environmental science & policy 11 (2008) 579–587

available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/envsci

Providing the evidence base for environmental risk assessments of novel farm management practices

J. Storkey[a,*], D.A. Bohan[a], A.J. Haughton[a], G.T. Champion[b], J.N. Perry[a], G.M. Poppy[c], I.P. Woiwod[a]

[a] Department of Plant and Invertebrate Ecology, Rothamsted Research, Harpenden, Herts, Hertfordshire AL5 2JQ, UK
[b] Broom's Barn, Higham, Bury St. Edmunds, Suffolk IP28 6NP, UK
[c] School of Biological Sciences, University of Southampton, Boldrewood Campus, Southampton, SO16 7PX, UK

ARTICLE INFO

Published on line 14 July 2008

Keywords:
Farm scale evaluations
FSEs
Genetically modified crops
Farmland biodiversity

ABSTRACT

An environmental risk assessment of a new agricultural management practice depends upon the provision of empirical evidence of cause and effect. This will invariably be derived from comparative experiments testing the null hypothesis that a change in management will have no effect on an assessment endpoint (the metric on which policy decisions will be based). Crucial to the design of these experiments is the answer to the question of 'what to measure?'. The selection of these measurement endpoints and the design of sampling protocols will be determined by the properties of the environmental stressors associated with the change in management practice and the taxa that are exposed to their effects, as well as logistic and financial considerations. The rationale for deciding what to measure in the context of these various criteria is reviewed. For a measurement endpoint to be a valid indicator of the risk of a negative impact of management on the assessment endpoint, a predictable and quantifiable link must be made between the two. It should also be recorded at the appropriate taxonomic resolution to safely assume that all the constituent parts will both respond in a similar way to the management stressor and have a similar effect on the assessment endpoint. Protocols must be designed with the spatial and temporal properties of the management stressor and the measurement endpoint in mind and a consideration of the statistical power of the experiment to detect changes. Where there is a lag in the response time of a measurement endpoint to a stressor due to inertia in the system, an accurate measurement of the effect of the novel management may require experiments running over several years. Throughout, care must be taken that the statistical and biological validity of a sampling regime is not compromised in the face of logistic and financial pressures. The Farm Scale Evaluations of the management of Genetically Modified Herbicide Tolerant crops are presented as a case study to illustrate the concepts discussed.

1. Introduction

Concern about the environmental impact of farmland management practices has led to calls for the development of an integrated risk assessment framework for agricultural systems (ACRE, 2006). Such an assessment will involve a number of logical steps linked within a conceptual framework (Poppy, 2003). For any new management practice, there may be a number of factors that could negatively impact the flora and fauna found in the farm landscape. The first step is, therefore,

* *Corresponding author.*
E-mail address: jonathan.storkey@bbsrc.ac.uk (J. Storkey).
1462-9011/$ – see front matter
doi:10.1016/j.envsci.2008.06.001

to clearly define each of these potential 'environmental stressors'. For example, stressors associated with the introduction of a novel crop could include a change in the timing of cultivation, nature of the agro-chemicals applied and biological interactions with non-crop organisms including crop-weed competition. The next step is to identify the potential impact of the stressor in terms of the taxa that may be affected either directly or indirectly and the 'exposure profile' of the stressor. The exposure profile is defined as the temporal and spatial coincidence of the stressor and the taxa at risk. At this stage a number of potential 'assessment endpoints' may be identified that could be impacted by the environmental stressor. Assessment endpoints are metrics that are used to inform policy and may be indicator species or components of the agricultural landscape that are relevant to specific policy objectives. For example, the UK Government has made a Public Service Agreement to reverse the long-term decline in farmland birds (DEFRA, 2007).

The environmental stressor and assessment endpoint may be linked via a separate 'measurement endpoint'. A measurement endpoint defines the indicator of change that will actually be recorded as part of a comparative study of the environmental impact of a new farm management practice and provides the evidence base for policy decisions. How this evidence base is established in terms of the choice of measurement endpoint and design of effective sampling protocols provides the focus of this paper. As well as being determined by the nature of the environmental stressor and its relationship with the assessment endpoint, the choice of measurement endpoints will also be a factor of practical constraints (whether sampling methodologies are available) and availability of human and financial resources (Qi et al., 2008). The rationale for taking into account all of these factors when selecting a measurement endpoint and designing sampling protocols is reviewed with a particular emphasis on plants and invertebrates. Finally, the Farm Scale Evaluations (FSEs) of the environmental impact of management changes associated with the growing of genetically modified herbicide tolerant (GMHT) crops are used as a case study to illustrate the conceptual steps discussed.

2. Selection of measurement endpoints

Policy decisions to do with the introduction of a novel farm practice will be influenced by an assessment of the impact of the change in management on key taxa that are important because of they have an inherent value (for example, rare species (Walker et al., 2007)), they perform an ecosystem service (Kremen et al., 2007) or are an indicator of wider farmland biodiversity (Gregory et al., 2005). In some cases, it may be possible to measure the assessment endpoint directly; for example, where the concern is for flowering arable plants (Sutcliffe and Kay, 2000) or pollinators (Biesmeijer et al., 2006). However, for other assessment endpoints (e.g. national populations of farmland birds (Siriwardena et al., 1998)) it may not be always possible to monitor their response to management changes directly. In this case, the question of 'what to measure' is crucial. When the measurement endpoint differs from the assessment endpoint it must be responsive to the environmental stressor and be a clear indicator of change in the assessment endpoint. A predictable and quantifiable link between the two must be made explicit either experimentally or from the literature. For example, a direct causal link between the abundance of weed seed (measurement endpoint) in arable fields and their value to farmland birds (assessment endpoint) has been clearly demonstrated empirically (Moorcroft et al., 2002; Whittingham et al., 2006). The impact of changes in weed population dynamics as a result of farm management changes on bird populations can, therefore, be predicted (Watkinson et al., 2000). In contrast, whereas a strong relationship would be expected between weed biomass and invertebrate herbivore density, this is not always apparent in the literature (Bohan et al., 2005).

2.1. *Taxonomic level of measurement*

The nature of the association between the assessment and the measurement endpoint will also determine the appropriate level of taxonomic information recorded. Where the effect of an environmental stressor is linked to the assessment endpoint via a number of species that perform a similar ecosystem function, it may be appropriate to amalgamate species in a functional group. For example, arable weeds have been grouped according to their competitive ability and potential value to higher trophic levels (Marshall et al., 2003; Storkey, 2006) and the response of invertebrates to agricultural management has been studied at the level of trophic group (Hawes et al., 2003). In contrast, if the assessment endpoint is species diversity or the abundance of a species of conservation concern, the appropriate level of measurement will be the species (Walker et al., 2007). Further, it has been argued that, where there is large intra-specific variation in eco-system response or function (due to genetic diversity or phenotypic plasticity), the unit of measurement should be the individual (Hawes et al., 2005).

The choice of the taxonomic level for the assessment of change in a measurement endpoint protocol may also be influenced by a number of pragmatic considerations. Species may be recorded at the level of genus or family because identification to species may be difficult and time consuming. In addition, species may be grouped because they are sampled by the same sampling protocol. However, whatever the reason for amalgamating species, the possibility of compromising the biological validity of the measurement endpoint must always be considered. For the amalgamated group to be an effective indicator of change, all of its components must both respond similarly to the environmental stressor and have a similar impact on the assessment endpoint. In the absence of data to validate this assumption, therefore, it should be assumed that detail is required at least at the level of the species and any amalgamation above this level should be fully justified from knowledge of the ecology and the linkages between the organisms concerned.

3. Biological metrics

Once a measurement endpoint has been selected, in terms of the taxa that will be recorded to monitor the impact of a