GM crops, volunteers, ferals and wild relatives
processes	population growth and dry matter accumulation among weeds, transfer to higher trophic groups	seed and pollen dispersal, spread and survival of populations, long-range geneflow
services	supporting: functional food web, biological control, pollination	provisioning: saleability of yield, economic cost of coexistence measures, penalty as source of impurity
transmission through scales:		
field	weed flora and food webs reduced in field (within production unit)	crops dropping seed and pollinating within the field
landscape	cumulative biodiversity (regional species pool) reduced among many fields in a landscape	many fields in a landscape connecting mainly through movement of seed
transference to other ES:	reduction of biocontrol and pollination leading to negative impact on provisioning services	change in field management or cropping sequence to control volunteers, e.g. reduction in broadleaf crops, leading to negative impact on supporting services

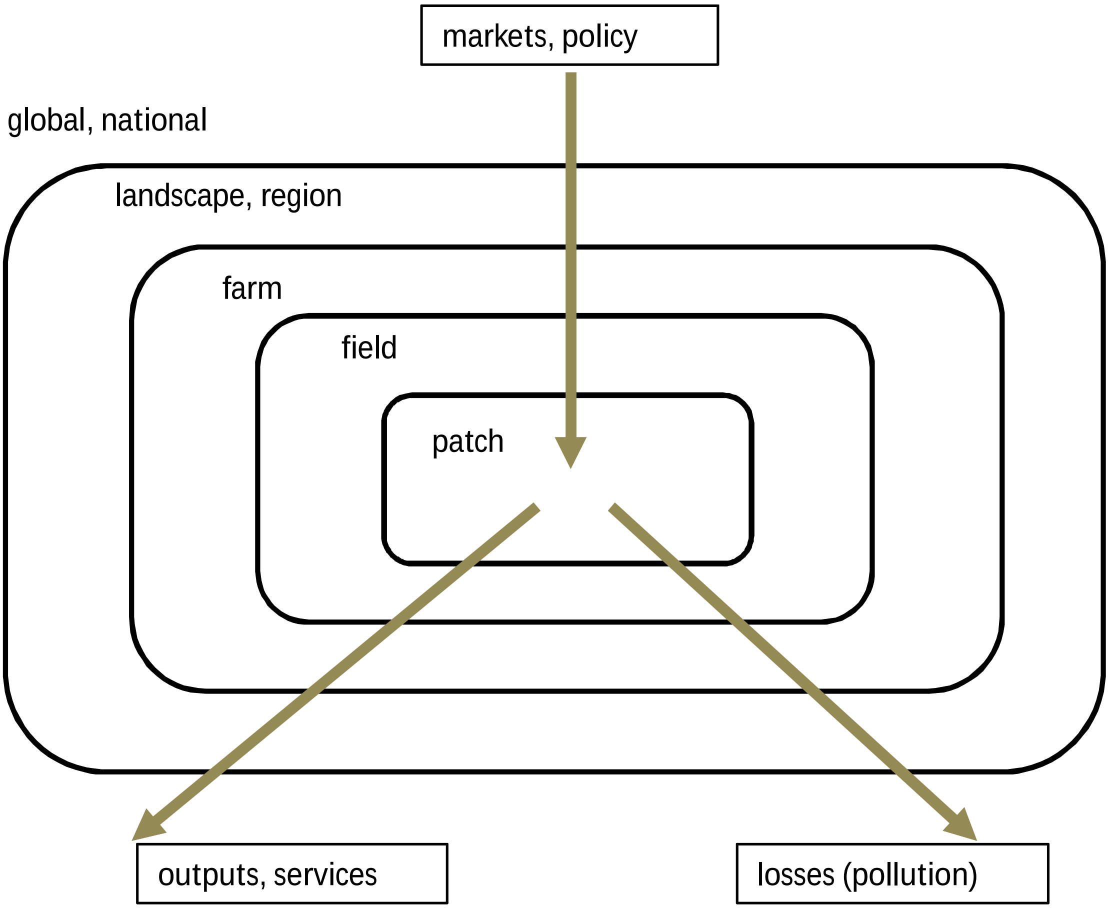

Fig. 3.1 Scales and influence (simplified from an interaction diagram constructed by Dr C Hawes at the James Hutton Institute).

Scale at which ecosystem services might be achieved

The manner in which the outputs and losses can be accumulated across scales depends commonly on the type of ecosystem service. For example, many supporting and provisioning services can be satisfied at the patch or field scales. Ecological processes may include soil formation, nutrient cycling, and primary production, the latter being used to provide food, fuel, fibre and other plant products. Losses from the system also usually measured at this scale include greenhouse gas emissions, nutrient leaching and fuel use. All of these processes are driven or influenced by management decisions that are made at the field scale, including how intensely the crop is managed in terms of fertiliser and crop protectants, field drainage, and use of machinery affecting soil physical structure.

Provisioning services, such as the economic viability of an enterprise, can usually be satisfied at the field and farm scales. They may be influenced by factors at higher scales (see below), but the chain operates within the smaller units. Yield, for example, is often up-scaled additively, in that the yield from a region is often the sum of the yield from all fields. A higher level output such as 'food security' has to be assessed at a national scale or wider, but is still determined by the combined production of all individual farms and fields. Even where a farm is reliant on imports of feed of fertiliser, the supporting and provisioning services are still satisfied at the field scale.

However, many regulating and cultural services cannot be satisfied by what goes on in a small unit in isolation. If the service is to regulate loss of water from a field and control flooding in a catchment, a concerted effort is needed across many fields and other forms of land use to achieve the aim. Similar considerations should be taken into account when regulating the movement of genetic material across landscapes, for example when devising plans to manage GM coexistence. The extent and type of these units and their spatial arrangement are influenced by many management decisions, particularly those affecting the placement of crops across fields (homogenous or heterogeneous), blocking of fields containing similar crops, cropping sequence or rotation and land use (e.g. forestry, game conservation, arable, grass or amenity). The properties of and outputs from a group of farms together produce a catchment with attributes relating to connectivity, habitat mosaics, biodiversity, hydrology and pollution rates. The generation of an attractive landscape is the result of many units operating sometimes individually, sometimes together over wide space and very long time scales.

Scales at which interventions may be influenced and implemented

The scale at which decisions are made to influence an intervention are often different to those at which the ecosystem service can be achieved. Decisions by farmers or landowners at the field and farm scale may be influenced by happenings on a neighbouring farm, elsewhere in the catchment (flooding), in the local market (processing plant availability, transport links, local demand, etc.) and policy at a national or EU level (availability of grants and subsidies, pesticide regulations, etc.).

The influence of such factors need not move through the scales in sequence. For example, a change in EU policy may result in operations at field and farm scales without affecting things at intermediate scales. Or a global deficiency of cereal grain caused by a widespread poor harvest, and hence the likely higher selling price of a crop, might lead to a decision to change a crop directly at a field scale. These matters of scale will be developed later in the study of *Long term effects* in WP3.5 and WP3.6.

Cross-referencing between scales

Many of the indicators deal with services, processes, organisms and interventions at one particular scale. However, some means to enable cross-referencing between scales should be built in to a comprehensive set of indicators. Ideally these indicators should enable performance of the GM crop and comparators at the trial site to be referenced against expected values in the receiving environments generally. For example, if yield at a trial site is as low as, say, one third of the expected regional or national average, or if twice as much fertiliser was used as the national average, then the representativeness of the site should be questioned.

This linking information at the higher scales is of the sort that can be obtained from regional and national statistics. Such data are not available for all types of indicator but there is usually sufficient coverage for three or four categories: crop growth and yield, agronomic inputs, site and soil characterisation, and economics. Examples are given in Table 3.2.

Table 3.2. Indicators that allow cross-referencing between plot/field and national scales.

category	**local – relevant to the comparison at the trial sites**	**wider – relevant to conditions in the receiving environments**
crop	Attributes such as dry matter in the whole	Data on crop yields from national statistical

performance	plant and yielding structures, and components of yield (number, mean mass of grain, etc.) can be used to understand the reason for any difference in yield between GM and comparator, possibly unrelated to the GM trait. Such attributes will also allow comparison across trial sites and with expected norms from national statistics.	records are usually available and should be adequate to enable comparison of general performance at the trial sites with normal or expected performance in the receiving environments with which the trial is associated. The current mean yield of the crop, its expected year to year variation and trends over time (e.g. two or three decades) will allow differences between crops measured at the site to be put in context.
agronomy	Timings of operations for soil tillage, sowing, application of fertiliser and pesticides, harvest, and all inputs of fertiliser and pesticide (kg/ha, formulations, condition of crop and soil, etc.) can be used to interpret crop growth and yield, even if management is intended to be similar between GM and non-GM. Records will allow the management to be compared across trial sites and years and with regional or national statistics (see right). Standard metrics such as 'spray area index' (spray hectares), toxicity indices and greenhouse-gas equivalents (carbon footprint) are valuable for comparison across crop types and systems.	In general, there is less information from government statistical records on actual fertiliser and pesticide input than on yield, but some information should be current, e.g. from crop levy boards and agronomy groups, as to what is best practice. Again, any deviation from what is expected in the receiving environment may be queried.
context – soil and weather	Soil biophysical factors such as textural class and % sand, silt and clay, sometimes with other variables such as carbon concentration, may be available and if not should be measured for the purpose of characterising the site. The assumption should be scrutinised that within-field replication is enough to account for local variation in soil conditions and microclimate. Soil and climatic data can also be used in crop models to normalise performance across treatments and to predict impacts. Previous crops over the last 10 years if available may allow estimates of management intensity if detailed records of inputs are not available.	Soil and weather data are usually available at regional and national scales from government departments. This information can be used to judge the relevance of the trial site to the intended receiving environments. Met records might reveal whether the conditions at the trial site were typical or extreme for the receiving environment. Any preference of a crop for a soil type in the receiving environment may be considered and compared with the soil type of the trial.

3.4 Conclusions

A wider range of indicators than is the norm in GM crop trials would be needed to assess trends caused by a major change in a cropping system or in order to design and plan a system that is sustainable in the long term. In particular, greater attention should be given to three main groups of indicators. One of these is to do with the crop - its phenology, growth, assimilate partition, productive output and resource-use efficiency. Characteristics of the crop – other than those directly associated with the GM trait - have perhaps surprisingly been assumed as neutral in some important GM studies. The two other groups of indicators are particularly important and are linked in that ideal states are essential in both of them if a system is to be sustainable in the long term. These groups are those covering the state and fluxes of energy and matter, those covering the functioning of soil and those necessary for an economic evaluation.

Scales of interest also need to be identified and indicators chosen to enable assessment across scales. Such indicators should not only enable judgement of whether an experimental site is representative of the region or receiving environment, but should facilitate up-scaling or prediction of consequences as an innovation is taken up over the landscape and over the years.

Attempts are made in the next section to achieve a comprehensive set of indicators for use in AMIGA WP3 to consider long term effects of GM cropping.

4. TOWARDS A COMPREHENSIVE SET OF INDICATORS

Chapters 2 and 3 argue the need for a comprehensive set of indicators for use in developing a system-centred approach. Those used widely at present for geneflow and GM coexistence are considered already generic, and through extensive experience, to be adequate for nearly all crops and purposes (Table 2.2). However, sets of indicators are proposed for the following:

- crop and management
- field structure
- energy and matter cycling
- soil biophysical status
- soil microbial and faunal status
- wild plants, food webs and focal species
- pests and integrated pest management (IPM)
- economic
- regional and national census

Examples for each group of indicators are presented in a set of Tables (Table 4.1 to Table 4.9) below. With reference to Fig. 1.2, the indicators are designed to provide information on ecological processes, life forms and interventions. The chain linking ecosystem services to interventions is not usually explicit in the indicators noted. However, the categories of ecosystem services that indicators are most likely to inform are indicated in each rightmost column as S = supporting, P = provisioning, R = regulating, C = cultural.

4.1 Crop performance and management

Some indicators of processes, organisms and interventions in the topics of crop and agronomy are usually measured in field trials. Crucial indicators are often omitted, however, resulting in poor discrimination between crops and treatments. Attributes such as 'components of yield' – numbers of plants, flowering heads, seed and mean seed mass - often contain clues to the cause of difference between sites, years and treatments. A proposed set of indicators in given in Table 4.1.

Table 4.1. Examples of indicators of crop performance and management. Examples: Champion et al (2003); Hawes et al. (2010).

Indicator type	**measurement / units**	**method**	**purpose**
Crop performance, properties and output			
mass output of enterprise	t, t ha^{-1}	yield recording per production unit	assess economic benefit (P)
crop composition	content %, composition, utility as food or feed	analysis on harvested material	assess crop quality and economic benefit (P)
crop tolerance to pesticide	% damage, etc.	field testing, e.g. with range of herbicide dosage	distinguish between crop types (e.g. GMHT and non-GM) in sensitivity to pesticide
crop toxicity to pest and non-target organisms	(various, a wide range of potential measures)	laboratory and field measures of toxic agent; impacts on survival and fecundity	distinguish between crop type (e.g. GM Bt and non-Bt) in ability to deter pests and affect non-targets
development,	time, thermal time,	field observations of	comparator of plant performance

dimensions	area, linear dimensions	emergence, vegetative, reproduction, maturity	(P)
primary production	mass per unit area, t ha^{-1}, per unit time	sampling, drying plants	comparator of plant performance; assess representativeness of crops/location (P, S)
dry matter partition	mass per unit area, t ha^{-1}	sampling, drying plants, separate and weight roots, leaves, stems, heads, etc.	comparator of plant performance – yield alone is not enough (P)
yield components	number or mass per plant, per unit area	separating and counting grain, tubers, etc.	fine scale comparator of plant performance (P)
crop and weed ground cover	% of ground covered	solarimeters or visual estimates	compare crop and weed canopy expansion, convert to intercepted solar radiation as a crop-weed and site comparator
root mass and litter returned to soil	mass per unit area, kg ha^{-1}, per unit time of various fractions	sampling plant dry matter and soil for matter / nutrient content	estimate residual carbon and plant nutrients (S, P, R)
Agronomy (copied from Table 2.1 with amendments)			
cropping history (5 to 8 years)	name of crop, season	reports from farmers or trial managers	background and context, surrogate for farming intensity (P, S)
timings of sowing, tillage, herbicide application, harvest, etc.	day in year	reports from farmers or trial managers, field observations	interpretation of crop, weed and invertebrate data (P, R)
pest incidence (usually by commercial agronomists)	(variable)	field walks	to define the timing and type of pest control in conventional and GM treatments (P, R)
fertiliser, also strategy for application	kg ha^{-1} of nitrogen, phosphate, potash and other main amendments	reports from farmers or trial managers	context and background, identification of untypical situations, interpretation of crop performance (S, P, R)
pesticide (usually determined by commercial agronomists)	active ingredients, formulations, mode of application	reports from farmers or trial managers	context and background, identification of untypical situations, interpretation of crop performance (P, R)
tillage	machinery, depth	reports from farmers or trial managers	context and background, identification of untypical situations, interpretation of crop performance (S, P, R)

4.2 Field structure

Structural characteristics have tended to be recorded where margins and boundaries might be at risk from pesticide treatment or where fields are intentionally partitioned to allow varieties or crops of different type to grow in proximity. An example of the latter is where refugia are established to reduce the rate at which pest targets become resistant to a Bt crop. More widely, aspects of field structure have not been routinely measured, but are valuable for setting baselines and in linking the site to the surrounding habitat for interpretation of IPM and food webs. Examples are in Table 4.2.

Table 4.2. Indicators used to assess structural relations of fields and their immediate surrounds.

Indicator	measurement / units	method	purpose / interpretation
type of boundary	e.g. none, wall, hedge, fence	observation, mapping	baseline characterisation, comparison between experiments and regions (R, C)
managed margin between field and boundary	e.g. none, managed width	observation, mapping	baseline characterisation, comparison between experiments and regions (R, C)
management of margin	species composition, inputs	observation, mapping	interpretation of plant and animal populations (S, P, R, C)
internal strips and patches	e.g. beetle banks, refugia	observation, mapping	interpretation of pest incidence, dynamics, resistance (P, R)
areas, lengths, widths	m, ha	observation, mapping	baseline characterisation, comparison between experiments and regions (S, P, R, C)
reason for margin or patch	e.g. for riparian buffer zone, habitat for biocontrol	discussion with farmers	division of cost and effect among services (S, P, R, C)
crop area lost	%, ha	observation, mapping	cost-benefit analysis of establishment of margins zones, buffer strips and refugia (P, R)
cost of establishing margins and boundaries	e.g. manpower, seed (euros)	farm records	cost-benefit analysis of establishment of margins zones, buffer strips and refugia (P, R)

4.3 Energy and matter cycling (biogeochemical cycles)

While not widely used in GM assessment to date, indicators of biogeochemical cycles are now part of the EFSA 2010a guidelines. Intensification of agriculture has had massive effects on these cycles, notably those for carbon, nitrogen, phosphorus and other plant nutrients. (e.g. Emmett et al. 2010; European Nitrogen Assessment 2011; Smith et al. 2013). This category of indicator is essential to a system-centred approach.

Many of the indicators of crop performance, agronomy and soil biophysical condition can be used in estimates of pools and fluxes of energy, carbon, nitrogen and the rest. However, additional indicators are necessary for a fuller account (Table 4.3). Typical requirements for energy, carbon and nitrogen are given. Other major ecosystem fluxes include water and phosphorus (not shown).

Table 4.3 Indicators used to assess stores and fluxes of energy and matter in agriculture.

Indicator	measurement / units	method	purpose / interpretation
solar and fossil energy			
solar income	GJ m^{-2}	solarimeters, met records	baseline comparator between sites, primary input to plant production
intercepted solar radiation	MJ m^{-2}	solarimeters, plant cover	plant performance, between species, sites, years
conversion efficiency	g MJ^{-1}	calculation from intercepted radiation	plant performance (as above)

		and plant mass	
fossil fuel usage	MJ per unit time, area	from agronomic indicators	used in calculations of carbon footprint, solar-fossil ratio
carbon			
concentration and content in soil, plants, etc.	%, g m^{-2}	see Tables 4.1, 4.4 and 4.6	baseline and general comparator, measured sequentially over time can provide measure of carbon loss, potential sequestration, etc.
decomposition rate in litter bags	mass per unit time	bags contained specified plant litter buried in soil or other media	simple assessment of plant matter and carbon loss
loss to air and water	mass per unit time	see section on N below for approaches	contribution to the overall carbon balance
nitrogen			
concentration and content in soil, plants, etc.	%, g m^{-2}	see Tables 4.1, 4.4 and 4.6	baseline and general comparator, necessary in calculation of N budgets
nitrogen fixation and nodulation in legumes	kg ha^{-1}	as a residual in a nitrogen balance; by direct measurement e.g. using the delta-13 C method	essential process in many low input agricultural systems, feature of high protein legume crops (soy bean, faba bean, pea); presence in a system makes major difference to N balance
nitrogen loss to the air	kg ha^{-1}	sampling of air from field chambers	contribution to calculations of the nitrogen balance; mineral N fertiliser is a major contributor to losses to air (greenhouse gas equivalents)
nitrogen loss to water	kg ha^{-1}	collection of solution in buried field lysimeters and runoff water	contribution to calculations of the nitrogen balance, efficacy of field tillage to reduce runoff
nitrogen deposition	kg ha^{-1}	national sampling scheme	addition to the N inputs with fertiliser and plant residues
phosphorus	(as for N)	as for N, with modifications	as for N, with modifications

4.4 Soil biophysical status

Indicators of the soil function and organisms were used in several studies on Bt crops because the toxin is present in roots and transferred to soil through dead roots and leaf litter (see 2.3). Few indicators of soil status have been applied more widely in studies of GM impacts, due to the GM-centred approach which primarily matches the chosen indicators in any study to the expected impacts. For example, a change to minimum or non-inversion tillage was considered unlikely following the introduction of GMHT oilseed rape to northern Europe due to the types of soil and weather; therefore soil was not examined in detail in most European GMHT impact studies.

Soil indicators have a major role in a system-centred approach since the functioning of soil delivers the most important of the supporting services, and anything that might influence its integrity will indeed have large ecological impacts. A large background literature exists on soil indicators (e.g. Black et al. 2011; Creamer et al. 2009; Powlson et al. 2011; Ritz et al. 2009), but their utility in detecting differences cause by GM crops or management still needs to be refined.

The recent large scale study of >100 fields in the UK, described at 3.1 (Valentine et al. 2012) built on existing knowledge to test the utility of soil variables as indicators of status and potential for degradation under different types of farming. Several indicators, particularly for soil carbon, bulk density and soil hardness, water holding capacity, and air filled porosity, proved robust enough to be able to show differences between intensity of crop management that had developed over two or three decades (Table 4.4).

Table 4.4 Range of indicators of soil biophysical status measurable in the field or in samples taken back to the laboratory, from Valentine *et al.* (2012) and other sources (e.g. Emmett et al. 2010).

Indicator	measurement / units	method	purpose / interpretation
depth, horizons, class, etc.	various	standard soil survey	baseline site characterisation (S, P, R)
particle size distribution	% sand, silt clay	standard laboratory processing of field soil samples	baseline site characterisation (S, P, R)
chemical content (P, K, Mg, etc.) and pH	various	standard laboratory processing of field soil samples	baseline site characterisation (S, P, R)
organic carbon and total nitrogen	% by mass, decomposition bags	various e.g. Element Analyser	carbon sequestration and loss, general indicator of soil condition (S, R)
bulk density	g cm^{-3}	field samples assessed in laboratory by drying and weighing	hardness of soil, indicative of ability of roots to penetrate, proneness to water runoff (S, P, R)
air-filled volume at a range of water contents (macroporosity)	cm^{-3} cm^{-3}	ex situ, on intact soil cores: matric potential adjustment by suction plate	ability of roots and small organisms to explore and move (S, P)
volumetric water content at a range of tensions	cm^{-3} cm^{-3}	ex situ, on intact soil cores: matric potential adjustment by suction plate	ability of soil to hold water for plant growth and other functions (S, P)
penetration resistance	MPa	ex situ, on intact soil cores: needle penetrometer and test frame	ability of soil to allow passage to roots (S, P)
ex situ root growth assay	length m, mass g	soil core removed from field, cereal seed placed on top and root length measured after 14 days	seedling root responses to all the above (P)
ex situ resistance and resilience test	change in function after stress and release	soil sample subjected to a stress (e.g. heat, compression), time course of function e.g. respiration measured	general indicator of soil condition (S, P, R)

The above indicators were applied to examine differences between types of cropping that had existed in fields for 15 to 30 years due to various degrees of intensification. At the high-intensity end of the scale were fields in winter cereals and other winter crops. At the low intensity end were fields in spring cereals, other mainly spring crops and sometime short-term grass. Carbon and nitrogen content, bulk density, volume of large pores, water holding capacity and penetrometer resistance were variously effective in detecting 15-30% differences between the highest and lowest levels of farming intensity.

Such differences between fields under different intensities of management probably take decades to develop, so are unlikely to be manifest in comparison of GM and non-GM crops over one or a few seasons. However, these indicators would be essential in defining the status of a trial site and the status of a receiving environment, especially in an approach that was system-led. The position of a site or region in relation to safe ranges would be an essential requirement in a system-led approach.

4.5 Soil microbiology and faunal status

A wide range of techniques have been used to characterise soils for biological properties and organisms, including general profiling of microorganisms by TRFLP, phospholipid fatty acid analysis and enzyme activities, induced respiration and carbon dynamics and the extraction and identification of microarthropods and nematodes (Black et al. 2011; Creamer et al. 2009; Ritz et al. 2009). As indicated at 2.3, some Bt experiments examined microbes, nematodes and larger organisms such as earthworms (e.g. ECOGEN 2007).

Four groups of indicators are being examined in AMIGA WP4 (Biological components of soil fertility): bacteria, fungi, nematodes and earthworms. The aims are consistent with those defined in this report: establish baselines, compare agroecosystems, define limits of concern, assess responsiveness to GM and deliver risk assessment tools. Table 4.5 lists some of the indicator-groups that are being examined and tested on soil from sites growing GM potato and maize. However, the list here should be regarded as interim since the work is likely to generate an advanced set of techniques.

Bacteria and Archaea, together with *nir*K and *nir*S genes, are being quantified (quantitative real-time PCR), following DNA extraction of soil samples and identified by advanced metagenomic approaches to assess the abundance of phyla and main genera. Fungal phyla and genera are also being quantified and identified. Nematodes are extracted and examined by visual and molecular methods to estimate total abundance per unit mass of soil and proportions of feeding groups, such as bacterial feeders, fungal feeders, omnivores, predators and plant parasites.

Earthworm species differ in vertical distribution, burrowing activity, food sources and typical abundance in arable soils. Two functional types were chosen for targeted study as focal species, both generally abundant: endogeic, creating network-like burrows, feeding on strongly decomposed organic matter; and anecic, creating permanent vertical burrows, feeding on decaying plant residues. Two or three species were selected of each type. (AMIGA Milestone 11, Report on Task 4.4: Selection of focal earthworm species for a novel test system of effects of GM crops on earthworms, considering the specificity of European biogeographical regions. Institute of Biodiversity, VTI Braunschweig, Germany).

In addition, soils at each site and treatment are characterised by standard methods for soil carbon, soil nitrogen and pH.

Table 4.5 Indicators of soil biological function under examination in AMIGA in WP4. This work is in progress and will result in refined and tested indicators.

Indicator	measures / units	method	purpose / interpretation
soil bacteria			
bacteria and archaea, quantity and diversity	16S rRNA gene copies/ng DNA, relative abundance of various taxonomic units	molecular methods (quantitative real-time PCR, ultra-deep sequencing)	comparison of baselines among sites and treatments; effects of GM on community composition (S, R)
specific genes – *nir*K and *nir*S, quantity and diversity	gene copies/ng DNA, relative abundance of various phylogenetic units	molecular methods (quantitative real-time PCR, ultra deep sequencing)	effects on genes regulating basic soil processes (S, R)
soil fungi			
quantity and diversity	ITS copies/ng DNA, relative abundance of various taxonomic units	molecular methods (quantitative real-time PCR, ultra deep sequencing)	comparison of baselines among sites and treatments; effects of GM on community composition (S, R)

nematodes			
identification	taxa	extraction, fixation, viewing on slides	comparison of baselines among sites and treatments (S, R)
focal species (phytophagous, bacterial feeders, predatory)	genus and species	search literature on species abundance to select species off global occurrence, etc.	enable consistent, manageable targeted studies across sites (S, R)
community analysis	abundance (e.g. number per g soil) proportions, quantities of feeding groups and taxa	molecular methods (PCR, T-RFLP)	comparison of baselines among sites and treatments; effects of GM on community composition (S, R)
ex-situ (laboratory) studies	(various)	various laboratory tests of responses to GM plants: biochemical, molecular	potential for nematodes to react differentially to GM and non-GM plant material (S, R)
earthworms			
focal species: endogeic	genus and species	search literature on species abundance to select species off global occurrence, etc.	enable consistent, manageable targeted studies across sites (S, R)
focal species: anecic	genus and species	search literature on species abundance to select species off global occurrence, etc.	enable consistent, manageable targeted studies across sites (S, R)
ex-situ (laboratory) studies	various e.g. number per treatment, per unit time	microcosm studies of mating, cocooning and hatching	potential for earthworms react differentially to GM and non-GM plant material (S, R)

4.6 Wild plants, food webs and focal taxa

Indicators of the seedbank, weed flora and marginal vegetation, of the various invertebrate functional groups in the farmland food web and of more specific non-target organisms in the field, have been tried and tested in studies of the type described at 2.1 and 2.3. The recent methodology for defining focal taxa (EFSA 2010b) has occasioned the need for several more formal indicators. For completeness, parts of the lists in Tables 2.1 and 2.3 are reproduced here, omitting field structural and management indicators, but including those for focal species. Indicators of mammals and birds, though relevant to a system-centred approach, are not included at this stage.

Indicators of non-target effects are a major part of AMIGA and are being subject to intensive field testing in WP 5 (trophic structures) and WP6 (pollinators). The list below should therefore be considered indicative only and provisional except for seedbank and vegetation.

Table 4.6 Summary of indicator groups for wild plants, including weeds, invertebrates and focal taxa. Factors of the crop, such as its toxicity, are covered in Table 4.1; see also Tables 2.1 and 2.3.

Indicator type	measures / units	method	purpose
Seedbank and vegetation			
seedbank, seed rain, by species	(number of individuals) m^{-2}	emergence from soil samples, seed rain traps	check baseline conditions, relate to previous cropping intensity, assess

		under the canopy	carry over effects to subsequent years, impact of weed management
plant functional groups	name	place species/genera in functional groups defined by competitiveness, season, life history traits	allow analysis related to potential role in food web; reduction in of number of taxa or categories
plant population and mass	(number) m^{-2}, g m^{-2} for each species or functional group	counts in a quadrat, sampling before harvest and drying of weed vegetation	impact of weed management during the season and subsequently, potential base for the arable food web.
crop and weed ground cover	% of ground covered	regular visual estimates	compare crop and weed canopy expansion, convert to intercepted solar radiation as a crop-weed comparator
vegetation of field margin	(various)	counts in designated strips	effect of herbicide drift, background and context for sites
Invertebrate trophic groups			
invertebrate population density (plant living, crawling, etc.	(number) m^{-2} for each species of group, diversity indices	wide range of techniques: plant bagging, vortice suction sampling, pitfall trapping	size and diversity of populations in relation to resource and habitat
functional groups e.g. herbivores, detritivores, parasitoids, predators	name of group, population size, diversity indices	identification of samples by the above collection methods	construction of potential food web, linkages, categories for selection of focal species
pollinators	(various)	observation in the crop, surrounding vegetation or hive (AMIGA WP5 is developing new methods of measuring pollinators)	size and diversity of populations in relation to resource and habitat
Focal taxa (criteria)			
species exposure	numerical scale	knowledge of life cycle stages in relation to exposure to crop	assess degree to which the non-target species will come into contact with the GM product
sensitivity to the expressed product	numerical scale	summarised information from laboratory and contained studies	the likelihood that the non-target species will be sensitive if exposed
dependence on the crop for food	numerical scale	summarise information from field studies	dependence of the non-target species on the crop for food
abundance	numerical scale	knowledge from population dynamics and demography in the field	the likelihood that the non-target species will occur in sufficient number to be detected
general vulnerability	numerical scale	knowledge from population dynamics and demography in the field	whether populations are already threatened (is the species vulnerable?)
impact on adjacent habitats	numerical scale	knowledge of demography in the field	whether the non-target species could transfer effects to semi-natural habitats

4.7 Pests and Integrated Pest Management

Indicators of pests, used generically here to include weeds, animal pests and diseases, are in widespread use in agriculture to assess the development of epidemics, timing and type of treatment and damage to crops. Such indicators have been probably among the most widely used of any in crop systems. Epidemiological models, which themselves generate indicators, are routinely used to predict pest pressure and best times to apply treatments.

Measures of pests have been made explicitly in GM crop trials where defining the comparator crop and treatment includes quantifying a pest. For example, in the GMHT field experiments (Table 2.1) the conventional and GM treatments were quantified through the number, type and timing of herbicide operations, which themselves were geared to agronomic assessment of, for example, trigger values. More widely, pests have been recorded commonly as part of the management of the whole field or farm, and being judged part of the background, have sometimes not been well recorded.

In AMIGA, interventions in the form of a GM crop are not considered in isolation of other control measures on the farm, which fall under the banner of integrated pest management (IPM). Therefore, indicators both within the field and beyond the field boundary may have to be assessed if the GM treatment is likely to interact with area-wide IPM. Measures to implement IPM have received variable attention in GM trials, except where IPM has been central to the management of pest-resistance, for example, through the inclusion of pest refuges.

The consideration of GM crop plants as being part of an IPM system is being taken forward in AMIGA Workpackage 8. Several indicators of pests are being measured at trial sites by various means, including visual assessments, field counts and traps. The categories in Table 4.7 should therefore be considered as general and indicative.

Table 4.7 Categories of indicator for quantifying pests and integrated pest management (IPM).

Indicator type	**measures / units**	**method**	**purpose**
crop characterisation	presence of resistance 'R' genes, physical resistance traits	information from breeders, plant phenotyping	comparing GM and non-GM crops,
pest pressure	numbers, rate of increase	population counts, epidemiological models	determine of whether and when to apply control measures
pest management / control	see Tables 2.1 and 4.1	see Tables 2.1 and 4.1	quantify efficacy and means of reducing crop damage
pest resistance to 'R' genes	see Table 2.3	see Table 2.3	see Table 2.3
pesticide usage	type, active ingredients, application method, toxicity, persistence	farm records, residue testing	assess potential cost effectiveness, wider environmental damage, impacts on food webs
crop loss	% damage, yield loss	field observation, weight difference at harvest, % spoilage	assess efficacy of pest control and IPM
field structure	area, linear dimensions, mosaics	field records, mapping	quantify IPM measures including refugia
landscape structure	field pattern, crop types, aggregation, connectivity	mapping, IACS records	quantify landscape as a descriptor of farming intensity or area-wide IPM

4.8 Economic indicators

The economic performance of crops and cropping systems rarely figured as an integral part of early ecological studies of GM impacts in Europe. When economic appraisals were made for GM crops and systems, they were more often related to coexistence than to GM impact (Berthaud, 2013; Messean *et al*.

2009). However, economic appraisals were integral to ECOGEN (2007, 2008) while estimates of the potential benefits of GM cropping have been made for a range of GM crops and countries (AMIGA report on D10.1).

Economic and biophysical appraisals are integrated with biophysical assessments in the approach taken in AMIGA. Typical indicators that can be used at field, farm and regional-national scales are summarised in Table 4.8, and as with other ongoing work in AMIGA, should be considered indicative at this stage.

The list is divided into three categories, beginning with 'National and regional – economics'. The EU nations all contribute to the Farm Accountancy Data Network (FADN), which is an instrument for evaluating the income of agricultural holdings and the impacts of the Common Agricultural Policy (see http://ec.europa.eu/agriculture/rica/). A range of data can be derived from this dataset and used to evaluate long term trends. Other data can be derived at either international level (for instance as a spot price on world commodity markets) or via local or regional information. Also data on land values and the value of rented land could provide useful indicators, although there is considerable variation within country and at regional level. Under the heading 'National and regional – adaption' are potential indicators of the propensity of farming to change (adaption). Data on the indicators listed will vary in availability, modernity and collection method although some will be available via FADN and some from other national data sets. National datasets on land quality for agriculture may also vary in their accuracy and accessibility.

At the scale of the field or farm, a range of financial data may be tracked over-time; and indeed many more forward looking farmers will use a range of farm and field level data for benchmarking purposes. For instance, in the UK, a range of per hectare level data can be derived from the Farm Business Survey for such purposes. Other countries have similar surveys which in part help to populate the FADN database. Indicators at the EU level could potentially be derived from FADN. Partners in WP9 have published on the use of FADN data for evaluating the effectiveness of agri-environmental schemes (see http://centaur.reading.ac.uk/16985/ and section 4.10 later).

Table 4.8. Indicators of economic performance being examined in AMIGA WP 10.

name	measures / units	source	purpose
National and regional - economics			
average net margin per hectare	euros	FADN (see text)	linking field economics to national status and trends (P)
price per tonne at farm gate	euros	market data	linking field economics to national status and trends (P)
level of SFP per hectare	euros	government statistics	linking field economics to national status and trends (P)
average environmental payments per hectare	euros	government statistics	linking field economics to national status and trends (S, P)
average price of freehold land per hectare	euros	land markets	linking field economics to national status and trends (S, P)
average price of rented land per hectare	euros	local survey	linking field economics to national status and trends (S, P)
National and regional – adaption			
average age of farmer	years	government statistics	adaption of farmers to change in practice (P, C)
average level of farmer education	secondary, tertiary	government statistics	adaption of farmers (P, C)
average farm holding size	ha	government statistics	adaption of farmers (P, C)
full time verses part time farmer	%	government statistics	adaption of farmers (P, C)
quality of land	land	government statistics	adaption of farmers (P, C)

	classification		
Field and farm			
average gross margin per hectare	euros per hectare	farm and trial records	economics of field operations (P)
co-existence costs	euros per hectare	farm and trial records	economics of field operations (P, R)
average fertiliser costs	euros per hectare	farm and trial records	economics of field operations (P, R)
average pesticide costs	euros per hectare	farm and trial records	economics of field operations (P, R)
average seed costs	euros per hectare	farm and trial records	economics of field operations (P, R)

4.9 National and regional agricultural output and inputs

As argued at 3.3, regional and national data are needed to be able to assess the representativeness of field experiments and on-farm trials. For example if a crop is normally grown in a certain area, on a type of soil or in combination with other crops, but the experiment is consistent with none of these, then its representativeness may be challenged.

These broad-scale indicators are also valuable for the first stage in up-scaling of a change due to GM cropping. So, for example, if experiments with a crop such as maize indicate an increase in mean yield and a decrease in mean pesticide usage, then as a first approximation, the higher effects can be estimated using data on national yield and pesticide usage.

In AMIGA Workpackage WP3.2, the availability of indicators at these large scales was examined for countries and for some more localised but discrete areas in the five regions – Atlantic, Boreal, Mediterranean, Continental and Balkan. National indicators from these regions will be used in the final task WP3.6 (24-48 months) to estimate GM impacts based on field trials. The types of information available are summarised in Table 4.9.

Table 4.9 Indicators at a country scale of agricultural activity, inputs and outputs.

indicator type	measures / units	source	purpose
areas of agricultural land, main crops	ha, 1000 ha	government census	potential food security, export potential of produce; with data on human population size, estimate of home-grown food intake (P, C)
land quality, land classification	qualitative or semi-quantitative grade	government census	when coupled with climatic data, potential output or worth (P)
soil classification	qualitative soil class, various other measures e.g. soil carbon, pH	national soil survey, soil archive (usually government funded)	linking conditions of experiment or field trial to conditions typical of specified agricultural activity (P, S, R)
number and size of holdings	number, %, mean area	government census	value of sector to rural livelihood, with data on related factors (P, C)
people employed in agriculture	number in total, mean per enterprise	government census	value of sector to employment (P, C)
yield per unit area of crops	ta ha^{-1}	government census, crop and animal levy boards	measure of agricultural output and efficiency (P)
fertiliser, NPK and other main	total usage, t; field rate, kg ha^{-1}	government census, trade	yield potential or efficiency (P); with data on source of fertiliser - reliance on imports

plant nutrient inputs		records	(P, C); estimate of internal and external carbon footprint, pollution potential (P, R)
animal feed products	tonnes of total product and origin	government census, trade records	reliance on imports for food security and agricultural viability (P, C); estimate of internal and external carbon footprint (R)
pesticides	type, active ingredient, formulation, mass, per unit area	government census	yield potential or efficiency (P); pollution and environmental damage, with additional data on toxicity and persistence (R)
pest incidence	crop, pest,	growers and trade records	crop yield and economic loss (P)
economic	-	see FADN in section 4.8	section 4.8

Anomalies in accounting

The use of national statistics in setting baseline and trends and in up-scaling has to be considered critically. Problems were encountered in census data when conducting Task WP3.2. Some problems were generic and some specific to particular countries or groups of countries.

The generic issues were to do with the methods of sampling farms to obtain census data. Quantification and analysis of trends in these broad-scale indicators require that data are collected and averaged consistently over the area or time period under investigation. However, national statistics on indicators such as area of crops, yield of crops, fertiliser inputs and pesticide usage tend to be collated from primary information obtained from different samples. For example, the area grown with different types of crop (e.g. maize, faba bean, cereals) might be available from an annual census taken of every farm, but estimates of yield, fertiliser and pesticide for those same crops may be taken from a sub-set of farms, but a different sub-set in each case. Where census practices are well developed, care is taken to ensure statistically sounds methods are used by the national or trade authorities to up-scale from each sub-set to a national average, such that crop areas, yields, fertiliser and pesticide are roughly comparable. However, the methods used in up-scaling are not always transparent.

A second, and more specific, inconsistency occurs where a country changed its boundaries, split into two countries or altered the areas it used for internal accounting. For example, data are available for Czechoslovakia up to 1992, then subsequently for the Czech Republic and Slovakia. And in Poland, the size and boundaries of internal accounting regions or voivodships changed in the early 1990s. A third type of problem arises where data are either missing or unreliable. For example, data do not seem to be available from some countries in Eastern Europe before the early 1990s, while AMIGA partners report that in some eastern European countries, yields and output are said to have been inflated during the soviet era. Care is therefore needed in interpreting trends in crop area, yield and total output before and after this period in parts of Eastern Europe.

An important recommendation to come out of WP3 on long term trends in cropping is that local expertise is essential when defining regional or national indicators and the use of these in up-scaling. In some countries, the original data-tables and explanations are in the national language and therefore inaccessible to many observers.

4.10 Higher level indicators based on combining and weighting

Most of the indicators laid out in Tables 4.1 to 4.9 above are actual things measured in the field or extracted from databases. However, such primary indicators can be worked to derive additional indicators. Examples include ecotoxicology indices derived from pesticide application data, erosivity indices based on soil type and slope, and diversity indices based on numbers and types of organisms.

A particularly valuable indicator is the Agri-environmental Footprint for the evaluation of European agri-environmental schemes, which is derived by the combination of several indicators of farming operations (AFI 2008; AE-Footprint 2008; Mauchline et al. 2012; Purvis et al. 2009). The evaluation outputs are a measure of the environmental impact or performance of farms, which can be used at a single farm level, aggregated across a larger sample of similar farms or repeated over time to establish a trend. The method is intended to be used by stakeholders, whose experience allows then to weight indicators in different categories and reach a conclusion about the impacts of various interventions. The AFI methodology is in many respects a system-centred approach to environmental assessment.

Some of the indicators are based on the more specific agronomic measures of the types in Tables 2.1 and 4.1, for example:

- inorganic fertiliser – water hazard indicator
- soil compaction indicator
- gaseous emissions hazard indicator
- proportion of river margin protected by semi-natural riparian zones
- proportion of farm extensively cropped
- proportion of field margins with conservation strips

The guidelines encourage stakeholders to check economic sources of the type in Table 4.8 (economic) and to make judgements of some attributes, soil compaction for example, from their own experience and even in the absence of indicators of the type in Table 4.4. The result of the analysis is a combined AFI Score which could be used to compare farms participating in schemes and those not.

Another type of higher level indicator is the carbon footprint, or estimate of the contribution of a crop, field or farming enterprise to greenhouse gas equivalents. The overriding factor in the estimation of carbon footprint in agriculture is the amount of nitrogen fertiliser which contributes through both manufacture and field application. Several 'calculators' can produce an index based on inputs of fertiliser, pesticide, fuel and related measures (e.g. http://www.coolfarmtool.org/; and see a PDF by Hillier et al. at http://www.scri.ac.uk/scri/file/PiP/Carbonfootprintingofcropproduction.pdf)

Decision trees and weighting of indicators

The AFI (2008) methodology is based on a decision tree that compares the effect of an environmental scheme on three different components - natural resources, biodiversity and landscape. The weighting of each of these three main branches is in turn determined by the indicators that combine to form lower nodes. Some research projects have used more complex and less symmetrical trees to assemble and weight indicators of integrated pest management (e.g. DEXiPM). Each node in one of these decision tree can be regarded as a higher level indicator. The final index (e.g. of the environmental sustainability of a cropping system) is itself an summary indicator of performance.

It is feasible within GM risk assessment to structure and weight indicators in some form of decision tree that ultimately requires some subjective assessment of the weights of impacts on qualitatively different end points.

4.11 Synthesis, gaps and future developments in AMIGA on *Long term effects*

The indicators listed in Tables 4.1 to 4.9, and those for geneflow and persistence in Table 2.2, will be applied in AMIGA WP3 on *Long term effects* to define baselines, comparators and impacts in a system-centred examination of GM cropping. Several categories of indicator are being developed, refined and tested in other workpackages, namely those for soil biological status (WP4), trophic groups (WP5, WP6), integrated pest management (WP8) and economics (WP10). As new or improved indicators become available, they will be incorporated into the indicator set summarised here.

This formal set of indicators is primarily assembled for the purpose of answering questions of the type illustrated in section 1.1: what large changes have occurred in European agro-ecosystems without GM, how ecologically safe are these systems, and how might GM crops contribute or otherwise to the long term security of these systems? It is not envisaged that these indicators in their entirety would become the norm in environmental risk assessment of GM crops in Europe. Certainly, the revised guidelines in EFSA 2010a introduced or re-emphasised certain topics, such as 'biogeochemical cycles', and a full assessment of the effect of a GM cropping system in Europe at present would therefore need knowledge of many of the indicators listed in Table 4.3. To reiterate, however, the main aim is to use the indicators in Tables 4.1 to 4.9 and 2.2 to populate a system-centred approach which will be developed in the next two phases of AMIGA WP3.

Next steps in AMIGA WP3

The following activities are scheduled for months 24 to 48:

- establish limits of concern and define safe ecological ranges (Fig. 1.3) for representative ecological processes and the life forms that mediate them (Task 3.5);
- assess the degree of long term change that might occur due to the introduction of GM crops, compared to change that has occurred in recent decades due to other factors;
- and assess whether the changes are likely to push agro-ecological systems towards or away from safe and sustainable ecological limits (Task 3.6).

REFERENCES

AMIGA reports

Regional protection goals in different European regions. Workpackage 2, Task 2.1. Authors: Miklau M, Dolezel M, Bartel A, Heissenberger A, Environment Agency, Austria.

Review of currently available economic data related to transgenic crops. Workpackage 10, Deliverable D10.1. Authors: McFarlane I, Park J, Ceddia G, University of Reading, UK.

General references

include Bt in Spain

AE-Footprint (2008) The AE-Footprint project. Development of a common generic methodology for evaluating the effectiveness of European agri-environmental schemes. http://www.footprint.rdg.ac.uk/

AFI (2008) Agri-environmental footprint index. Users' manual, version 1. An output of the AE-FOOTPRINT Project SSPE-CT-2005_006491. Contact: Simon Mortimer (s.r.mortimer@reading.ac.uk).

BEETLE (2007) Long-term effects of genetically modified (GM) crops on health and the environment (including biodiversity). Final Report.
http://ec.europa.eu/food/food/biotechnology/reports_studies/docs/lt_effects_report_en.pdf

Begg GS, Hockaday S, McNicol JW, Askew M, Squire GR (2006) Modelling the persistence of volunteer oilseed rape (*Brassica napus*). *Ecological Modelling* 198, 195-207.

Bertheau Y (ed.) 2013. *Genetically modified and non-genetically modified food supply chains: coexistence and traceability*, First edition. Blackwell Publishing Ltd.

Bohan DA, Boffey CWH, Brooks DR and 15 others (2005). Effects on weed and invertebrate abundance and diversity of herbicide management in genetically modified herbicide-tolerant winter-sown oilseed rape. *Proceedings of the Royal Society of London, Series B* 272, 463-474.

Black HIJ, Ritz K, Harris JA, Cameron C, Campell CD, Chamberlain PM, R Creamer, Pawlett M, Wood C, Singh BK (2012) Scoping biological indicators of soil quality. Phase II. Defra Final Contract Report SP0534.
http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=2&ProjectID=11765

Firbank LG, Heard MS, Woiwod IP and 16 others (2003) An introduction to the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Journal of Applied Ecology* 40, 2-16.

Champion GT, May MJ, Bennett S and 15 others (2003) Crop management and agronomic context of the Farm Scale Evaluations of genetically modified herbicide-tolerant crops. *Philosophical Transactions of the Royal Society of London B*. 358 (1439), 1801-1818.

Creamer RE, Bellamy P, Black HIJ, Cameron C, Campbell CD, Harris J, Parekh N, Pawlett M, Poskitt J, Stone D, Ritz K (2009) An inter-laboratory comparison of multi-enzyme and multiple substrate-induced respiration assays to assess method repeatability in soil monitoring. *Biology and Fertility of Soils* **45**, 623-633.

den Nijs HCM, Bartsch D, Sweet J (eds) (2004) Introgression from genetically modified plants into wild relatives. CABI Publishing.

Donald PF, Green RWE, Heath MF (2001) Agricultural intensification and the collapse of Europe's farmland bird populations. Proceedings of the Royal Society London (B) 268, 25-29.

ECOGEN (2007) Soil ecological and economic evaluation of genetically modified crops. (Ed. BS Griffiths). Pedobiologia 51, 171-269 (Special Issue).

ECOGEN (2008) Final report of the ECOGEN Project.
https://pure.au.dk/portal/files/15949533/ECOGEN_Final_report_Introduction_and_Summary_web.pdf

EFSA (2010a) Guidance on the environmental risk assessment of genetically modified plants. EFSA Journal 2010;8(11) 1879 (111 pp.) (doi:10.2903/j.efsa.2010.1879)

EFSA (2010b) Scientific opinion on the assessment of potential impacts of genetically modified plants on non-target organisms. EFSA Journal 2010;8(11):1877 (72 pp.) (doi:10.2903/j.efsa.2010.1877)

EFSA Scientific Colloquium (2007). Environmental Risk Assessment of Genetically Modified plants, 20-21 June 2007, Tabiano, Italy. EFSA Scientific Colloquium 8, European Food Safety Authority.

Emmett BA, Reynolds B, Chamberlain PM and 12 others (2010) Countryside Survey: Soils Report from 2007. Technical Report No. 9/07 NERC/Centre for Ecology and Hydrology, 192pp.

European Nitrogen Assessment (2011) The European nitrogen assessment: sources, effects and policy perspectives. Edited by Sutton MA, et al. Cambridge University Press.

FADN. Farm Accountancy Data Network. http://ec.europa.eu/agriculture/rica/

Firbank L. Bradbury RB, McCracken DI, Stoate C (2013) Delivering multiple ecosystem services from closed farmland in the UK. *Agriculture Ecosystems and Environment* 166, 65-75.

Hawes C, Haughton AJ, Osborne JL and 16 others (2003) Responses of plant and invertebrate trophic groups to contrasting herbicide regimes in the Farm Scale Evaluations of genetically-modified herbicide-tolerant crops. *Philosophical Transactions of the Royal Society of London B*. 358 (1439) 1899-1913.

Hawes C, Haughton A, Bohan DA, Squire GR (2009). Functional approaches for assessing plant and invertebrate abundance patterns in arable systems. Basic and Applied Ecology 10, 34-42.

Hawes C, Squire GR, Hallett PD, Watson CA, Young MW (2010). Arable plant communities as indicators of farming practice. *Agriculture, Ecosystems and Environment* 138, 17-26.

Heard MS, Hawes C, Champion GT and 10 others (2003) Non-crop plants in fields with contrasting conventional and genetically modified herbicide-tolerant crops. 1. Main effects of treatments. *Philosophical Transactions of the Royal Society of London B*. 358 (1439), 1819-1832.

Hilbeck A, Andow DA, Fontes EMG (eds) (2006) Environmental risk assessment of genetically modified organisms. Volume 2. Methodologies for assessing Bt cotton in Brazil. CABI Publishing.

Kjellsson G, Simonsen V, Ammann K (eds) (1997) Methods of risk assessment for transgenic plants. II Pollination, gene transfer and population impacts. Basle: Birkhauser Verlag

Marshall EJP, Brown VK, Boatman ND, Lutman PJW, Squire GR, Ward LK (2003) The role of weeds in supporting biological diversity within cropped fields. *Weed Research* 43, 77-89.

Mauchline AL, Mortimer SR, Park JR and 20 others (2012) Environmental evaluation of agri-environment schemes using participatory approaches: experience of testing the Agri-Environment Footprint index. *Land Use Policy 29*, 317-328 http://centaur.reading.ac.uk/21548/

Messean A, Squire G**,** Perry J, Angevin F, Gomez M, Townend P, Sausse C, Breckling B, Langrell S, Dzeroski S, Sweet J (2009) Sustainable introduction of GM crops into european agriculture: a summary report of the FP6 SIGMEA research project. *OCL* 16(1)**,** 37-51.

Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-being: Biodiversity Synthesis.* Washington, DC: World Resources Institute.

Perry JN, Rothery P, Clark SJ, Heard MS, Hawes C (2003) Design, analysis and statistical power of the Farm Scale Evaluations of genetically modified herbicide tolerant crops. *Journal of Applied Ecology* 40, 17-31.

Powlson DS, Gregory PJ, Whalley WR, Quinton JN, Hopkins DW, Whitmore PR, Hirsch PR, Goulding KWT (2011) Soil management in relation to sustainable agriculture and ecosystem services. *Food Policy* 36, 572-587.

Purvis G, Louwagie G, Northey G and 13 others (2009) Conceptual development of a harmonised method for tracking change and evaluating policy in the agri-environment: the Agri-environmental Footprint Index. *Environmental Science & Policy* 12, 321-337 http://centaur.reading.ac.uk/9073/

Ritz K, Black HIJ, Campbell CD, Harris J, Wood CM (2009) Selecting biological indicators for monitoring soils: A framework for balancing scientific and technical opinion to assist policy development. *Ecological Indicators* 9, 1212-1221.

Royal Society (2009) Reaping the Benefits: Science and the Sustainable Intensification of global agriculture. London: Royal Society.

SIGMEA (2007) Final report of the SIGMEA project. http://www6.inra.fr/sigmea

Smith P, Ashmore MR, Black HJ, Burgess PJ, Evans CD, Quine TA, Thomson AM, Hicks K, Orr HG (2013) The role of ecosystems and their management in regulating climate, and soil, water and air quality. *Journal of Applied Ecology* 50, 812-829.

Squire GR, Brooks DR, Bohan DA and 12 other authors (2003). On the rationale and interpretation of the farm-scale evaluations of genetically-modified herbicide-tolerant crops. *Philosophical Transactions of the Royal Society of London B* 358 (1439), 1779-1800.

Squire GR, Hawes C, Begg GS, Young MW **(**2009) Cumulative impact of GM herbicide-tolerant cropping on arable plants assessed through species-based and functional taxonomies. *Environmental Science and Pollution Research* 16**,** 85-94.

Squire GR, Hawes C, Bohan DA, Brooks DR, Champion GT, Firbank LG, Haughton AJ, Heard MS, May MJ, Perry JN, Young MW (2005) Biodiversity effects of the management associated with GM cropping systems in the UK. Final Report of Project EPG 1/5/198. Defra, London.

Squire G R, Lecomte J, Husken A, Soukup J, Messean A (2013) Contributions of pollen and seed to impurity in crops – a comparison of maize, oilseed rape and beet. In: *Genetically modified and non-genetically modified food supply chains: coexistence and traceability,* First edition. Edited by Y Bertheau. Blackwell Publishing Ltd.

Stoate C, Boatman ND, Borralho RJ, Carvalho CR, de Snoo GR, Eden P (2001) Ecological impacts of arable intensification in Europe. *J. Environ. Manage.* **63**, 337-365.

UK National Ecosystem Assessment (2011) The UK National Ecosystem Assessment Technical Report. UNEP-WCMC, Cambridge.

Valentine TA, Hallett PD, Binnie K, Young MW, Squire GR, Hawes C, Bengough AG (2012) Soil strength and macropore volume limit root elongation rates in many UK agricultural soils. *Annals of Botany* 110, 259-270.