# Post Market Environmental Monitoring of Genetically Modified Crops

Report of an expert working group of the Advisory Committee on Releases to the Environment

12 March 2013

# Post Market Environmental Monitoring of Genetically Modified Crops

Report of an expert working group of the Advisory Committee on Releases to the Environment

## Executive Summary

This report provides scientific advice on the Post Market Environmental Monitoring (PMEM) of genetically modified (GM) crops. It examines whether existing environmental surveillance networks (ESNs) could be used for General Surveillance for unanticipated adverse effects. The findings have wider relevance for environmental monitoring of agro-ecosystems by making use of ESNs of the United Kingdom to investigate causes of change.

European legislation requires that GM crops undergo a pre-market environmental risk assessment before they are authorised for commercial cultivation. In addition, the legislation describes two types of PMEM. Case-specific monitoring may be required, depending on the outcomes of the environmental risk assessment, to address specific hypotheses. General Surveillance for unanticipated adverse effects is required in all cases, although there is no reason to expect that GM crops would have adverse effects if risks have not been identified in the environmental risk assessment. This report provides advice on both types of monitoring, focussing on the use of two tools for General Surveillance: ESNs and the Farm Questionnaire.

Many different drivers operate synchronously in the arable landscape resulting in both positive and negative impacts. Existing ESNs have not been used routinely to investigate relationships between cause and effect owing to the challenge of identifying individual causes of change against background variability. Despite these challenges, this report analyses the capabilities and limitations of existing ESNs and their statistical power to detect change correlated with GM crop cultivation. Four ESNs are used as case studies: the UK Butterfly Monitoring Survey, the Breeding Birds Survey, the Countryside Survey and the Water Quality Monitoring Programme.

The analysis presented in this report shows that with well chosen indicators and specific data analysis, these ESNs could be used to detect unanticipated adverse effects. There are, however, limitations with respect to the size of the effect that may be detected with high probability, and the speed with which effects could be detected.

This report recommends that GM crops are considered in the context of the wider impacts of agriculture on the environment and initiatives to identify drivers of change through use of existing ESNs. It does not recommend reconfiguring existing ESNs for monitoring GM crops. It would, however, be useful to establish an agreed set of assessment endpoints, monitored by ESNs. This would enable the type of analysis presented here to combine reporting on the health of the farmed environment with searching for correlations with any drivers of change.

General Surveillance can provide a tripwire for identifying changes if they were to result from cultivation of GM crops. It can be seen as offering an additional safety net in that if unanticipated adverse effects were to occur they could be identified earlier. It must be recognised, however, that it will not be possible to use General Surveillance to definitively establish a relationship between cause and effect. If an effect is observed, which could result from GM cultivation, expert opinion would be needed to determine what further action is needed to investigate the cause.

## Key messages and recommendations

### Case specific monitoring

1. CSM provides an appropriate route for the testing of specific hypothesis which can only be addressed under the range of conditions that are represented by commercial farming

2. Identifying the need for CSM should not be seen as an indication that unacceptable risks remain and that authorisation should be refused.

### General surveillance

3. GS should focus on environmental parameters in close contact with the GM crop.

4. As GS is not hypothesis-driven the same approach could be taken for all arable GM crops.

5. GS should use a combination of tools, to maximise the chances of detecting any adverse effects.

6. An agreed set of assessment end points and corresponding measurement endpoints that can be used routinely to report on the status of the farmed environment and identify correlations with possible influencing factors should be defined

### The Farm Questionnaire

7. The FQ is a useful tool which can be employed as part of GS of GM crops and as such its potential as an information source should be maximised wherever practical

8. The FQ should also collect information on cultivation and agronomic practices in fields in the years following GM crop cultivation

### Existing surveillance networks

9. Power analysis is a useful tool for determining the probability of detecting change using ESNs and therefore for interpreting a finding of no significant effect

10. In the UK ESN could be used to detect unanticipated effects resulting from the cultivation of GM crops

11. GS should make use of ESNs which are already in place or will be introduced in line with general reviews of national monitoring requirements.

12. To maximise the value of existing ESNs for GS of GM crops, specific analysis of the data collected would be required.

13. To detect effects more quickly, at a lower level of change, or at a stage when GM crops were less widely cultivated, supplementary monitoring would be needed.

### Further action if potential adverse effects are detected

14. GS may be used to identify correlations between adverse effects and potential drivers of change, but is unlikely to conclusively demonstrate the cause of an adverse effect

15. If a potential adverse effect is detected, independent scientific advice should be sought to interpret data and determine what kind of further investigation should be triggered

**ACRE working group membership:**

ACRE members
Prof. Rosie Hails (Chair), Dr. James Bullock, Prof. Les Firbank, Prof. David Hopkins, Dr. Ieuan Joyce, Mr. Jim Orson, Prof. Chris Pollock

Invited experts
Mr. Chris Chesterton (Natural England), Mr. Mark Clook (Chemicals Regulation Directive), Dr. Jonathan Davey (Science and Advice for Scottish Agriculture), Dr. Terry Parr (Centre for Ecology & Hydrology), Mr. Phil Smith (Environment Agency), Dr. Jonathan Storkey (Rothamsted Research), Dr. Lawrence Way (Joint Nature Conservation Committee)

Statistical analysis
Dr. Gavin Siriwardena and Dr David Baker (British Trust for Ornithology), Dr Stephen Freeman and Dr Peter Henrys (both of the Centre for Ecology & Hydrology), Jenny Lannon, Andrew Davey and Edward Glennie (all of WRc).

Published 12 March 2013

**Abbreviations**
ACRE: Advisory Committee on Releases to the Environment
BC: Butterfly Conservation
BBS: Breeding Birds Survey
BTO: British Trust for Ornithology
CAMERAS: Coordinated Agenda for Marine, Environment and Rural Affairs Science
CCW: Countryside Council for Wales
CEH: Centre for Ecology & Hydrology
CS: Countryside Survey
CSM: Case Specific Monitoring
Defra: Department for the Environment, Food and Rural Affairs
EFSA: European Food Safety Authority
ERA: Environmental Risk Assessment
ES: Environmental Stewardship
ESNs: Environmental Surveillance Networks
FC: Forestry Commission
FERA: the Food and Environment Research Agency
FQ: Farm Questionnaire
FSA: Food Standards Agency
FSE: Farm Scale Evaluations
GM: Genetically Modified
GMO: Genetically Modified Organism
GS: General Surveillance
HT: Herbicide tolerant
JNCC: Joint Nature Conservation Committee
LWEC: Living with Environmental Change
NE: Natural England
PMEM: Post Market Environmental Monitoring
RSPB: Royal Society for the Protection of Birds
SNH: Scottish Natural Heritage
UKBMS: United Kingdom Butterfly Monitoring Survey
UKEOF: United Kingdom Environmental Observation Framework
WCBS: Wider Countryside Butterfly Survey
WQMP: Water Quality Monitoring Programme

## Contents

## Section 1
## Scope and purpose of the report

This report provides scientific advice on the practical implementation of Post Market Environmental Monitoring (PMEM) of Genetically Modified (GM) crops as required by Directive 2001/18/EC. It represents the outputs of an Advisory Committee on Releases to the Environment (ACRE) expert working group (for further information see Annex 1). ACRE provides statutory advice to the UK government and devolved administrations on the potential risks of releasing Genetically Modified Organisms (GMOs) to the environment[1].

This expert working group was established in response to recent initiatives from the European Commission to strengthen the PMEM of GM crops (see Annex 2). In a series of meetings, the ACRE working group considered how PMEM could be practically implemented using scientifically robust principles against the existing EU legislative framework. In this, the need for monitoring to be proportionate to the level of risk was taken into account. This report provides decision makers with advice on options for implementing PMEM in line with the requirements of EU legislation. The main focus of the report is General Surveillance and the use of Environmental Surveillance Networks. To set this in context other aspects of PMEM are also considered, with advice provided on Case Specific Monitoring and the Farm Questionnaire.

## Section 2
## Background information relevant to post market environmental monitoring

### 2.0 Introduction
This section provides background information relevant to Post Market Environmental Monitoring of genetically modified crops. The requirements of the GM legislation are briefly introduced with emphasis on PMEM. Information is provided on which GM crops are currently approved for commercial cultivation in Europe and which crops have applications in the regulatory pipeline. More detailed information is provided in annexes to the report.

### 2.1 Regulatory context
In Europe the commercial marketing of GM crops is regulated according to a set of interconnecting EU legislation. The GM Food and Feed Regulation (EC) 1829/2003[2] and the GM Deliberate Release Directive 2001/18/EC[3] require an assessment of food and feed safety and environmental safety. The Traceability and Labelling of GMOs Regulation (EC) 1830/2003[4] requires that all GM products are clearly labelled and traceable. EU level authorisation is needed before a GM crop can be commercially marketed.

The Department for the Environment, Food and Rural Affairs (Defra) is the UK Competent Authority for Directive 2001/18/EC. The Directive is implemented under the Genetically Modified Organisms (2002) Regulations in England and equivalent regulations in Scotland, Wales and Northern Ireland.

This report considers the requirements of Directive 2001/18/EC. This Directive requires that all GM crops undergo an Environmental Risk Assessment (ERA) prior to authorisation for commercial marketing. GM crops are assessed on a case by case basis and compared to the non-GM equivalent. This includes a comparison of cultivation, management or harvesting practices. A GM crop is only authorised if the risk assessment indicates it is safe for human health and the environment.

Directive 2001/18/EC also sets out requirements for post market environmental monitoring of GMOs, which is the subject of this report. PMEM plans must be submitted as part of applications to market GM crops. Monitoring, and the submission of monitoring reports, is the responsibility of the consent holder. The regulatory authorities must come to a view as to whether monitoring is fit for purpose.

Two types of monitoring are required:

> Case specific monitoring (CSM): to confirm that any assumption regarding the occurrence and impact of potential adverse effects of the GMO or its use in the environmental risk assessment are correct;
>
> General Surveillance (GS): to identify the occurrence of adverse effects of the GMO or its use on human health or the environment which were not anticipated in the environmental risk assessment.

CSM is therefore not required in all cases. It is only needed in situations where, following the environmental risk assessment, a specific hypothesis remains as to how GM crops could cause adverse effects. GS is required in all cases irrespective of the outcomes of the ERA.

The concepts of CSM and GS are expanded on in Council Decision 2002/811/EC[5] and in guidance from the European Food Safety Authority[6] (see Annex 3 for more information on the role of EFSA). EFSA has recently revised its guidance on PMEM of GM plants[7]. The EFSA guidance recommends that GS should make use of three different approaches: a Farm Questionnaire (FQ), the use of existing Environmental Surveillance Networks (ESNs) and a literature review. The advice in this report is based on the assumption that GS will take this form. In general this report takes the EFSA guidance into consideration to provide advice on the practical implementation of PMEM in the UK.

### 2.2 GM crops in Europe

In the EU an application to commercially market a GM crop can be restricted to import and processing or can include cultivation in its scope. Authorisations for commercial marketing of GM crops are valid across the EU. Although more than thirty different GM crops are authorised for import and processing, only three are currently authorised for commercial cultivation. At present GM crops are not commercially cultivated in the UK as there are no approved varieties available with traits of potential interest to UK farmers.

The three GM crops which are authorised for commercial cultivation in the EU are a potato with altered starch properties (Amflora), an insect resistant maize (MON810) and a herbicide tolerant maize (T25). Following a commercial decision, T25 maize has never been commercially cultivated in the EU, although the authorisation remains valid. Information on the PMEM which is implemented for Amflora and MON810 is provided in Annex 4.

Currently 18 applications for the commercial cultivation of GM crops in Europe are at different stages in the regulatory pipeline. These applications are for five types of crop: maize, potato, sugar beet, soybean and cotton. The crops are modified to be resistant to insect pests (maize), have modified starch content or resistance to late blight (potato) or to be herbicide tolerant (maize, sugar beet, soybean and cotton). At present there are no applications for cultivation of GM oilseed rape in the EU.

Not all of these crops would be of interest to UK farmers. In particular the insect resistant maize varieties do not offer significant benefits at present owing to absence or low incidence of the insect pest. Starch potatoes are also not currently cultivated in the UK owing to the lack of starch processing facilities. There is very limited cultivation of soybeans in the UK and no cultivation of cotton.

At present there are applications in the regulatory pipeline for cultivation of three types of GM crop which could be of interest to UK farmers; these are GM herbicide tolerant (GMHT) maize, GMHT sugar beet and late blight resistant potato.

## Section 3
## General recommendations on case specific monitoring and general surveillance

### 3.0 Introduction

This section of the report summarises the ACRE working group's advice on case specific monitoring and general surveillance. It considers situations where CSM would be required, describes the different objectives of CSM and GS and considers the different approaches required.

### 3.1 Case specific monitoring

*"to confirm that any assumption regarding the occurrence and impact of potential adverse effects of the GMO or its use in the environmental risk assessment are correct"* [3]

CSM is not required in all cases. It is needed in situations where, following the environmental risk assessment, a specific hypothesis remains as to how GM crops could cause adverse effects. This must include a pathway as to how harm could occur. It is important that situations which require CSM are clearly defined in the environmental risk assessment. CSM must be well designed to effectively test remaining hypotheses.

The design of CSM will depend on the hypothesis being tested. It is likely to require additional sampling or in specific (limited) cases may be most effectively addressed by additional targeted questions in the Farm Questionnaire. The design of CSM should be proportionate to the level of risk.

CSM may be needed due to uncertainty about whether an effect could occur. It may not be possible to fully resolve uncertainty in pre-market field trials if, for example, an effect would only become apparent over time or owing to the wide range of environmental and agricultural conditions across the EU. One reason for the difficulties in resolving uncertainty is the comparative risk assessment (of GM and non-GM crops) required by the GM legislation. This includes a comparative assessment of the cultivation and management practices.

Although it might be impossible to resolve experimentally, such uncertainty also indicates that aspects of cultivation of GM crops could be better for the environment. It would therefore not be proportionate to refuse authorisation on this basis. **CSM provides an appropriate route for the testing of specific hypothesis which can only be addressed under the range of conditions that are represented by commercial farming.**

When designing CSM, it is important to take into account the management strategies which will be put in place to mitigate any risks identified in the ERA. Evidence that risk management strategies will be effective should be provided in the ERA. If sufficient evidence is provided to demonstrate this, CSM will not be needed. If there is uncertainty as to whether management strategies will be effective, CSM may still be needed. In these cases CSM should be designed to test and resolve uncertainty about effectiveness of the management strategy. In all cases the implementation of risk management measures should be monitored. The most suitable tool for monitoring their implementation is likely to be the Farm Questionnaire.

CSM provides a way in which adverse effects can be detected at an early stage so that preventative action can be taken. It should be considered an additional safety measure put in place to mitigate risks by detecting any adverse effects at an early stage of commercial use so that action can be taken. **Identifying the need for CSM should not be seen as an indication that unacceptable risks remain and that authorisation should be refused.**

### 3.2 General Surveillance

*"to identify the occurrence of adverse effects of the GMO or its use on human health or the environment which were not anticipated in the environmental risk assessment"* [3]

GS is required to monitor for unanticipated adverse effects. There is no reason to expect that GM crops would have adverse effects if risks have not been identified in the environmental risk assessment. GS is, however, in line with the precautionary approach set out by the legislation. As there is no hypothesis as to how adverse effects could occur it is challenging to determine what should be monitored. With finite resources it is not possible to monitor all aspects of the environment and it will be necessary to focus monitoring to maximise the potential to detect adverse effects should they occur.

**GS should focus on environmental parameters in close contact with the GM crop.** As applications currently within the regulatory pipeline are for cultivation of GM arable crops, at present these will be parameters associated with arable farmland. These are the aspects of the environment which would have maximum exposure to the GM crop and where any adverse effects would be expected to become evident first. In the future other GM plants, such as trees, could be developed for commercial cultivation. In this case GS would need to be adapted to monitor environmental parameters in close contact with such GM plants.

**As GS is not hypothesis-driven the same approach could be taken for all arable GM crops.** Adopting the same approach for all crops would have the added advantage that GS would then be suitable for monitoring multiple GM crops and their interactions. It would not be appropriate to modify GS according to the introduced GM trait as this would be hypothesis-driven.

The EFSA guidance recommends the use of three main tools for GS: the Farm Questionnaire, the use of existing ESNs and a literature review. To be most effective**, GS should use a combination of tools, to maximise the chances of detecting any adverse effects.** Each tool has different strengths and weaknesses, but together they can be used to monitor a range of protection goals. Certain environmental parameters will be more effectively monitored by one tool than another.

Further consideration is given to the use of Farm Questionnaires and ESNs in the following sections of the report. Although the use of literature reviews is not discussed further, they provide an important tool as part of GS. Literature reviews should follow standard, methodological approaches according to clearly defined protocols and there is value in using a 'systematic review approach' where appropriate[8] [9]. A clear explanation of the relevance of papers which are identified to the commercial cultivation of GM crops should be included in the reports submitted by consent holders.

In some cases there may be overlap and the same environmental parameter may be monitored by more than one tool. If the same effect were detected by more than one tool, this would provide a stronger indication that this could be associated with the cultivation of a GM crop and could help to inform the type of subsequent investigation and action needed.

## Section 4
## Specific Recommendations on General Surveillance

### 4.1 Introduction

This section of the report discusses concepts for defining assessment endpoints to focus General Surveillance. Specific recommendations are provided on the use of the Farm Questionnaire and criteria are established for identifying suitable existing Environmental Surveillance Networks.

### 4.2 Protection Goals

*Protection goal: natural resources or natural resource services which are to be protected as set out by EU legislations*[6]

In the absence of a hypothesis as to how harm could occur, determining what should be monitored is challenging. Resources will be finite and GS needs a proportionate and pragmatic approach. EFSA suggest that in the absence of a hypothesis GS can be focussed by defining a set of valued protection goals[7].

The EFSA guidance provides examples of protection goals with corresponding assessment endpoints and measurement endpoints[7]. These terms originate from the field of ecological risk assessment[10]. Together they describe: the general concept of what is to be protected (protection goal); a specific definition of entity and attribute of the environment to be protected (assessment endpoint); and the environmental parameters or indicators which should be measured to determine whether protection is effective (measurement endpoints).

As suggested by EFSA, wider EU legislation or domestic policy could be used to define protection goals and assessment endpoints and to identify corresponding measurement endpoints monitored by existing surveillance networks. This approach is discussed further in Annex 5. An alternative effective way to define assessment endpoints and measurement endpoints would be to use an ecosystem function or ecosystems services approach. The working group recommends that an ecosystems approach is developed and used to **define an agreed set of assessment endpoints and corresponding measurement endpoints that can be used routinely to report on the status of the farmed environment and identify correlations with any possible influencing factors.** This would allow GM crops to be monitored in the context of wider impacts of agriculture on the environment. It would involve defining what to protect in the farmed environment in order to protect ecosystem services such as pollination, desirable biodiversity and soil quality. The development of protection goals and endpoints using an ecosystems services approach would be a major undertaking, which goes beyond the remit of what should be required under GM legislation, and has therefore not been attempted here.

As an alternative, the protection goals identified by EFSA[7] have been used in this report to identify corresponding assessment and measurement endpoints in the farmed environment (see Table 1). The assessment and measurement endpoints have been defined according to the activities of existing ESNs in the UK and to what can reasonably be achieved using the Farm Questionnaire.

### 4.3 Background information on the Farm Questionnaire

The Farm Questionnaire (FQ) has been developed as a tool for GS of GM crops for focused monitoring at the level of production, i.e. fields and farms[11,]. Parameters relating to protection goals are monitored. Recommendations on the design and analysis of the FQ are provided in the EFSA guidance[7]. A very basic description of the FQ is provided in this section of the report. More detailed information is available in the references provided.

FQs have been developed and used for monitoring GM crops currently cultivated in Europe[12] [13]. As a result of efforts made by industry to develop a harmonised protocol, the FQ plans for most crops in the regulatory pipeline follow a similar template. This FQ asks a series of structured questions aimed at determining whether there are differences between the GM crop and the conventional alternative and whether cultivation of the GM crop could have adversely affected certain protection goals. Information on influencing factors, such as pest pressure and soil characteristics, is also collected.

FQs ask questions based on a qualitative, comparative scale (i.e. more than, the same, less than). Farmers are asked to answer based on their experience of growing the GM crop relative to the conventional, non-GM crop grown on the same farm or their historical knowledge of growing the non-GM crop. An effect is recorded if above a certain threshold of farmers (usually 10%) report a deviation from “the same.”

### 4.4 ACRE working group recommendations on the Farm Questionnaire

**The FQ is a useful tool which can be employed as part of GS of GM crops**. FQs represent a proportionate way of monitoring for certain unanticipated effects which could result from the cultivation of GM crops. Although there are limitations as to what this tool can deliver, there are also certain advantages.

A key advantage of the FQ is that observations are carried out at the farm level by those working most closely with the crop i.e. farmers and/or agronomists. The FQ can be effectively used to collect information on agronomic parameters (e.g. incidence of pests, disease and weeds) which farmers observe closely. The FQ offers a relatively low cost and proportionate approach for detecting any direct or indirect effects of cultivating GM crops.

Basic information can also be collected on other parameters such as biodiversity. It would not, however, be reasonable to expect farmers to return detailed, species-level information on biodiversity. Environmental Surveillance Networks would be more effective for collecting this type of information.

**The potential of the FQ as an information source should be maximised wherever practical**. The FQ should be used to collect basic information on parameters such as biodiversity, despite its limitations. As described in Section 3.2 of this report, overlap in the parameters monitored by different tools can be advantageous. If a negative effect were recorded by both the FQ and an ESN this could provide a clearer indication that an effect correlated with GM cultivation may have occurred. If an effect were only recorded by one of the tools this should not, however, remove the need for further investigation.

A disadvantage of the FQ is that it does not collect quantitative data and the detection of an effect does not provide any indication of its size. It would not, however, be feasible to use this tool to collect quantitative data. The threshold (usually 10%) for recording an effect is arbitrarily defined, but is a reasonable trigger point for determining when further investigation would be needed. If an effect is recorded it is important that the regulatory authorities are informed immediately and consulted about the form further investigation should take.

GS is not hypothesis-driven and so ideally all FQs should comprise the same core set of questions. It may be relevant to create some additional questions based on crop type as different species, pests and diseases are closely associated with different crops. Questions included for purposes of GS will not need to be considered on a case-by-case basis as monitoring for unanticipated effects should not be affected by the trait.

In certain (limited) cases, the addition of specific questions to the FQ to support CSM could be advantageous. This would need to be determined on a case by case basis, designed to address a specific hypothesis and identified as case-specific questions. The FQ would be particularly useful for collecting information on the implementation of management measures put in place to mitigate risk, although it could not be used alone to determine the effectiveness of these measures. It could, for example, also be used to support the testing of certain hypotheses on the impacts of herbicide use with GM herbicide tolerant crops. It could not be used to collect detailed information on weed diversity, but could be used to collect information on the extent of weed problems, herbicide usage and timing of applications.

**The FQ should also collect information on cultivation and agronomic practices in fields in the years following GM crop cultivation**. This would enable the detection of any effects at later stages in the rotation and so maximise the FQ's effectiveness as a monitoring tool. While the main focus should remain on collecting information during the year in which the GM crop is cultivated, it may be possible to target a smaller subset of farmers in years following cultivation.

The FQ must be implemented in a way which ensures accurate and impartial information is recorded. Trained representatives from third party organisations have been employed to conduct the questionnaire for MON810 maize and support growers in understanding and accurately answering the questions[14]. This approach has advantages in promoting consistency, quality and return of responses. To further test the reliability of the FQ, findings could be independently verified with a subsample of farmers.

The FQ is a valuable tool, and its use to collate information from different sources should be encouraged. It would be advantageous for agronomists, as well as farmers to contribute to the FQ. This might enable more detailed information to be collected in particular given agronomists' expertise in identification, for example, of a wide range of weeds, diseases, pests, and predators and soil conditions. In addition agronomists would typically work across a number of different farms, which could help to validate unusual effects and determine whether or not they could be associated with GM-cropping.

### 4.5 Background information on Environmental Surveillance Networks

*"surveillance could, if appropriate, make use of already established routine surveillance practices"* [3]

The concept of using existing systems for surveillance of GM crops is introduced in Directive 2001/18/EC. This idea of making use of Environmental Surveillance Networks (ESNs), which are already in place in EU countries for other purposes, has been developed further in the EFSA guidance[7]. Although the design, implementation and reporting of monitoring remains the responsibility of applicants, it is acknowledged that authorities in EU countries should play a role in determining how ESNs can be used for GS. This report aims to contribute to this discussion.

ESNs have been described as an additional tool to strengthen independent monitoring[15]. ESNs are not currently used for GS, although efforts have been made to make use of them in Germany[16]. In the Netherlands, the availability and potential usefulness of ESNs for GS has been investigated[17]. This report presents the findings of a similar exercise in which the availability of ESNs in the UK and their capabilities and limitations have been investigated.

In the UK there is a well developed set of existing ESNs, used to monitor the status of the environment. Many of these have extensive coverage and long term data sets. At present, in the UK, ESNs are rarely used to investigate links between cause and effect; to achieve this, hypotheses are usually established and then tested experimentally. Many stressors have impacts on the environment and it is difficult to identify individual causes of change against background variability. GM crops would be just one of a range of potential drivers of change within the agricultural landscape. In addition, more recently, many initiatives have been put in place to mitigate the negative impacts of agriculture on the environment. It will therefore be challenging to determine whether any change reported by ESNs is linked to the cultivation of GM crops. These challenges are discussed further later in this report.

### 4.6 Criteria for selecting Environmental Surveillance Networks

Although a large number of ESNs exist in the UK, not all of these will be well suited for use in GS. ESNs suitable for GS of arable crops should ideally fulfil all of the following criteria:

a) Have monitoring points in arable farmland in both areas where the GM crop is grown and control sites where GM crops are not grown

b) Monitor parameters that are, directly or indirectly, sensitive to change in the farming system

c) Allow inclusion in the analysis of a range of influencing factors (covariates) which could affect the parameters monitored, either by collecting these data or by cross reference to other ESNs

d) Have temporal and spatial coverage that is appropriate to the environmental parameter being assessed. This will differ according to which environmental parameters are monitored

e) Use standardised protocols including well defined, scientifically robust sampling strategies and data verification and validation procedures

f) Collect data that enables numerical analysis

g) Undertake long term surveillance providing baseline data and continuing for at least the duration of the authorisation

It should be noted that there will be a limited number of networks which fulfil all of these criteria. In practice it may be necessary to select networks which meet as many of the criteria as closely as possible. Some networks will not prove useful for GS because whilst they collect data which is relevant to protection goals, they do not meet enough of the remaining criteria.

### 4.7 Identification of existing Environmental Surveillance Networks

Table 1 lists examples of existing ESNs in the UK mapped against the protection goals identified by EFSA. These are ESNs which monitor in, or close to, arable farmland and collect information on relevant parameters. Not all of these networks meet all of the above criteria and the list is not exhaustive. The value of these ESNs for GS will differ depending on the extent to which the criteria are met. In addition, the networks available do not cover all aspects of the arable environment. Gaps and limitations of existing networks are further considered in Annex 6 of this report.

In the UK there are existing initiatives to focus monitoring to meet the evidence needs of dealing with environmental change, to achieve value for money, and to make better use of monitoring data (Box 1). This work is independent of the regulatory requirements for GS of GM crops, but in the future could consider GM crops as one of a range of potential influencing factors.

In some cases, Table 1 records measurement endpoints, which differ from the assessment endpoint. For example the Countryside Survey collects detailed information on vegetation. This can be used to derive information on the measurement endpoint ‘abundance of key seed-bearing plant species for food provision in arable and horticultural fields.’ This measurement endpoint is related in the table to the assessment endpoint ‘farmland birds’. In addition to measuring farmland birds directly, it may also be possible to use other measurement endpoints to identify possible impacts on farmland birds. If a decline in the abundance of key seed-bearing plant species were seen, this could be an early indication of potential impacts on farmland birds, which depend on these plants for food.

**Table 1 ESNs in the UK which could be used as part of General Surveillance of GM crops.** Further information on these ESNs is provided in Boxes 2 to 5.

| **General Protection Goal** | **Specific Protection Goal** | | **Environmental surveillance network** |
|---|---|---|---|
| | **Assessment endpoints** | **Measurement endpoints** | |
| **Conservation of biodiversity** | Farmland birds | Population growth rates, population trends or abundance relative to other locations/habitats | Breeding Birds Survey |
| | Farmland birds | Abundance of key seed-bearing plant species for food provision in arable and horticultural fields | Countryside Survey |
| | Butterflies | Butterfly species abundance on fixed transects | UK Butterfly Monitoring Scheme |
| | Butterflies | Abundance of butterfly food plants (e.g. nectar plants) | Countryside Survey |
| | Plant diversity | Several (e.g. dicot species richness & cover) | Countryside Survey |
| | Bats | Population trends for 11 common resident species | National Bat Monitoring Programme |
| **Soil quality/ functionality** | Soil quality | Physical: bulk density; Chemical: loss on ignition, pH, % carbon, nutrients; Biological: invertebrate composition, abundance, diversity | Countryside survey; Land Information System; OPAL (Open Air Laboratory) earthworm census |
| **Water** | Water quality | Chemical: concentration of a range of substances including nutrients and pesticides<br>Biological: Macro-invertebrate, diatom, macrophyte and fish taxanomic composition, abundance, diversity and sensitive taxa. | Environment Agency monitoring programme (plus Water Company monitorng programmes for Drinking Water Protected Areas) |
| **Sustainability of agro-ecosystems, including plant health** | Pollination | Abundance of butterfly food plants in the wider countryside | Countryside Survey |
| | Permanent pasture | No decline in permanent pasture (ha) | June Survey |
| | Pest and disease outbreaks | Pest and disease incidence in winter wheat, spring beans and oilseed rape. | Crop Monitor |
| | Pesticide use | Area treated / amount applied / amount applied (ha) / # times treated | Pesticides Usage Survey |
| | Fertiliser use | Average application rates of nitrogen, phosphate, potash, sulphur, organic manures & lime on agricultural crops | British Survey of Fertiliser Practice |

**Box 1 Focusing monitoring in the UK**

A major existing initiative to focus monitoring is the UK Environmental Observation Framework (UK EOF). Launched in 2008, the UK EOF provides a catalyst for the stakeholders in environmental monitoring to review and adjust monitoring so that it helps provide the evidence needed to tackle the societal challenges associated with environmental change. The UK EOF is funded by the major sponsors of observations in the UK and is a self contained programme of Living With Environmental Change (LWEC), a partnership of government department's agencies, local government and research councils. UK EOF provides tools to help stakeholders balance the regulatory and scientific requirements for monitoring with the need to detect and predict environmental change. These include providing a comprehensive overview of monitoring programmes and activities, criteria and evaluation frameworks and support to initiatives that improve the accessibility of monitoring data and their results.

Many monitoring responsibilities are held by countries within the UK, and co-ordination at this level draws on UK EOF resources. In England the Defra network is undertaking a review of monitoring whilst in Scotland an environmental monitoring strategy is being developed under the CAMERAS (Co-ordinated Agenda for Marine, Environment and Rural Affairs Science) umbrella. Similar co-ordination of monitoring is developing under the Welsh Natural Environment Framework and within Northern Ireland. These broader scale co-ordination initiatives provide the context for thematic and more specific reviews of monitoring where the need to be able to detect drivers of change in the agricultural landscape can be considered in more detail. These include an LWEC initiative to drive innovation in measuring change in the countryside, a review of the Countryside Survey monitoring programme, and the UK Terrestrial Biodiversity Surveillance Strategy which provides co-ordination across UK and the devolved administration's biodiversity monitoring responsibilities and activities.

### 4.8 Information on four ESNs chosen as case studies

Four networks have been selected and used as case studies in the following section of this report; the Countryside Survey (CS), the Breeding Bird Survey (BBS), the UK Butterfly Monitoring Survey (UKBMS) and the Environment Agency Water Quality Monitoring Programme (WQMP). These four networks meet the majority of the criteria described in Section 4.6 of this report. These ESNs were chosen as case studies as they have long term data sets from a large number of monitoring points. They also monitor protection goals which are less effectively monitored by the Farm Questionnaire and could therefore benefit from monitoring using an alternative means. Further information on these networks is provided in Boxes 2 to 5.

### Box 2 UK Butterfly Monitoring Scheme (UKBMS)

**Website:** http://www.ukbms.org/

**Aims:** The UKBMS provides an annual estimate of the relative abundance of butterfly species at sites. Site indices are combined to derive regional and national collated indices and are used to estimate trends over time. The scheme is important, not only as the provider of information on butterfly population trends and status, but also as it enables the UK to meet its obligations under the European Habitats Directive (92/43/EEC) (for Marsh Fritillary and Large Blue), and to report on, implement or deliver UK Biodiversity Action Plan (UK BAP), country biodiversity strategies and biodiversity indicators. The data are also used to support various research initiatives and partnerships to help better understand butterfly populations and to start to understand factors that may be affecting populations.

**Protection goal:** Biodiversity
**Parameters measured:** Butterfly species abundance

**Established:** 1976. In 2009 the Wider Countryside Butterfly Survey (WCBS) was incorporated into the UKBMS. Further information on the WCBS is provided in Annex 7.
**Led by:** Butterfly Conservation (BC) and the Centre for Ecology & Hydrology (CEH)
**Funded by:** JNCC, Defra, NE, FC, CCW and SNH.

**Data collected by:** Volunteers
**Number of sites:** Initially less than 50, now approximately 1800 per year
**Number of sites in arable farmland:** ~30
**Sampling area:** Fixed line transects of width of 5m and length of approximately 3km
**Selection of sites:** Sites are selected by the observers (volunteers), with guidance from BC and CEH. There is a predominance of sites hosting localised habitat specialist species and an under-representation of arable land, which tends to attract largely ‘wider-countryside’ habitat generalist species. There is considerable site turnover, although a substantial proportion of sites have time series of >10 years.
**Frequency of sampling:** Annual. Up to 26 weekly counts are made, subject to suitable weather conditions, each year.
**Outline of protocol:** Standardised protocols. Observers walk the length of the fixed transect, making counts of butterfly species encountered within a specified distance from the observer (typically 5m). For the majority of transects all species of butterfly are recorded, but there are also a small number of single species transects.

**Analysis of data:** Data are collated each year and undergo a series of quality assurance checks before analysis. Analysis provides species trends which are aggregated to form indicators e.g. the UK Biodiversity indicator and country indicators. Data can also have local applications and can be used in analyses of pressures. For example to measure the impacts indices of annual changes in local, regional and national abundance, which are used to measure the status of species (e.g. for UK BAP reporting), multispecies community indicators and to measure the impacts of environmental stressors, for example, climatic change. The principle results of the scheme are annual estimates of population abundance for each combination of site, year and species. These data are used to calculate a national (or regional) collated index of abundance by fitting a log-linear Poisson regression model to account for site and year effects, and to account for missing values. A linear regression model is fitted to the collated index to measure the trend over time.
**Data availability and reporting:** All data are stored in a centralised relational database. The results of the scheme are published via the UKBMS website (www.ukbms.org.uk) the spring following data collection and later in an annual report to contributors and funders. Data are available on request via the UKBMS website for use subject to terms of a licence agreement.

**Box 3 Breeding Bird Survey (BBS)**
**Website:** http://www.bto.org/volunteer-surveys/bbs

**Aims:** To monitor population changes among the UK's common breeding birds. BBS data also underpin the bird indices used in the UK Biodiversity Indicators, the England Biodiversity Strategy indicators and the Scotland Biodiversity Indicators to inform about the sustainability of policy and management nationally. BBS data from the UK are a component of Pan-European bird indicators produced annually for the European Commission. The EU farmland bird indicator is an EU Structural Indicator and also an Indicator of Sustainable Development of the EU. The UK farmland bird indicator has been also approved as the indicator for Regulation in EU´s Rural Development Plans (Council Regulation (EC) No 1698/2005). These indicators are also used by the Organisation for Economic Co-operation and Development (OECD), United Nations Environment Programme (UNEP), European Environment Agency (EEA) and are included in the Living Planet Index (LPI).

**Protection goal:** Biodiversity
**Parameters measured:** Abundance of breeding birds

**Established:** 1994
**Led by:** British Trust for Ornithology (BTO)
**Funded by:** Jointly funded by the BTO, the Joint Nature Conservation Committee (JNCC) (on behalf of the Countryside Council for Wales, Natural England, Scottish Natural Heritage and Northern Ireland's Council for Nature Conservation and the Countryside) and the Royal Society for the Protection for Birds (RSPB), on the basis of a six-year rolling contract.

**Data collected by:** volunteers
**Number of sites:** approximately 3200 per year
**Number of sites in arable farmland:** 950 (+ 700 in mixed farmland)
**Sampling area:** 2km fixed transect within a 1km$^2$ area
**Selection of sites:** chosen as a random sample (stratified by observer density) by the BTO. Interested volunteer observers are assigned a site near to where they live; target levels of coverage are currently exceeded for lowland Britain.
**Frequency of sampling:** Annual. Two visits each year.
**Outline of protocol:** Standardised protocol. The transect is walked in the early morning at two different time points in the year (April-mid-May and mid-May-June). Birds are recorded in three distance categories, or as 'in flight.' Habitat information is recorded on a separate, third visit.

**Analysis of data:** Data undergo quality assurance before analysis. Where possible population trends are calculated at the national, country and regional levels. Population changes are estimated using a log-linear model with Poisson error terms.
**Data availability and reporting:** Annual reports: showing smoothed and unsmoothed trends and recording significant changes at the species level and considering composite indicators across broad habitats and countries or regions, approximately one year after data collection. Population changes are used to inform the UK Red and Amber lists of Birds of Conservation Concern, which are updated every six years. In addition, an alert system designed to draw attention to developing population declines that may be of conservation concern is applied annually, with alerts recorded given a >25% or >50% population decline over a 5-year (or longer) time period. Alerts are advisory and do not supersede the agreed UK conservation listings. At the UK level alerts were triggered for ten species between 2004 and 2009. http://www.bto.org/about-birds/birdtrends/2011.

**Box 4 Countryside Survey (CS)**
**Website:** http://www.countrysidesurvey.org.uk/

**Aims:** CS data are used for quantifying and understanding the processes of change in the countryside. This has enabled the detection of gradual and subtle changes in the UK's countryside over time and their connection with changes in farm-management practices associated with set-aside, grazing regimes and fertiliser use. The use of data from CS to look at the impacts of recent agri-environment schemes is also currently being explored. CS data is used to report on extent and condition of UK Broad habitats, including some priority habitats, water and soil quality, informing a range of EU directives (further information is provided in Annex 5b).

**Protection goals:** Biodiversity, soil quality/functionality, water, sustainability of agro-ecosystems
**Parameters measured:** A comprehensive assessment of the natural resources in the UK's countryside habitats, landscape features, soil chemistry, soil invertebrates (to broad taxonomic groups), freshwater macro-invertebrates, hydrochemistry and plant communities of ponds and headwater streams, vegetation (species and cover)

**Established:** 1978 (most recent survey 2007)
**Led by:** Centre for Ecology & Hydrology (CEH)
**Funded by:** a partnership of research councils, government departments and agencies

**Data collected by:** professionals
**Number of sites:** Approximately 600 squares in 2007
**Number of sites in arable farmland:** In 2007, a total of 86 plots in maize and potato fields specifically and a total of approximately 450 plots in arable land generally.
**Sampling area:** 1km$^2$
**Selection of sites:** Stratified random sample. Fixed sites revisited.

**Frequency of sampling:** Approximately every eight years
**Outline of protocol:** Relevant data collected by CS for PMEM comes from fixed location vegetation plots located in farmland environments, particularly those that are in or adjacent to arable fields. These record crop and non-crop species, including arable weeds at the time of survey (June – October). Information on cropping and land management in years between surveys is not directly available from the survey but can sometimes be obtained from other sources.

**Analysis of data:** Data undergo quality assurance before analysis. Estimates of stock and change are calculated for multiple metrics collected from the field survey using a generalised linear mixed model framework. This allows for the nested nature of plots within squares and the uneven sample size over time. Poisson, gamma, normal and binomial error distributions are used dependent on the metric of interest.
**Data availability and reporting:** The data and results from countryside survey are released in a series of reports covering the UK nationally and its regions and most are available through the website at http://www.countrysidesurvey.org.uk/data-access. Reporting takes place the year following the main field season. With the exception of information on the location of squares, data are publicly available via the CS website, for use subject to terms of a licence agreement.

**Box 5 Environment Agency water quality monitoring programme (WQMP)**
**Website:** http://www.environment-agency.gov.uk/research/planning/34383.aspx

**Aims**: The Environment Agency monitors water quality in order to determine status; direct action; and measure improvement, in support of the wider goal to 'improve and protect inland and coastal waters'. There is a wide range of national and international reporting requirements, including those associated with the EC Water Framework Directive (http://www.environment-agency.gov.uk/research/planning/33362.aspx).

**Protection goal:** Water quality
**Parameters measured:** A range of parameters is monitored, including biology (taxanomic composition, abundance, diversity and sensitive taxa for phytoplankton, diatoms, macrophytes, macroinvertebrates and fish); physico-chemical (oxygen, temperature, pH, nutrients); priority substances and specific pollutants.

**Established:** Most sites possess long term (5 years or more) records with some records dating back to the 1970s.
**Led by:** The Environment Agency, an Executive Non-departmental Public Body, responsible for delivering the environmental priorities of central government and the Welsh Government
**Funded by:** Central government

**Data collected by:** Professionals (trained Environment Agency members of staff)
**Number of sites:** Currently around 5,000 chemical and 1,900 biological river monitoring sites in England, with a more limited network of 700 chemical and 600 biological sites where long-term monitoring is in place (these include *Surveillance* sites for the EC Water Framework Directive; the Environmental Change Network and Harmonised Monitoring Scheme)
**Number of sites in arable farmland:** Approximately 15 per cent of freshwater monitoring sites are directly adjacent to arable farmland.
**Sampling area:** Sample points or stretches in rivers, lakes, transitional, coastal, and ground waters.
**Selection of sites:** Sites are selected to be representative of water bodies. Regular reviews of monitoring requirements have resulted in loss and additions of specific sites.
**Frequency of sampling:** Water chemistry is typically monitored at a monthly frequency and most long-term monitoring sites possess at least five year data records. Biological monitoring is typically undertaken 1-3 times per annum on a 3-year rolling cycle.
**Outline of protocol:** Standardised protocols specific to the measured parameters

**Analysis of data:** Data are subject to analytical and statistical quality assurance checks. Data are analysed to determine compliance assessment and quality trends.
**Data availability and reporting:** Data are held on the Environment Agency public registers and are available on request. The outcomes of analysis are made available on the Environment Agency website:
Quality trends:
http://www.environment-agency.gov.uk/research/library/data/58818.aspx
Water Framework Directive Classification:
http://www.environment-agency.gov.uk/research/library/data/97343.aspx
What's in your backyard?:
http://www.environment-agency.gov.uk/homeandleisure/37793.aspx

## Section 5
## The data collected by ESN and options for analysis

### 5.0 Introduction

The types of analyses which are possible depend on the nature of the data collected by ESN. This section of the report provides a description of the data collected by the four ESN detailed above and discusses options for analysing them.

### 5.1 The data collected by ESN

Two of the ESN described above (the UKBMS and BBS) record species-level counts during the course of each year. Both surveys use transects, which in arable areas may cross a number of fields. Birds are much more mobile in the landscape than butterflies, ranging over a wider area, so fields or habitats immediately adjacent to the transect would therefore have more of an influence on butterfly than on bird counts.

In contrast, the CS takes place only approximately every eight years. Data are recorded in a number of formats including counts of species number or percentage cover of vegetation. A number of plots are located within a 1km$^2$. The precise location of these plots is known. Given the sessile nature of vegetation the greatest influence will be the management of the field in which the plot is located.

Water quality is recorded as measurements of chemical concentrations and ecological abundance, diversity and taxonomic composition. Water quality is monitored at specific locations. The most useful sampling sites for GS would be streams and rivers adjacent to arable land. Sampling sites would be influenced not only by the adjacent fields, but by the management of the whole upstream water catchment area.

Although all of these networks monitor a large number of sites, only a subset of these sites occurs in arable farmland. A still smaller subset of sites are associated with fields where maize, sugar beet or potato are cultivated. Analysing the data from sites in arable farmland would be most informative for determining whether cultivation of GM crops resulted in unanticipated adverse effects.

As illustrated by Figure 1 there is typically a large amount of between year and between site variation in the abundance, or levels, of the environmental parameters monitored by these networks. There may also be existing trends of either deterioration or improvement, which add further complexity. The plots in Figure 1 are based upon UKBMS butterfly data for two generalist species based on counts recorded at arable sites (large white and small tortoiseshell). Significant changes must be identified against this background variability. Changes may be caused by a number of different drivers and it will be difficult to separate these from any effects caused by the cultivation of GM crops.

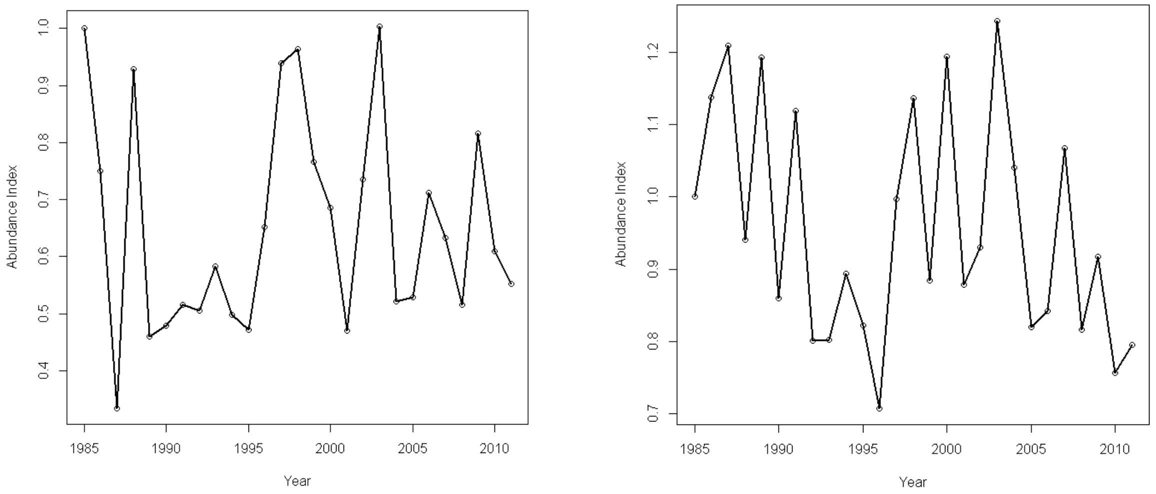


Fig 1a

Fig 1b

**Figure 1.** Indices of relative abundance for two common species recorded under the UK Butterfly Monitoring Scheme: a) the small tortoiseshell and; b) the large white. All indices are scaled to unity in the first year.

In the UK individual ESNs carry out the collation, analysis and publication of their data as detailed in Boxes 2 to 5. Although trends are monitored, data are not routinely used to investigate the relationship between cause and effect. In cases where the data are used to answer specific questions (e.g. see Box 6, Section 6.2) additional funding may be sought to enable ESNs to undertake this analysis.

## 5.2 Options for analysing data collected by ESN

### 5.2.1 Temporal analysis

Most ESNs in the UK report on trends, or changes, over time. If changes exceed an expected magnitude, this may trigger an alert. For example, alerts are raised if analysis of the BBS data reveals a greater than 25% decrease in abundance of a species over a five year period, or over longer periods of time[18]. For the period 2004 to 2009 alerts were raised for ten species at the UK level[19]. This illustrates a difficulty of using this kind of analysis for GS. If a 'red flag' were raised every time an alert was triggered there would be many false positives (i.e. situations where an effect is observed which is not associated with GM cultivation) as there are many drivers of change. As there is no way of predicting when an unanticipated adverse effect might occur, any significant effect which occurred following the introduction of GM crops would need to be investigated further.

In the Netherlands, it has been proposed that trend analysis could be used to interrogate the data collected by ESNs and to identify whether a change in a general trend occurs following the introduction of GM crops[17]. Similar analysis has been undertaken to identify change points in trends for breeding bird data in the UK[20] [21]. This is helpful where a species is in decline prior to the introduction of GM crops or if a species is increasing, but the increase slows. It determines whether the trend of decline continues at the same rate or is affected by the introduction of GM crops. It looks for a change in the trend which corresponds to the time point of the introduction of GM crops.

Analysis which looked for a change in trend would eliminate some false positives. For example a decline of greater than 25% which represented a continuation of a previous trend might be discounted. Any significant change in trend occurring after the introduction of GM crops would still require further investigation. This would be the case even if it occurred a number of years after the first cultivation if there was a plausible mechanism by which there could be a delay between cause and effect.

### 5.2.2 Comparison of two regimes (spatial or spatial plus temporal analysis)

It is also possible to analyse the data collected by ESN by comparing the effects of two different regimes i.e. comparing the data collected at sample sites in areas of GM cultivation with those in areas where GM crops are not cultivated. The comparison could be undertaken at a single snapshot in time or may compare trends over time under two different regimes. The latter approach, spatial plus temporal analysis, would be the most effective in identifying a change which was actually correlated to the introduction of GM crops.

In the Netherlands it has been suggested that this could involve a regional level comparison, comparing trends in areas where GM crops are grown with areas where there are no GM crops [17]. If the same trend were observed in both cases, it is less likely that this would be caused by GM crops. This spatial component provides a useful addition to the temporal analysis described above, but comparisons at a regional level will have limited replication.

If more detailed information were available on the location of cultivation of GM crops, a spatial analysis at the one kilometre square or at the field level would be possible. Trends at sample sites in fields or one kilometre squares where GM crops were cultivated could be compared with those in fields or areas where there was no cultivation of GM crops. Alternatively, when considering areas larger than a single field, the proportion of the area covered by a GM crop may be considered as a continuous explanatory variable, as this allows a more powerful analysis. If a statistically significant difference were seen between the two regimes, this would indicate a correlation with the cultivation of GM crops. It would still not demonstrate a link between cause and effect, but would identify a need for further investigation.

Spatial plus temporal analysis represents the most powerful option for analysing data from ESNs. Where possible, this is used in the following sections of the report to analyse the power of ESNs to detect change. In some cases, however, the design of the survey means that spatial plus temporal analysis is not possible and alternatives must be considered.

## Section 6

## Analysis of the statistical power of ESNs to detect change

### 6.0 Introduction

This section of the report provides an analysis of the statistical power of the four ESNs to detect change. It describes the power analysis methodology, results and the implications of the findings. The power analysis presented here uses a set of measurement endpoints, or indicators, selected as representative of the farmed environment. Any set of measurement endpoints could equally be used.

Data on the current cultivation patterns of three crops (maize, potato and sugar beet) is used in parts of the analysis in this section of the report. This is because, as described in Section 2.2, applications for cultivation of GM varieties of these crops are currently in the EU regulatory pipeline and, if approved, UK farmers may decide to cultivate these GM varieties commercially in the future

**Power analysis is a useful tool for determining the probability of detecting change using ESNs and therefore for interpreting a finding of no significant effect.** It provides an estimation of the probability of being able to detect a difference between two data sets if it occurs. If an ESN records no effect of a driver on a measurement endpoint it could be that there is no influence of that driver or that the driver does influence the endpoint but the analysis has failed to detect this. A power analysis is needed to distinguish between these two possible conclusions and to provide an estimate of the size of change ESNs are capable of detecting.

The size of change it would be desirable to detect can be determined to an extent based on expert opinion. This is informed by knowledge of the size of change which would be biologically relevant based on the extent to which a resource or species can tolerate and recover from negative impacts. Limits set in legislation may also dictate the magnitude of change it is desirable to detect (e.g. for compliance with the EC Water Framework Directive). Given the limits of current knowledge it will not always be possible to provide decision makers with a scientific recommendation on the size of change it would be desirable to detect.

Factors which affect power include sample size, variance in the abundance or levels of the measurement endpoint being monitored and the size of the difference in the datasets being compared. A value of ≥ 80 % (0.8 probability) is usually considered to represent a reasonable value for power. The Farm Scale Evaluations were designed to have the statistical power to detect a 50% change with 0.8 probability[22].

### 6.1 Generic power analysis for ESNs recording species counts

A model has previously been developed (the Freeman and Newson model) which can be used to determine whether there are differences in the annual growth rates of populations at sites subject to different treatments[23]. This combines a spatial and temporal approach. Using this model it is possible to test the null hypothesis of no difference or quantify the magnitude a difference between two sites under different regimes (e.g. a GM crop and a conventional crop). This model is valid for any survey recording counts of species on an annual basis. To illustrate how this model would apply to the UKBMS and BBS, the power to detect change of bird and butterfly species of known abundance is indicated on the graphs presented below. Further details of the model are provided in Hails *et al.* 2013[24].

In this report, the power of this model to detect such differences was estimated assuming a range of different circumstances. These circumstances were defined by the background population trend (on ‘untreated’ sites) and the value of eight more key factors: (i) the total number of sites; (ii) the proportion of those sites ‘treated’ (i.e. GM); (iii) the average abundance at a site at the start of the period; (iv) the variance of these initial abundances; (v) the proportion of site visits missed; (vi) the duration of the survey since the introduction of the treatment (e.g. GM crop) (years); (vii) the magnitude of the difference it is desired to detect and (viii) a measure of the overdispersion (unexplained variance in excess of that predicted by the Poisson distribution) in the data with respect to the Poisson model fitted. This analysis can be used for any measurement endpoint in any survey recording counts of species on an annual basis if values for the factors are known.

The information inputted to this model was simulated and was selected to cover a realistic range of values for the two surveys. Artificial data were randomly simulated for each of a given set of values for the above variables; in these artificial data sets differences between treated and control sites are deliberately engineered into the data. These data were then analysed using the Freeman and Newson model in an attempt to detect differences between treated and control sites. The proportion of occasions in which the differences between treatments were detected was taken as an estimate of the power of the analysis. A simple model was then derived to allow power to be predicted for any specified set of the nine factors described above. Further information on the methodology used for the power analysis is provided in Hails *et al.* 2013[24].

### 6.1.1 Starting abundance and the power to detect change

Figure 2 shows an example of how the mean starting abundance of the measurement endpoint (e.g. bird or butterfly species) influences the power to detect change. The effect size is set as a 1% per year decrease (over ten years) at sites where GM crops are grown. It also assumes a set number of sample sites available in arable farmland (n=50); other parameters are set as defined in the figure legend. The power to detect change is low for species with a low starting abundance (i.e. locally rare species) and increases for species with a higher starting abundance (i.e. locally more common species).

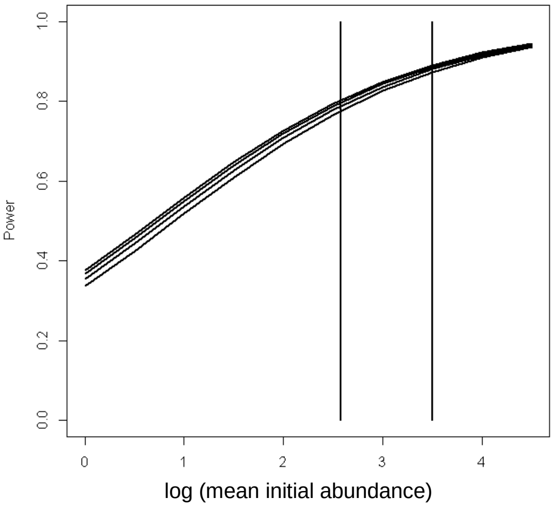


**Figure 2. Estimated power as a function of initial abundance.**

Other parameter values are: 50 sites visited over 10 years, with 40% of visits missed at random. Variance of initial log-abundances = 4 and the overdispersion parameter = 5. Power curves, top to bottom, correspond to 50%, 40%, 30% and 20% of all sites treated, and represent the power to detect a value of α = -0.01 compared to a stable trend at control site (i.e. a decline on treated sites of approximately 1% p.a.) .
The vertical lines represent the average abundance on arable UKBMS sites of large white (left) and small tortoiseshell (right) by way of example. These species are common and found in most habitats, although the small tortoiseshell is declining rapidly.

### 6.1.2 GM uptake and the power to detect change

The number of sites which are exposed to the treatment (i.e. GM cultivation) will depend, in this case, on the uptake of the crop by farmers. The four power curves in Figure 2 show how power increases as the proportion of sites where GM crops are cultivated approaches 50%. Figure 3 shows an alternative representation of how the GM uptake influences the power to detect change. This shows how power to detect change is greatest if approximately 50% of sites are cultivated with GM crops. If levels of GM uptake exceed 50%, the power is lower again as fewer untreated control sites are available for comparison.

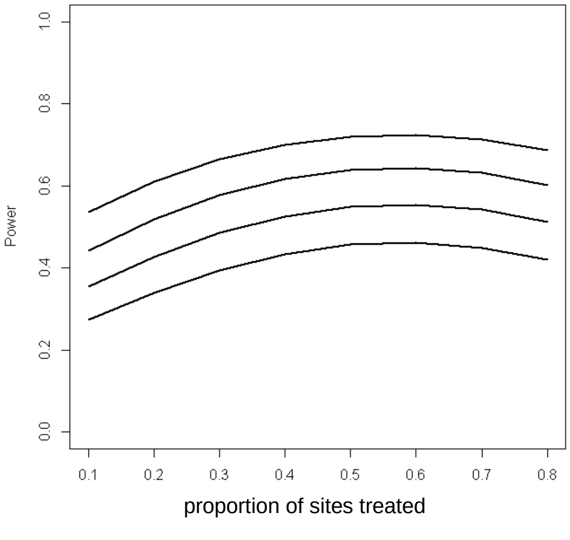


**Figure 3. Estimated power as a function of the proportion of 100 surveyed sites treated.**

Other parameter values are: 100 sites visited over 10 years, initial log (mean abundance) = 1.5, 1, 0.5, 0 from top to bottom (with variance = 4.0 and overdispersion parameter = 5) and 40% of visits missed at random. Power curves represent the power to detect a value of α = -0.01 compared to a stable trend at control site (i.e. a decline on treated sites of approximately 1% p.a.).

### 6.1.3 The number of sample sites and the power to detect change

Figure 4 shows how the number of sample sites affects the power to detect change. The effect size is again set as 1% per year decrease at GM sites with respect to non-GM sites. It also assumes cultivation of GM crops occurs at 20% of the sites. The initial abundance is set to match that of a moderately abundant species. For a species with a lower starting abundance, a larger number of sample sites would be needed to achieve the same probability of detecting change. Using these parameters it is notable that a large number of sites (> 130) are needed before the power to detect change reaches 80%.

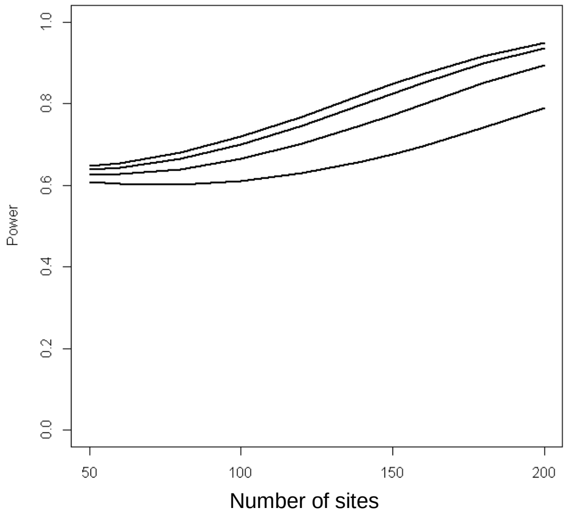


**Figure 4. Estimated power as a function of sample size (number of sites surveyed).**

Other parameter values are: Sites visited over 10 years, initial log (mean abundance) = 1.5 (with variance=4.0 and overdispersion parameter = 5) and 40% of visits missed at random. Power curves, top to bottom, correspond to 50%, 40%, 30% and 20% of all sites treated, and represent the power to detect a value of α = -0.01 compared to a stable trend at control site (i.e. a decline on treated sites of approximately 1% p.a.).

A key factor which is not considered explicitly in the generic power analysis is crop type. In reality, the number of sample sites which occur in areas where GM crops are being cultivated will depend on the types of GM crops which are available as well as the levels of uptake. There are only three types of GM crop currently in the regulatory pipeline which may be of interest to UK farmers for cultivation: maize, sugar beet and potato. Only a subset of sample sites in arable farmland will contain each crop type, and a smaller subset will overlap with any one indicator species over a number of years. In a spatially explicit analysis these factors will reduce the number of sample sites which are relevant.

### 6.1.4 Effect size and the power to detect change

Figure 5 shows how the power to detect change is influenced by the size of effect which occurs. It assumes a set number of sample sites available in arable farmland (n= 100) and that cultivation of GM crops occurs at 20% of the sites. The power to detect change is highest where large effects occur. The graph shows that for a species with a moderate starting abundance (mean abundance) of 1.5, the power to detect change becomes high (>80%) if an annual decline in abundance of over 2% occurs consistently for 10 years.

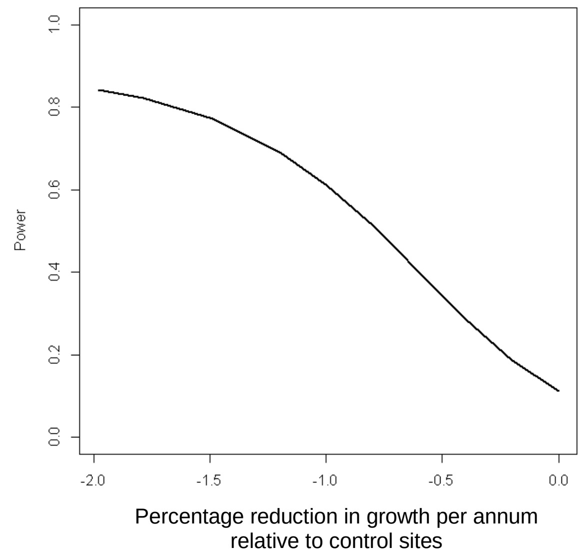


**Figure 5. Estimated power as a function of the strength of treatment effect.**

Other parameter values are: 100 sites visited over 10 years with the GM crop cultivated at 20% of those sites, initial log (mean abundance) = 1.5 (with variance=4.0 and overdispersion parameter = 5) and 40% of visits missed at random. The power curve represents the probability of detecting a difference at GM sites compared to a stable trend at control sites as the size of that difference varies. The x axis is the % reduction in growth relative to control sites *per annum* (varying from 0 to 2% *per annum* reduction).

### 6.1.5 Time lags and the power to detect change

All the graphs above have illustrated the power to detect change after 10 years of year on year change at the stated magnitude. However, the time lag between introduction of the 'treatment' and the analysis of survey data will also influence the power to detect change as illustrated in Figure 6. It is notable, that using these parameters it would take nearly 15 years before there would be power of ~80% to detect change of 1% *per annum*.

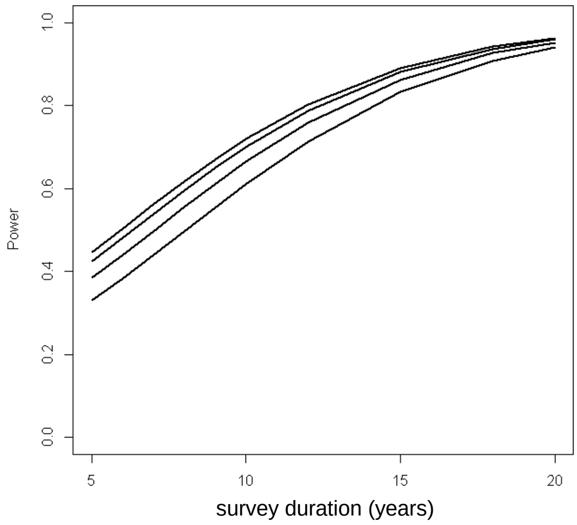


**Figure 6. Estimated power as a function of survey duration.**

Other parameter values are: 100 sites visited, with α = -0.01, initial log (mean abundance) = 1.5 (with variance = 4.0 and overdispersion parameter = 5) and 40% of visits missed at random. Power curves, top to bottom, correspond to 50%, 40%, 30% and 20% of all sites treated, and represent the power to detect the given value of α compared to a stable trend at control site.

### 6.1.6 Generic analysis: key findings

This analysis illustrates a number of key points which apply to all networks. Firstly, power to detect change depends on the starting abundance of the species and the size of change which occurs. It will be possible to detect large changes in the population growth rates of relatively common and widespread species using these networks. It will not be possible to detect small changes or changes in the population growth rates of rare or localised species.

Power to detect change also depends on the level of GM uptake. A maximum power to detect change is reached if GM crops are cultivated at around 50% of the sites. This is because in order to detect significant differences between two distinct groups sample size needs to be high in both. Detecting change can take several years. If a small change occurs in a population each year, it could take at least ten years before the cumulative effects become apparent.

The power to detect change always increases if a greater number of sample sites are available. The number of sample sites which are needed to reach a high power to detect change depends on the distribution and abundance of the species and the size of change which we wish to detect. The type of GM crop and levels of GM uptake will affect how many of the total number of sample sites actually contribute data to the analysis. In all cases the power to detect change could be increased by adding extra sample sites. For some species this would, however, require a very large increase in sampling effort.

**Box 6 The impacts of stubble management on population growth rates of farmland birds: analysis using the BBS data**

BBS data have been specifically analysed for the dependence of farmland bird abundance on stubble management under the Environmental Stewardship (ES) scheme in England from 2005 to 2010. This analysis is an example of where data collected by an ESN have been successfully used to detect a likely relationship between cause and effect. It should be noted that this was based on a clear hypothesis as to where a relationship would be expected. In this case a number of positive effects of management actions on populations were seen.

The impacts of stubble management on the population growth rates of three farmland bird species were revealed as part of a wider analysis of the possible effects of ES management likely to affect individual farmland bird species (i.e. each species was tested with respect to the management variables that its ecology indicated it might respond to *a priori*). Among other relationships, it was found that the growth rates of breeding skylark, linnet and yellowhammer populations were positively associated with areas of ES cereal stubble management options providing cereal stubbles in the local area over winter (Baker *et al.* 2012[27]). The analysis detected small, but statistically significant, differences in population growth rates.

### 6.2 Case study using BBS data

Section 6.1 provides a theoretical analysis of the capabilities of the UKBMS and BBS to detect change. This section of the report presents the outputs of an analysis using real data sets from the BBS. These data sets had been used in a previous analysis described in Box 6.

The Environmental Stewardship data set has been used here as a proxy to investigate the power of the BBS to detect small changes in population growth rate if they occurred as a result of the cultivation of GM crops. Known relationships between areas of ES stubble management options and the population growth rates of three farmland bird species, *Carduelis cannabina (*linnet), *Alauda arvensis (*skylark) and *Emberiza citrinella (*yellowhammer), between 2002 and 2010 were used. Note that these known relationships were positive, whereas effects of GM crops of interest here would be negative; however, inference about power to detect effects of land-use change remains valid and the simulated change relates to a subtle shift in crop management that is arguably analogous to a change from a conventional to a GM crop variety. The known relationships between stubble management area and average annual population growth rates (expressed as ratio modifiers) varied from 1.04 to 1.19 (i.e. 4-19% increases).

The analysis used the same population growth rate model framework described in section 6.1 of this report. Full details are provided in Hails *et al.* 2013[24]. Crop type was taken into account. “GM crop” areas were simulated (by re-sampling BBS data) to match the regional distributions of maize, beet and potatoes in order to approximate realistic bird data sets for the geographic distribution of each crop. Periods of three and six years since the inception of the cropping change (ES) were considered. This was done using data from the whole of England and dividing the data set into arable, pastoral and mixed farmland. Analyses were based on resampling within the actual areas of the different crop types found in different Government Office Regions to maintain realism in simulated distributions of GM crops.

Table 2 summarizes the power of the analysis to detect changes of the same size as those seen from stubble management in ES, with rates of GM uptake simulated between 20 and 80% of the total areas of the three different crop types and considering the population growth rates of the three species. Results are shown for the whole of England and for a period of six years following the change in management, with statistical power estimates for the crops providing the least and most power for each species (as a brief summary). Complete results for each crop and species at both the all-England level and in areas with predominantly arable, pastoral or mixed land-use separately, are provided in Annex 8. Annex 8 also presents results for both three and six years following the change in management, which show the expected pattern of greater power given a longer period of potential effect.

Power to detect effects varied between species and crop types but increased with the percentage conversion of crop to GM (Table 2), and was generally greatest when considering the whole of England or arable landscapes alone, although power was also high for some species-crop combinations in pastoral or mixed landscapes alone (see Annex 8). High power in arable-dominated landscapes reflects the distributions of these birds, which are biased towards arable farmland.

Specifically, for linnet, power was low for the data representing maize cropping and, with the other crops, lower in pastoral and, especially, mixed landscapes. For yellowhammer, power was generally low in mixed landscapes, but high for all crops in the national analyses, in arable landscapes and, for maize, in pastoral ones. Power for skylark was lower everywhere, especially for maize and for potatoes in each individual landscape type, but was high for beet in arable and mixed landscapes.

| | Power to detect an effect of GM | | |
|---|---|---|---|
| | BBS (minimum/maximum power)[1] | | |
| %age uptake | Linnet[2] | Skylark[2] | Yellowhammer[3] |
| 20 | 58/94 | 22/79 | 48/85 |
| 40 | 71/100 | 29/96 | 71/97 |
| 60 | 73/100 | 29/100 | 81/100 |
| 80 | 81/100 | 41/100 | 78/100 |

**Table 2. The approximate power with which given percentage conversions of conventional crops to GM are expected to be detectable based on the power analysis results.** Note that the figures are based on important assumptions that are described in Annex 8 and Hails *et al.* 2013[24].

1. Maximum power here refers to the most powerful analysis conducted among the beet, potato and maize cropping patterns for the whole of England and minimum to the least powerful. Power varies because of both sample size and location i.e. the number of survey squares in which both the focal species and focal crop were found. Treating squares in arable, pastoral and mixed landscapes individually generally provided lower power, with some species-crop combinations failing to reach 50% power even with 80% predicted conversion to GM.
2. For linnet and skylark, maximum power = beet, minimum = maize.
3. For yellowhammer, maximum power = beet, minimum = potato.

**Key findings:** the results of the BBS power analyses illustrate the potential of this ESN to detect small effects on bird populations of changes in crop management. For this analysis, realistic scenarios were developed by taking the current regional distribution of each crop type into account in postulating future distributions of a GM variety by re-sampling the real bird and ES data from BBS squares with respect to these current distributions.

This analysis illustrates the key point that the power to detect change increases with the length of time since the change of management, as would be expected if the impact on the population is cumulative. The probability of it being detected therefore increases each year.

Differences in power were seen between species and among landscape types (arable, pastoral and mixed). This reflects variation in the level of coincidence between species' distributions and those of the (simulated) distributions of GM crops, as well as (perhaps) regional variation in the impact of the environmental change, i.e. the introduction of ES stubble (such variation in impact may or may not occur with a real GM crop). The greater power in analyses at the all-England level reflects the larger sample sizes in those analyses. The variation in power with respect to crop type and bird species illustrates that habitat preferences or distribution of a species can influence their sensitivity to specific environmental changes and therefore the likelihood that it will be affected by a GM crop. This suggests that it would be wise to consider GS of multiple species that might plausibly be affected by a given GM crop in assessing whether environmental impacts could be occurring.

A major caveat of this analysis is that it assumes that the impacts of GM crops will operate in the same way and be of similar magnitude as found for ES stubble management options. In addition, it assumes that either the same species would be affected or that other species affected and monitored would respond in the same way. This is unlikely to be the case. If GM crops were to result in unanticipated effects, different species might be affected and the size of effect might be different. However, if cultivation of GM crops were to have a biological effect of a similar (or greater) magnitude on one or more common bird species, such effects on population growth rates could be detected by analyses analogous to those conducted here. It will be important to consider, *a priori*, the species that are most likely, given their ecology and distribution, to be sensitive to the particular cropping changes that occur in practice.

In conclusion, these analyses show that BBS data have the potential to detect small effects on the populations of some bird species if these resulted from the cultivation of GM crops. The power of analyses of individual species varies with the distribution of the crop of interest. Thus, a wise application of the ESN data in BBS would be to identify those bird species whose ecology and distribution suggest that they will be exposed to a GM crop and to undertake the kind of analyses demonstrated here a few years after the crop is released to the market.

### 6.3 Case study using Countryside Survey data

A different approach was taken to analyse the CS data, as data are not collected annually. Therefore there is no temporal component here, and a spatial analysis is used to compare two treatments.. A complete description is provided in Annex 9 and Hails *et al.* 2013[24].

A set of power analyses were undertaken for data sampled only in maize and potato as there were insufficient CS plots located in sugar beet fields. Rather than species counts, data on vegetation species richness were analysed (see Annex 9). Vegetation subplots located within the crop field in the 2007 Countryside Survey were used to determine the power of detecting changes in species richness resulting from uptake of each GM crop.

Three different levels of GM uptake were assumed (20%, 40% and 60%). Species richness was modelled to reflect varying levels of difference between GM and non-GM crop fields. This was for 20%, 30%, 40%, 50% and 75% changes (reductions) in species richness. After taking multiple samples of plots (100 simulations) for each uptake and change scenario, the proportion of simulations where this uptake term (i.e. GM/non-GM) was statistically significant provided the power of detection for the proposed change and uptake. Table 3 shows the power to detect a 50% decrease in vegetation species richness in maize and potato fields at different levels of GM uptake.

Analysis of plots in potato fields had higher power, probably due to the greater sample size as compared to maize (i.e. more CS plots were located in maize than in potato fields). For potato a relatively large effect (50% change in mean species richness) can be sufficiently detected at 20% GM uptake (Figure 7a). For maize there is a high probability of detecting differences only at higher levels of GM uptake. When considering small differences of around 20%, the power was very low. Power is higher for 40% and 60% uptake than 20% and 80% because in order to detect significant differences between two distinct groups sample size needs to be high in both. This is clearly optimal (and equivalent) in the 40% and 60% uptake scenarios (Figure 7b).

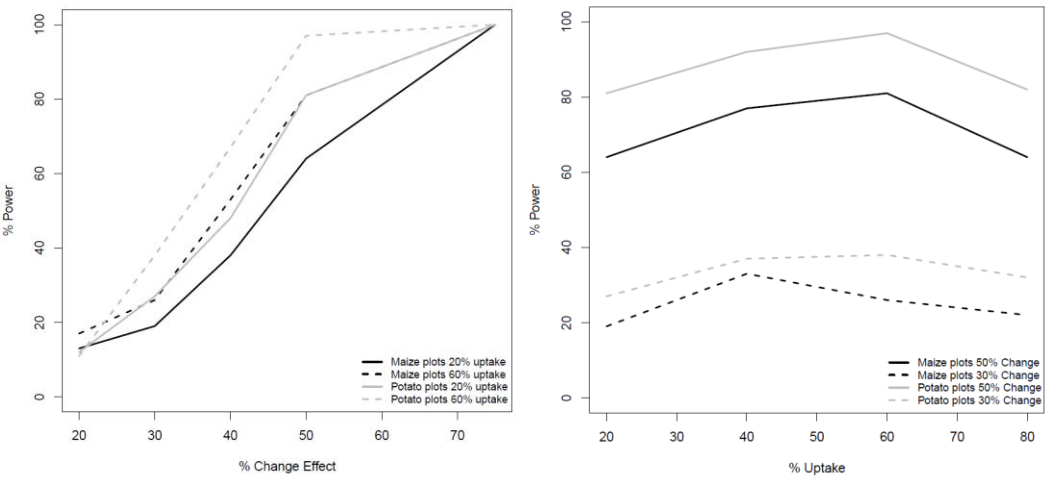
This analysis was repeated under the assumption that the location of GM cultivation or levels of GM uptake would only be known for each 1km square, not for individual fields and sub-plots. Table 4 shows that in this case the power of the CS to detect change is lower than when the location of fields where the GM crop is cultivated is known (as shown in Table 3).

| | Power | |
|---|---|---|
| % GM uptake | Maize | Potato |
| 20% | 0.64 | 0.81 |
| 40% | 0.77 | 0.92 |
| 60% | 0.81 | 0.97 |

**Table 3. Power to detect the effect of a hypothetical 50% change in mean plant species richness at Subplot (field) level given the proportion of plots that adopted GM.** (Test data were simulated from 2007 survey data, figures show the % of simulations in which a significant effect was detected).

| | Power | |
|---|---|---|
| % GM uptake | Maize | Potato |
| 20% | 0.11 | 0.32 |
| 40% | 0.18 | 0.71 |
| 60% | 0.39 | 0.72 |

**Table 4. Power to detect the effect of a hypothetical 50% change in plant species richness due to the introduction of GM cropping at the 1km square level, given different proportions of plots that adopted GM.** (Test data were simulated from 2007 survey data, figures show the % proportion of simulations in which a significant effect was detected).



(A) (B)

**Figure 7. Power to detect changes in mean species richness against % change effect (A) and % GM cultivation uptake (B).** In (A) relationships are shown for maize plots (black lines) and potato plots (grey lines) and uptake scenarios of 20% (solid lines) and 60% (dashed lines). In (B) relationships are shown for maize plots (black lines) and potato plots (grey lines) and change scenarios of 50% (solid lines) and 30% (dashed lines).

**Key findings:** The results from the CS power analysis show that this survey could be used for GS of GM crops, although it occurs only once every approximately eight years. The spatial analysis used here demonstrated a high power to detect relatively large (50%) differences in vegetation species richness between two treatments in maize and potato crops (Figure 7a). There were insufficient sample sites in sugar beet fields to conduct a power analysis. This illustrates the point that for less widely cultivated crops, or crops with localised distribution, the number of CS sample sites available may be insufficient.

The CS analysis also illustrates that for this survey the level at which the location of GM cultivation is known affects the power to detect change. Power was greatly reduced if it was assumed that the location of GM cultivation was known only at the 1km square level, rather than at the field level. This has a large influence on the power of the CS to detect change due to the fine resolution of the data collected. The UKBMS and BBS data is analysed at the 1km square level and so field level information on the location of GM cultivation would not increase the power of the analysis for these networks. Likewise, the water quality data is sampled at specific points, but is influenced by the entire water catchment. Therefore, information on the levels of GM uptake at the 1km square or catchment level would probably be equally informative as field level information.

### 6.4 Water quality monitoring programme analysis

The WQMP has a number of distinctive features which demand a different approach to data analysis to that adopted for the other ESNs. Full details are provided in Annex 10. In this report a spatial and temporal analysis was used to look at the power to detect changes in mean nitrate concentration at a national or regional scale following a hypothetical introduction of GM maize.

Water quality monitoring sites, with a reasonably complete 10-year time series of measurements, were identified in the main areas of maize cultivation. A mixed-effects model was fitted to the data to characterise the temporal variation in water quality at these sites. Time, day of the year and rainfall were included as explanatory variables to minimise the unexplained variation and help reveal any changes in water quality arising from cultivation of GM crops. The coefficients from the model were then used to stochastically simulate 500 replicate time series with the same properties as the original data. The synthetic data was modified by the inclusion of a hypothetical GM crop impact proportional to the coverage of that crop upstream of each site. The 500 replicate time series were then analysed using the same mixed-effects model as before, but with an additional term representing the GM crop impact. The proportion of the time series yielding a statistically significant GM crop effect was taken to indicate the power of the test. The simulation was repeated for a range of scenarios to examine how power changes with key factors such as the level of GM uptake, duration of monitoring and number of monitoring sites.

The analysis undertaken for the WQMP necessarily makes a number of simplifying assumptions about the future uptake and impact of GM crops. Full details are provided in Annex 10.

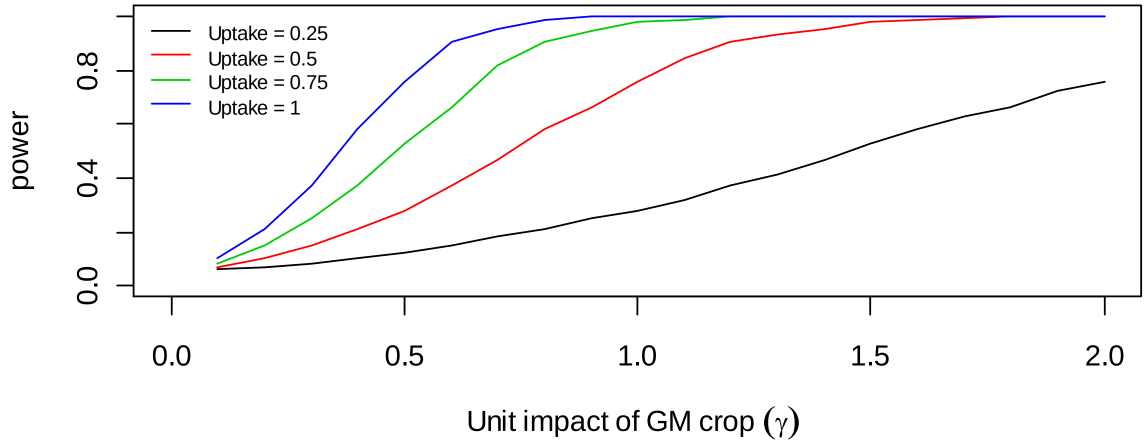


**Figure 8. Power for a range of values of unit impact of the GM crop (γ) and level of uptake of the GM crop (U)**. For a 10 year time period using 190 sites.

Figure 8 shows how the power to detect change over a 10 year period (5 years of data before and 5 years of data after the introduction of the GM variety) varies with the level of uptake and the unit impact of the GM variety. Power is high (>80%) if all the maize is GM and the GM variety increases losses of nitrate by at least 50% (γ>=0.5). At lower levels of GM uptake (<=50%), only very large increases in nitrate loss (>=100%) will have a high chance of being detected.

The effects of the time period and the number of sample sites were also modelled and details of this analysis are presented in Annex 9. Increasing the time period of monitoring (10 years before and 10 years after introduction of GM crops) increased the power to detect change. For example, the power to detect a 50% increase in nitrate loss was high (>80%) at 75% GM uptake. Likewise increasing the number of sample sites to 380 increased the power to detect change. The power to detect a 50% increase in nitrate loss was again high (>80%) at 75% GM uptake.

**Key findings:**
The results of the simulations suggest that the existing WQMP can detect adverse impacts of GM maize on mean nitrate concentration, but that power will be high (>80%) only if (i) GM maize is widely adopted (uptake is at least 75%), (ii) GM varieties cause a large (>= 50%) increase in pollutant losses relative to conventional varieties, and (iii) at least 10 years of monitoring data is available from at least two hundred affected monitoring sites.

The analysis focussed on the power of the WQMP to detect changes in mean nitrate concentration arising from GM maize. There are reasonable grounds for believing that power will be lower for other crops and other chemical determinands. This is because maize is more commonly grown than potatoes or sugar beet (the other crops for which applications for cultivation of GM varieties are currently in the EU regulatory pipeline) and nitrate has the lowest temporal variation in concentration of the three determinands examined.

Relative to other ESNs, the WQMP has a high number of monitoring sites and a high frequency of sampling, and so generates a very large volume of data with which to analyse GM crop impacts. On the other hand, individual crop types (maize, sugar beet and potatoes) rarely cover more than 10% of a catchment's area, and more typically cover just 1 - 5%. This means that GM crops can theoretically have a large effect on pollutant losses at the field scale and yet have an only minor impact on water quality at catchment scale. Power to detect changes in water quality can therefore be low. By contrast, the relatively small plots surveyed by the CS and BBS could have very high levels of GM crop coverage, and the localised impacts of those crops could theoretically be very pronounced and easier to detect.

The Environment Agency already uses results from the WQMP to track national *changes* in water quality. It is important to recognise that this work addresses the more difficult problem of attempting to *attribute* changes in water quality to a specific cause (in this case the introduction of GM crops). As this is an assessment of hypothetical impacts arising from a very specific, hypothetical change in agricultural practice, the results should not be interpreted as indicating the power of the network to detect other types of water quality change.

### 6.5 Conclusions from the power analysis

The results of the power analysis show that **in the UK there are ESNs which could be used to detect unanticipated effects resulting from the cultivation of GM crops**. However, **the limitations of what ESNs can be expected to deliver should be clearly acknowledged**. It will be important to select ESNs to contribute to GS with care, considering metrics that are likely to respond and species whose geographical distributions provide good overlaps with those of focal crops. It is likely that the uptake of GM crops will need to be quite extensive and the local effects quite significant before effects are detectable. In addition, there are likely to be time lags before detectability becomes high. Thus, the data collected by ESNs are not likely to be appropriate for detecting impacts of localized cropping changes, small effects or impacts on rare species. Nevertheless by specifically analysing the data collected by ESN for effects which could be correlated to GM crops, it may be possible to detect any adverse effects at an earlier stage than would otherwise have been possible.

# Section 7
# Using Environmental Surveillance Networks for General Surveillance

## 7.0 Introduction

This section of the report sets out three options for how ESNs could be used for GS: making use of existing reporting, specific data analysis and supplementary monitoring. The benefits and limitations of these approaches are discussed. Recommendations on key factors to consider in determining a strategy for the use of ESN are then made.

## 7.1 Options for the use of ESN for General Surveillance

### 7.1.1 Making use of existing reporting

Each of the four ESNs report on a set of key indicators and highlight where significant changes have occurred. In the UK, the devolved administrations (England, Scotland, Wales and Northern Ireland) establish biodiversity indicators. A revised set of indicators for England has recently published[25]. This includes specific indicator species for farmland including butterflies, birds, bats and plant species richness as measured by the networks described in this report. Work is underway to refine and further develop biodiversity indicators.

Making use of annual reports, or general indicator statistics, represents a minimal resource approach to GS. As described in Section 5.2.1 of this report, ESNs typically undertake a temporal data analysis, which seeks to identify significant declines or increases over time. If large unanticipated adverse effects did result from the cultivation of GM crops over extensive areas, it is possible that these would be detected by the data analysis currently undertaken by ESNs. The effect would need to very be large to be visible against background trends or variability. Another factor to consider is that although significant effects are often recorded by ESNs, these may be as a result of a range of different drivers and further analysis would be needed to determine correlates of change. For example, the England farmland bird abundance indicator may show a decline based on the national average, but it would not show if this effect was focussed in GM cropped areas. Following the introduction of GM crops, all such effects would need to be investigated further if this approach were adopted to determine whether they could be caused by GM cultivation.

### 7.1.2 Specific data analysis

Specific analysis of the data collected by ESNs can increase the probability of detecting if unanticipated effects are correlated with the cultivation of GM crops. This would have the advantage that smaller adverse effects could be detected more quickly. It may also be possible to identify if adverse effects occurred only at a regional or local scale. There are, however, still limits as to the advantages accrued by undertaking such additional analysis. It will not be possible to detect small impacts rapidly or impacts on rare species and such analysis will have limited power at low levels of GM uptake.

Options for undertaking such analysis are discussed in Section 5 of this report. An analysis which includes a spatial comparison of two treatments (GM and non-GM) would generally be most effective in addressing the question of whether the cultivation of GM crops correlates with the occurrence of adverse effects. It is important that such an analysis is only undertaken if the results will be meaningful. Therefore the power of an analysis to detect change and the size of change which could be identified should be established *a priori,* this aids the interpretation of a non-significant result.

As noted in Section 4.2 of this report, it would be useful to define a common set of assessment end points and corresponding measurement endpoints that can be used routinely to report on the status of the farmed environment and identify correlations with possible influencing factors. Data analysis could then focus on these indicators.

In some cases raw data collected by the networks is publically available or could be made available to the consent holder. In the UK, many of the existing ESNs would be well placed to undertake such analysis if required. There will, however, be costs associated with any specific data analysis. If it is considered that such analysis should be undertaken it will be necessary to clearly assign responsibilities for undertaking and funding such work.

### 7.1.3 Supplementary monitoring

GS is intended to detect unanticipated adverse effects, which were not identified in the environmental risk assessment or in pre-market trials. Such effects are, by definition, likely to be small, at least initially. Small effects are less likely to be detected by existing ESNs and may only become significant, and detectable, if the crop were widely cultivated and the effects were cumulative over a number of years.

As illustrated in Section 6, there are limits to the size of effect and the speed at which that effect could be detected using ESNs. In addition, there are likely to be significant time lags before the probability of detection becomes high. In all cases, the power to detect change is higher if a greater number of sample sites are available.

To detect effects more quickly, at lower levels of change, or at a stage when GM crops were less widely cultivated, supplementary monitoring would be needed. This would involve supplementing existing ESNs with additional sites, measurements or time points. This would be most efficiently done by choosing sites to facilitate a direct comparison between GM and conventional crops.

For example, at low levels of GM uptake, extra sample sites could be added to ensure that sufficient were available in areas of GM uptake. Supplementary monitoring would aim to detect effects before cultivation became widespread and provide an early warning system. If levels of GM cultivation became more widespread, supplementary monitoring may no longer be needed as existing sample sites would be effective in detecting adverse effects.

Such supplementary monitoring has previously been implemented to increase the power of existing ESNs to address specific hypotheses, for example to increase the power of the BBS to detect effects of entry level ES on breeding farmland bird populations in England[26]. In the latter case, clear biological effects were not detected, but subsequent analyses using an alternative analytical design maximizing temporal rather than spatial replication proved more sensitive[27]. The latter provided the basis for the BBS power analysis described in Hails *et al.* 2013[24] and gives an important message that some changes in sampling intensity (and therefore effort and investment) have greater effects in some contexts than others (i.e. increased temporal replication can have greater effects on sensitivity than increases in the spatial sample: statistically significant effects were detected using multiple years of historical ESN data and without any additional monitoring).

## 7.2 Recommendations on the use of ESNs for General Surveillance

The analysis presented in this report demonstrates existing ESNs in the UK could be used as part of GS for unanticipated effects of GM crops. It will be for decision makers to determine if or how ESNs should be used. In developing an approach, it is important to note that there is no *de facto* reason to expect that GM crops would have adverse effects if risks have not been identified in the environmental risk assessment. Any change in agricultural practice could potentially have impacts on the environment and it is important that the use of ESNs for GS of GM crops is considered in this context. There will be limits to the benefits that will accrue through implementation of monitoring relative to the costs of gathering the information. It is therefore recommended that **GS should make use of ESNs which are already in place or will be introduced in line with general reviews of national monitoring requirements.**

As described in Box 1 existing initiatives in the UK are examining the roles of ESNs. It is recommended that GM crops should be considered as one potential influencing factor as part of wider reviews of ESNs. For example, whilst it may be desirable to address gaps in the monitoring of environmental parameters (see Annex 6) it would not be proportionate to require this only for purposes of monitoring GM crops, given that GM crops also undergo a rigorous environmental risk assessment prior to authorisation.

**To maximise the value of existing ESNs for GS of GM crops, specific analysis of the data collected will be required**. A combined spatial and temporal, comparative (GM vs non-GM) analysis will maximise the power to detect change correlated with GM cultivation whilst reducing the number of false positives. Before an ESN is used as a formal monitoring tool, the statistical power it can provide to detect change should be determined.

This approach will need a set of assessment endpoints and corresponding measurement endpoints to be determined. As described in Section 4.2, it would be desirable to use an established set of measurement endpoints which could be used routinely to report on the status of the farmed environment and identify correlations with any driver of change. Preferably this should make use of an ecosystems function or ecosystem service approach. Such work should form part of wider initiatives on agriculture and environmental policy.

Power analysis could then be used to determine the size of change which it would be possible to detect and the speed with which it could be detected for given levels of uptake of GM crops. Even using an ecosystems function or service approach, it may be difficult to determine scientifically the types of change it would be desirable to detect given current knowledge of arable ecosystems.

The expectations of what GS using ESNs can achieve should be made clear in advance. **To detect effects more quickly, at a lower level of change, or at a stage when GM crops were less widely cultivated, supplementary monitoring would be needed**. Decision makers will need to come to a conclusion as to how sensitive they wish monitoring for unanticipated effects using ESNs to be.

## Section 8
## Further action if potential adverse effects are detected

**GS may be used to identify correlations between effects and potential drivers of change, but cannot be used to conclusively demonstrate a relationship between cause and effect**. Adverse effects could have many causes (e.g. unusual weather conditions, changing management of non-GM crops) and it would not be pragmatic to take further action every time an adverse effect was observed.

If a correlation between GM cultivation and an impact on an environmental parameter were to be identified further investigation would be needed. This further investigation should aim to establish whether the effect is caused by GM crops and if it is harmful.

GS, using a combination of tools - the Farm Questionnaire, ESNs and a literature review - provides a “tripwire” which would indicate if an adverse effect could have resulted from GM cultivation. If a correlation between an adverse effect and GM cultivation were observed by any one of these tools, further investigation would be needed to determine the cause.

The following trigger points should instigate further action:

- greater than 10% of respondents to the FQ report an adverse effect
- an adverse effect that could be associated with GM crop cultivation is detected by analysis of the data collected by ESNs
- a literature review, conducted according to the methodology recommended by EFSA reveals an adverse effect which would be relevant to the commercial cultivation of GM crops

The nature of the trigger will inform what type of further investigation would be most appropriate. In many cases further analysis of available data would be an appropriate first step. In some cases additional studies, such as manipulative field experiments may be needed. Any such studies would need to follow an experimentally robust design to test the relationship between cause and effect and determine whether the observed adverse effect could be due to the cultivation of the GM crop. Depending on how strong the evidence of an effect is and the type of effect which has occurred it may be necessary to halt the commercial cultivation of the GM crop whilst this is investigated.

It is not always straight forward to determine whether an observed effect is actually adverse. In addition it must be noted that a correlation with cultivation of GM crops does not demonstrate that GM crops have caused the effect. For these reasons it is recommended that **if a potential adverse effect is detected, independent scientific advice should be sought to interpret data and determine what kind of further investigation should be triggered**.

## Section 9
## Conclusions and recommendations

In the UK the data collected by ESNs could be used to identify correlations between adverse effects and drivers of change. Any change in agricultural practice could have environmental impacts and it is important that GM crops are seen in this context. It is desirable to make best use of ESNs to identify relationships between agricultural practices and environmental impacts. Against a background of sustainable intensification making optimal use of ESNs for monitoring the impacts of any changes in agricultural practice would be desirable. It would not be proportionate to reconfigure ESNs solely for purposes of monitoring GM crops.

The use of ESNs is only one part of PMEM. It is important that a combination of tools is used for GS to maximise the chances of detecting adverse effects should they occur. In addition it is important that situations which require CSM are clearly identified and appropriate monitoring is put in place. Finally PMEM follows a robust environmental risk assessment prior to the authorisation of any GM crop. There will be limitations as to what PMEM can deliver and it is therefore important that the emphasis remains on ensuring effective environmental risk assessment.

**Key messages and recommendations:**

**Case specific monitoring**

1. CSM provides an appropriate route for the testing of specific hypothesis which can only be addressed under the range of conditions that are represented by commercial farming

2. Identifying the need for CSM should not be seen as an indication that unacceptable risks remain and that authorisation should be refused.

**General surveillance**

3. GS should focus on environmental parameters in close contact with the GM crop.

4. As GS is not hypothesis-driven the same approach could be taken for all arable GM crops.

5. GS should use a combination of tools, to maximise the chances of detecting any adverse effects.

6. An agreed set of assessment end points and corresponding measurement endpoints that can be used routinely to report on the status of the farmed environment and identify correlations with possible influencing factors should be defined

**The Farm Questionnaire**

7. The FQ is a useful tool which can be employed as part of GS of GM crops and as such its potential as an information source should be maximised wherever practical

8. The FQ should also collect information on cultivation and agronomic practices in fields in the years following GM crop cultivation

**Existing surveillance networks**

9. Power analysis is a useful tool for determining the probability of detecting change using ESNs and therefore for interpreting a finding of no significant effect.

10. In the UK ESN could be used to detect unanticipated effects resulting from the cultivation of GM crops

11. GS should make use of ESNs which are already in place or will be introduced in line with general reviews of national monitoring requirements.

12. To maximise the value of existing ESNs for GS of GM crops, specific analysis of the data collected would be required.

13. To detect effects more quickly, at a lower level of change, or at a stage when GM crops were less widely cultivated, supplementary monitoring would be needed

**Further action if potential adverse effects are detected**

14. GS may be used to identify correlations between adverse effects and potential drivers of change, but is unlikely to conclusively demonstrate the cause of an adverse effect

15. If a potential adverse effect is detected, independent scientific advice should be sought to interpret data and determine what kind of further investigation should be triggered.

---

## References

[1] ACRE terms of reference: http://www.defra.gov.uk/acre/about/terms-of-reference/

[2] Regulation (EC) No 1829/2003 of the European Parliament and of the Council of 22 September 2003 on genetically modified food and feed http://ec.europa.eu/food/food/animalnutrition/labelling/Reg_1829_2003_en.pdf

[3] Directive 2001/18/EC of the European Parliament and of the Council of 12 March 2001 on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC http://europa.eu/legislation_summaries/agriculture/food/l28130_en.htm

[4] Regulation (EC) No 1830/2003 labelling of genetically modified organisms and the traceability of food and feed products produced from genetically modified organisms and amending Directive 2001/18/EC http://europa.eu/legislation_summaries/environment/nature_and_biodiversity/l21170_en.htm

[5] COUNCIL DECISION of 3 October 2002 establishing guidance notes supplementing AnnexVII to Directive 2001/18/EC of the European Parliament and of the Council on the deliberate release into the environment of genetically modified organisms and repealing Council Directive 90/220/EEC http://www.biosafety.be/PDF/2002_811.pdf

[6] EFSA Panel on Genetically Modified Organisms (GMO). Guidance on the Environmental Risk Assessment of genetically modified plants. EFSA Journal 2010;8(11):1879 http://www.efsa.europa.eu/en/efsajournal/doc/1879.pdf

[7] EFSA Panel on Genetically Modified Organisms (GMO). Guidance on the Post-Market Environmental Monitoring (PMEM) of genetically modified plants. EFSA Journal 2011;9(8):2316 http://www.efsa.europa.eu/en/efsajournal/doc/2316.pdf

[8] Centre for Evidence-Based Conservation (2010) Guidelines for Systematic Review in Environmental Management. Version 4.0. Environmental Evidence: www.environmentalevidence.org/Authors.html

[9] EFSA. Application of systematic review methodology to food and feed safety assessments to support decision making. EFSA Journal 2010; 8(6):1637 http://www.efsa.europa.eu/en/efsajournal/doc/1637.pdf

[10] Suter, G. W. (2000) Generic Assessment Endpoints Are Needed for Ecological Risk Assessment. Risk Analysis, Vol 20 (2): 173-178

[11] Schmidt K., Wilhelm R., Schmidtke J., Beissner L., Mönkemeyer W., Böttinger P., Sweet J., Schiemann J. (2008) Farm questionnaires for monitoring genetically modified crops: a case study using GM maize. Environ Biosafety Res 7 (3): 163-179

[12] Monsanto Company (2011) Post market monitoring of insect protected Bt maize MON810 in Europe 2010, Appendix B http://ec.europa.eu/food/food/biotechnology/docs/2010_mon810_farmer_questionnaire_survey_en.pdf

[13] BASF Plant Science Company GmBH (2011) Post-market monitoring report 2010 Amylopectin Potato EH92-527-1 Annex 6 http://ec.europa.eu/food/food/biotechnology/docs/post_market_monitoring_report_annex_6_en.pdf

[14] Monsanto Europe, S.A. (2010) Annual Monitoring Report of the cultivation of MON810 in 2010 http://ec.europa.eu/food/food/biotechnology/docs/report_mon_810_en.pdf

[15] Environmental monitoring of GM crops. Hearing organised by the European Commission. March 2012. http://ec.europa.eu/food/food/biotechnology/docs/agenda_29032012_en.pdf

---

[16] Monsanto Europe, S.A. (2009) Annual monitoring report on the cultivation of MON810 in 2008.

[17] Glandorf, D.C.M. (2012) General Surveillance of genetically modified plants: possibilities for implementation in the Netherlands. RIVM report. http://www.rivm.nl/dsresource?objectid=rivmp:118436&type=org&disposition=inline

[18] Baillie, S.R. & Rehfisch, M.M. (eds) (2006) National and site-based alert systems for UK birds. Research Report 226. BTO, Thetford. http://www.bto.org/sites/default/files/u196/downloads/rr226.pdf

[19] BBS population declines and increases. http://www.bto.org/about-birds/birdtrends/2011/summary-tables/bbs-population-declines-increases

[20] Siriwardena G.M., Baillie S.R., Buckland S.T., Fewster R.M., Marchant J.H., Wilson J.D. (1998) Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common birds census indices. Journal of Applied Ecology 35: 24-43

[21] British Trust for Ornithology (2010) Assessing the impact of the loss of set-aside stubbles on the winter ecology, behaviour and distribution of farmland birds and their breeding population trends. Defra report BD1641. http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&ProjectID=16375&FromSearch=Y&Publisher=1&SearchText=BD1641&SortString=ProjectCode&SortOrder=Asc&Paging=10#Description

[22] Perry J. N., Rothery P., Clark S.J., Heard M.S., Hawes C. (2003) Design, analysis and statistical power of the Farm-Scale Evaluations of genetically modified herbicide-tolerant crops. Journal of Applied Ecology 40, 17-41 http://www.rothamsted.ac.uk/pie/sadie/reprints/perry_rothery_clark_heard_hawes_2003.pdf

[23] Freeman S.N, Newson S.E. (2008) On a log-linear approach to detecting ecological interactions in monitored populations*.* Ibis*,* **150**, 250-258.

[24] Hails, R.S., Freeman, S., Henrys, P. and Siriwardena, G. (2013). Determining and increasing the sensitivity of existing environmental surveillance monitoring networks to detect unanticipated effects that may occur in the environment in response to the cultivation of genetically modified crops. Defra final report CB0304 http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&ProjectID=18171&FromSearch=Y&Publisher=1&SearchText=CB0304&SortString=ProjectCode&SortOrder=Asc&Paging=10#Description

[25] Defra. A strategy for England's wildlife and ecosystem services. Biodiversity 2020 indicators. 2012 assessment. http://www.defra.gov.uk/statistics/files/England-Biodiversity-indicators-2012-FINALv2.pdf

[26] Davey C.M., Vickery J.A., Boatman N.D., Chamberlain D.E. Parry H.R. & Siriwardena G.M. (2010). Assessing the impact of Entry Level Stewardship on lowland farmland birds in England. Ibis*:* 152: 459-474.

[27] Baker D.J., Freeman S.N., Grice P.V. & Siriwardena G.M. (2012) Landscape scale responses of birds to agri-environment management: a test of the English Environmental Stewardship Scheme. Journal of Applied Ecology 49: 871-882.