Risk-based modelling of diffuse land use impacts from rural landscapes upon salmonid fry abundance

Sim M. Reaney[1], Stuart N. Lane[1], A. Louise Heathwaite[2] and Lucy J. Dugdale[3]

1 Institute of Hazard, Risk and Resilience and Department of Geography, Durham University, Durham, DH1 3LE, U.K.

2 Centre for Sustainable Water Management, Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, U.K.

3 Eden Rivers Trust, Units O&Q, Skirsgill Business Park, Penrith CA11 0DP, U.K.

Abstract

Research has demonstrated that landscape or watershed scale processes can influence instream aquatic ecosystems, in terms of the impacts of delivery of fine sediment, solutes and organic matter. Testing such impacts upon populations of organisms (i.e. at the catchment scale) has not proven straightforward and differences have emerged in the conclusions reached. This is: (1) partly because different studies have focused upon different scales of enquiry; but also (2) because the emphasis upon upstream land cover has rarely addressed the extent to which such land covers are hydrologically-connected, and hence able to deliver diffuse pollution, to the drainage network. However, there is a third issue. In order to develop suitable hydrological models, we need to conceptualise the process cascade. To do this, we need to know what matters to the organism being impacted by the hydrological system, such that we can identify which processes need to be modelled. Acquiring such knowledge is not easy, especially for organisms like fish that might occupy very different locations in the river over relatively short periods of time. However, and inevitably, hydrological modellers have started by building up piecemeal the aspects of the problem that we think matter to fish. Herein, we report two developments: (a) for the case of sediment associated diffuse pollution from agriculture, a risk-based modelling framework, SCIMAP, has been developed, which is distinct because it has an explicit focus upon hydrological connectivity; and (b) we use spatially-distributed ecological data

to infer the processes and the associated process parameters that matter to salmonid fry. We apply the model to spatially-distributed salmon and fry data from the River Eden, Cumbria. The analysis shows, quite surprisingly, that arable land covers are relatively unimportant as drivers of fry abundance. What matters most is intensive pasture, a land cover that could be associated with a number of stressors on salmonid fry (e.g. pesticides, fine sediment) and which allows us to identify a series of risky field locations, where this land cover is readily connected to the river system by overland flow.

Key words: diffuse pollution, hydrological connectivity, land cover, salmonids, fine sediment, risk

Introduction

There is growing realisation that the localised restoration of individual reaches of river can be undermined due to larger scales of influence, such as the delivery of fine sediment from eroding agricultural land. This approach is enshrined in the EU Water Framework Directive, which advocates holistic analysis (e.g. Newson, 1997), and it applies to land management activities like agriculture that drive diffuse responses but which collectively create particular point problems (e.g. increased flood risk, nutrient loading, fine sediment accumulation in river gravels). It has proved exceptionally difficult to demonstrate the extent to which diffuse activities are responsible for these point problems, not least because statutory monitoring agencies rarely design data collection strategies that reveal the characteristics of diffuse pollution (Harris and Heathwaite, 2005). For this reason, the use of mathematical modelling to identify the sources of diffuse pollution has been dominant, commonly in a risk based framework. In this framework, sources of risk are imagined to be distributed across a river catchment. Human activities (e.g. fertiliser additions) may combine combined with landscape attributes (e.g. soil type, local slope) to make certain sites more important sources of risk than others. Thus, reducing the risk to rivers is concerned with identifying the locations of the important sources of risk and embarking upon appropriate management interventions. This paper is concerned with two developments to this risk-based approach. First, it recognises that the sources of risk need to be connected hydrologically to the river network if they are to deliver their ‘risk’. This may be within a storm event, or through a series of storm events by repeated erosion and transport along a hydrological flow path. Existing analyses of diffuse pollution contain only a very rudimentary representation of this process. Second, some previous attention has been given to the spatial patterns of the sources of risks in landscapes, surprisingly little attention has been given to the collection of in-river data, distributed spatially across an entire catchment to test such predictions. Using conventional water quality samples to do this is difficult because of temporal variability, which necessitates many sampling points, measuring through time in order to obtain unbiased estimates of water quality parameters. The aim of this paper is to develop a reformulated approach to model the impacts of diffuse pollution (notably material such as fine sediment eroded from the landscape). We inform this approach by using ecological data (salmonid fry) collected from across the study catchment.

Modelling fish populations at the landscape scale

The potential role of landscape scale factors in river management is based upon the premise that they influence aquatic communities, in terms of chemistry, hydrology and the production and transfer of organic matter (Allan and Johnson, 1997). Reflecting the observation of Hynes (1975), that the valley rules the stream, landscape factors (e.g. soil type, land use) have been shown to influence instream water quality (Hunsaker and Levine, 1995; Johnson *et al.*, 1997). This influence has been extended to impacts upon instream organisms whether explored directly (e.g. Roth *et al.*, 1996) or indirectly, through the effects of landscape scale factors upon relevant reach-scale parameters (e.g. Richards *et al.*, 1997) such as food availability (e.g. Townsend *et al.*, 1997). The recognition that landscape scale factors matter has been part of a move towards the hierarchical interpretation of aquatic communities, in which factors that range in scale from microhabitat to the entire river basin interact to impact upon both where habitat is suitable and the degree to which an organism can move between suitable habitat sites (Poff, 1997; Armstrong *et al.*, 1998, 2003; Wang *et al.*, 2003; Durance *et al.*, 2006). Such work recognises the fundamental structuring effect of river basin drainage networks (e.g. Benda *et al.*, 2004), necessitating an upscaling of the focus of river restoration efforts (Harding *et al.*, 1998; Durance *et al.*, 2006) which have traditionally emphasised the riparian environment alone (Johnson and Gage, 1997; Folt *et al.*, 1998).

Incorporating the landscape-scale has not proved to be straightforward (Durance *et al.*, 2006). For instance, in relation to fish, a trade off has to be made between: improving reliability of ecological data at-a-point in time through more time-consuming repeat pass electrofishing; and capturing population variability in space, requiring less time at individual locations. Sacrificing time leads to better spatial resolution but enhanced spatial variance (Wiley *et al.*, 1997), or sampling-enhanced noise. However, assessment of a watershed or catchment factor has to have a spatial component, especially river basins with a range of land use activities and practices, where the mosaic of land uses found will create substantial spatial variability in instream water quality. These issues are compounded by: (1) possibly many limiting habitat influences; (2) spatial variability in the extent of habitat limitation as compared with other population controlling factors; (3) spatial variability in exactly what aspect of habitat is limiting (Pess *et al.*, 2002) and (4), in particular, inter-annual variation in recruitment that means there will be substantial temporal variability in any of the spatial data that are acquired. If reliable spatial datasets can be generated, even a partial explanation of their spatial structure by any one watershed factor is a significant challenge (Pess et al., 2002, Johnson and Gage, 1997). When larger-scale factors have been considered, results have been attributed to the scale implicit in the design of the study (Wang *et al.*, 1997; Stauffer *et al.*, 2000; Lammert and Allan, 1999; Durance *et al.*, 2006). Thus, the reach/riparian focus of much conservation work is not surprising, notably in the presence of results that can at times be contradictory (Rich *et al.*, 2003), and the difficulty of getting ecological data at a scale that matches the landscape emphasis.

This situation aside, landscape-scale factors still provide meaningful hypotheses for explaining both historical and current patterns of instream organism populations. The landscape hypotheses are founded upon the assumption that upstream factors influence the delivery of water, sediment, solutes and organic matter to locations that are suitable for a particular organism and so impact upon the local habitat suitability of that site in both positive (e.g. delivery of the organic matter required to support benthic macroinvertebrate populations) and negative (e.g. siltation of spawning redds) ways. Larger scale processes are assumed to be creating the template within which the small scale operates (Armstrong *et al.*, 1998; Stauffer *et al.*, 2000). Thus, one explanation for differing findings in relation to the importance of landscape factors is not methodological but due to substantive differences in the sources of risk in catchments and their propensity to be delivered to the river network. Indeed, emphasis upon landscape scale attributes has focused almost entirely on either abiotic metrics such as geology or relief and/or land use and management practices. It has not recognised the extent to which there is delivery of material from those land uses to the river system (Meador and Goldstein, 2003). Recent work in both hydrology (Kirchner *et al.*, 2000) and biology (Poff, 1997) has emphasised how landscapes can operate as large-scale filters (Burt and Pinay, 2005) in which the scales of variability of inputs to the system (e.g. rainfall) are fundamentally restructured by the time they become outputs (e.g. water quality). The focus in this paper is upon developing a modelling approach that can capture this effect, in a risk-based framework.

Model principles: diffuse pollution risk, connectivity and instream impacts

In the absence of spatially-distributed in-river data, and faced with the need to identify the locations within the drainage basin that are most likely to be sources of catchment risks, a number of modelling approaches have been developed. These can be classified into three main groups (Lane et al., 2006): (1) transfer function modelling – which predicts material export on the basis of simple empirical transfer functions driven by known inputs such as fertiliser and manure applications coupled with soil nutrient status (e.g. Jordan et al., 1994; Johnes, 1996; Johnes and Heathwaite, 1997; Herrmann et al., 2003; Ekholm et al., 2005); (2) land unit modelling – which applies physically-based ('mechanistic') models of sediment and nutrient cycling to individual land units in order to determine export (e.g. Priess et al., 2001; Weber et al., 2001; Binder et al., 2003; Wolf et al., 2005; Matthews, 2006; Vatn et al., 2006); and (3) land transfer modelling – which combines the kind of analysis described in (2) with a physically-based, sometimes dynamic, treatment of how material is transferred across the landscape (e.g. Adams et al., 1995; De Roo and Jetten, 1999). A proper treatment of the transfer process is necessary (Beven et al., 2005) as local, often small scale, hydrological pathways can exert a major control on whether or not material is delivered to drainage networks (e.g. Blackwell et al., 1999; Burt et al., 1999; Quinn, 2004) as well as the deposition and transformation processes that result (Harris and Heathwaite, 2005). Hydrological modelling, for instance, suggests a complex pattern of overland flow generation during an individual storm event (Lane et al., 2004), with saturated parts of the landscape both connected and disconnected by overland flow to the drainage network. These patterns of runoff generation and hydrological connection occur at spatial scales of the order of 10 m or less (Lane et al., 2004; Heathwaite et al. 2005), often related to quite subtle topographic attributes. This local scale hydrological structuring of the landscape may exert an important control upon the connectivity of sources of risk to the landscape (e.g. Figure 1), leading to the idea of Critical Source Areas (CSAs, Heathwaite et al., 2000) which are parts of the landscape that generate risks that can be readily delivered to the drainage network. Delivery is a critical process in determining whether or not a risky land cover produces material that can reach the river network. Once delivered, additional hydrological processes may impact upon the level of instream risk, such as when tributaries with different suspended sediment concentrations meet, resulting in the dilution of one by the other (e.g. Figure 2).

The extent to which these three modelling approaches capture delivery is variable. There remains a tendency either: (1) to treat delivery processes in a simplified way (e.g. as some function of distance from the nearest stream, Munafo et al., 2005); or (2) to apply models with a potentially sophisticated treatment of delivery but at coarse spatial scales (e.g. 1 km, Adams et al., 1995), losing much of the spatial detail known to drive hydrological response. In this paper, we present a new approach to the modelling of landscape risks that is suited to large rural river catchments, but that also recognises that the drivers of the delivery of risks at the catchment scale include processes that occur at small, potentially sub-field, spatial scales. This need to transcend scale is well-established (e.g. Muscutt et al., 1993; Haycock and Muscutt, 1995; Kuusemets and Mander 1999; McKergow et al., 2003; Quinn, 2004). In theory, physically-based dynamic water quality models (see Borah and Bera, 2003, 2004) ought to do this. They represent delivery implicitly and continuously as a catchment wets and dries using appropriate, commonly distributed, process rules. Borah and Bera review both the mathematical bases and parameterisation issues associated with a number of such models. Critically, whilst these models are physically-based, they commonly depend on model calibration. The information demands of model calibration, arising in complex models, may significantly exceed the information content of available data (Young et al., 1996; Heathwaite, 2003). This becomes more acute at large spatial scales: the poor availability of calibration data does not allow unambiguous estimation of the spatial distribution of key unknowns (e.g. soil depth); and many different model realisations may yield similar levels of model success (i.e. equifinality, Beven, 1989). This problem can be addressed in two ways. The first couples conventional predictive models with differing levels of process complexity for different scales (e.g. Quinn, 2004). Commonly, the finer the scale, the more complex is the process resolution. Models applied at the fine-scale are applied over smaller spatial units and for shorter time periods. This information is then transferred to coarser scales and longer time periods using generalisation tools in which each fine scale treatment is representative of other locations in the landscape (Quinn, 2004). Such an approach remains dependent upon suitable calibration data but also requires appropriate rules for generalising across scales. The second, adopted in this paper, uses a risk-based analysis in which the fine scale representation is applied to all locations in the landscape and then integrated up to the particular scale of enquiry.

Risk based analysis of this kind for diffuse pollution well-established (e.g. Jordan et al.¸1994; Johnes, 1996; Johnes and Heathwaite, 1997; Heathwaite, 2003; Heathwaite et al., 2003a, b; Jordan and Smith, 1994; Munafo et al., 2005). Early applications of these methods focused upon determining the export of fine sediment and nutrients associated with particular land covers (e.g. Johnes, 1996), but gave much less attention to delivery, the process by which material produced at a location in the landscape is transported to the stream network. Herein, we focus upon incorporating a treatment of delivery into a risk-based analysis in the form of a single modelling framework, SCIMAP (Sensitive Catchment Integrated Modelling and Analysis Platform), applied across spatial scales, and with reference to salmonid fry. We take as our first premise the fundamental property of catchments: they can be conceived as a set of flow paths that accumulate distributed sources of pollutants from across the landscape into the river corridor, where diffuse pollution may become visible, either during routine monitoring or through the occurrence of water quality problems (e.g. eutrophication). Given an observed downstream problem, and provided this can be attributed to diffuse sources, the primary challenge is to determine which parts of the landscape are most likely to be contributing to that problem. Our analysis is relative, in that we aim to judge the riskiness of one location in the landscape for locations in the downstream water environment as compared with all other locations in the landscape. This is what export coefficient models do implicitly, but they differ because they aim to translate their estimates of relative risks into absolute loadings to water courses (e.g. Johnes, 1996; Johnes and Heathwaite, 1997). Subject to data availability, the analysis can be run for any size of catchment, and predictions are made for all landscape locations relative to each other upstream of the catchment outlet chosen. By starting at the coarse scale, and running the model for progressively smaller spatial units (sub-catchment, tributary, stream, field) it allows successive identification of the sub-catchments that merit prioritisation, followed by the tributaries within a prioritised sub-catchment, the streams within a prioritised tributary and finally the fields connecting to a prioritised stream.

The second premise is that analyses like these need to more carefully consider how to incorporate an assessment of delivery, which matters in both a physical and a biochemical sense. In physical terms, the ease of hydrological connection will control the delivery of both conservative (e.g. particulate) and non-conservative (e.g. nutrient) parameters. In biochemical terms, the type of connection (e.g. overland flow, versus pipeflow versus matrix flow) will determine the nature of the biochemical transformations that result. To date, most approaches have specified connectivity in terms of simple landscape attributes. For instance, Johnes and Heathwaite (1997) used a simple distance-decay function to model the impact of land cover change on nutrient concentrations in streams draining the Slapton catchment, southwest England. Gburek et al. (2000) used a conceptual approach parameterised with empirical data based upon a contributing area function for different stream reaches. They found that the risk of phosphorus (P) delivery in surface runoff decreased with increasing distance from the stream, reflecting spatial variation in saturation-excess surface runoff: a rising stream water level in response to storm flow resulted in a rising water table in near-stream zones, so initiating P transport in surface runoff. The initiation of surface runoff from areas some distance from the stream required large magnitude long return period storms with the probability of surface runoff generation is low in such areas. Childress *et al*. (2002) defined relative connection in inverse proportion to the downslope distance from a given land unit to the drainage network.

The third premise is a challenge to classical approaches to modelling the impacts of diffuse pollution which tend to follow a hydrological process cascade (Lane, 2008). They begin by identifying the cascade of processes that might lead to a particular impact (e.g. soil erosion, leading to fine sediment delivery, leading to siltation of salmonid redds, leading to problems of fry emergence) and then break these down into the processes that need to be modelled (e.g. rainfall, evapotranspiration, infiltration, runoff generation, soil erosion, instream sedimentation). The cascade of processes leads to a hierarchy and the emphasis on the parts that make up this cascade make it reductionist. Furthermore, it is often unknown which aspects of the cascade matter to the impacted organisms. The focus of modelling is upon what is perceived to matter to an organism, sometimes supported by field or laboratory evidence, and despite conflicting ecological evidence over what might matter. Thus, herein, we use and inverse analysis based upon Bayesian methods. We include in the model the most rudimentary representations of processes that we think are sufficient: a treatment of erosion of material from the land surface; and a measure of the likelihood that eroded material can reach a river. We then use spatially distributed ecological data, in this case for salmonid fry, to determine which connected land covers seem to explain the ecological patterns.

Model development and application

Formulation of the model requires: (1) determination of the generation risk (p_i^g), here for material that can be eroded; (2) determination of the delivery index, or connection probability (p^c) for that eroded material; (3) convolution of (1) and (2) to get the locational risk (p^{gc}); (4) routing of the locational risk to determine a risk loading (L_j); and (5) transformation of the risk loading to a risk concentration (C_j). An overview of the processing steps for the generation of the risk map is shown in Figure 3.

Generation Risk

The focus of this paper is a formulation of our modelling approach for risks that need to be eroded, such as fine sediment, rather than risks that are dissolved in water. The generation risk for material that must be eroded will be determined by: (i) the energy available for erosion (the hydrological risk); and (2) the resistance to erosion or erodibility, which is used to weight the hydrological risk. Thus, we define p_i^g as the product of the risk of there being sufficient energy available to erode (p_i^h) and the risk of the material on the surface being erodible (p_i^e):

$$p_i^g = p_i^h \cdot p_i^e \quad [1]$$

The energy available to erode is assumed to be positively correlated with: (1) the area draining through a point in the landscape per unit contour length (which will determine the depth of water and hence contribute to soil erosion potential), A_i; and (2) the local slope, β_i; as represented by a stream power index (Ω_i):

$$\Omega_i = A_i \tan \beta_i \quad [2]$$

This index is linearly scaled to give a hydrological risk of erosion between the largest 5% and smallest 5% of values defined by [2], and this defines p_i^h. Determination of p_i^h requires the use of the topographic data to determine the upslope area and local slope in [2].

In this paper, we present two methods for estimating p_i^e in this example, we do two things. First, we develop a logical approach where we use intuitive argument to estimate the effects of land cover on erodibility: (1) erodibility might be expected to be negligible or zero under woodland cover (p_i^e=0.00); (2) it might rise slightly under moorland (p_i^e=0.05); (3) rise further under extensive pasture (p_i^e=0.10); (4) rise again under intensive or improved pasture (p_i^e=0.20); and (5) rise significantly, to the maximum risk for any land cover (e.g. arable) where the land cover might be bare for part of the year (p_i^e=1.00). It should be noted that an emphasis upon land cover may be warranted given that land cover is commonly correlated with soil type which also influences the erodibility. This approach then allows available land cover to be mapped onto p_i^e. However, it contains an implicit assumption that what matters is the erodibility of material, as conditioned by land cover. This may be relevant to an instream organism such as a salmonid, but the same amount of eroded material from disprate land covers may impact salmonids differently if the chemicals transported with the material are driving the degradation. So, second, we invert the problem and use a Bayesian approach to to identify the values of p_i^e that best reproduce the spatial structure of distributed salmonid fry counts: i.e. we make no *a priori* assumptions about the hydrological risk of erosion according to probable surface erodibility.

Delivery Index for eroded material

Our treatment of delivery has two primary assumptions. First, conceptually, connectivity within a landscape can be viewed over a range of spatial and temporal scales. At a point in time, there will be a binary relationship between two points in space, either there is currently a connection between the two points or there is not. As the temporal scale is increased, there will be a distribution of connection durations which gives information on the frequency and length of the connected periods for each point in the landscape with the receiving waters. The shape of this distribution will be governed by interaction between the temporal structure of storm events, both within storms and between separate storms events, and the structure of the landscape. It will also determine the amount of material that will reach the channel (Reaney et al., 2007) and the transformations that the material will undergo during transport. The primary assumption in our analysis is that these temporal distributions will be spatially structured, leading to a variable connection strength across the catchment. If we can find a reliable description of this spatial structure then we can use it to determine the likelihood that generated material is delivered to the drainage network. The nature of the required description will be dependent upon the type of material that is being delivered. For eroded material, the description must recognise that since eroded material is predominantly transported by overland flow, all of the flow path must be generating overland flow and this flow must be towards the drainage network for its entire length, in order for there to be connection. If a point on the flow path is not generating such flow, the water will infiltrate at that point and the eroded material will be deposited leading to the disconnection of the upper part of the slope (e.g. Figure 1). Thus, the point along a given flow path that is least likely to generate overland flow becomes the controlling location for the connection of all points upstream.

Our second assumption is that the topographic wetness index (Beven and Kirkby, 1979) can be used to describe the propensity to generate saturation excess overland flow for each point in the landscape: the higher the wetness, the greater the propensity for overland flow generation. The topographic wetness index expresses the propensity to saturation as the ratio of the upslope area per unit contour length draining through a point in the landscape and the tangent of the local slope, the latter assumed to represent the hydraulic gradient. Lane et al. (2004) show that the propensity to surface hydrological connectivity can then be described by the lowest value of the topographic wetness index along a flow path: the network index. Lane et al. (2009) show that the network index is effectively a measure of the propensity to vertical as opposed to lateral flow. In a system where delivery is dominated by surface or shallow subsurface flow, vertical flow reduces the propensity to disconnection.

Following the assumption made above, we make an ergodic hypothesis and assume that the network index implicitly contains a temporal dimension, one that applies equally to water, as it does to the material transported by that water. As the landscape wets up, more of the landscape will become connected as points that were previously disconnected areas start to generate and transmit runoff and hence connect their upslope areas to the river channel. The reverse will happen during drying. Thus, a point with a higher network index is less likely to be disconnected from the drainage network and hence is more likely to be connected for a longer duration. Under the assumption of a topographic control on overland flow generation, the major challenge is how to map the network index onto the duration of connection, the latter expressed as a probability. Here, we assume that the mapping between network index and duration of connection is linear between the largest 5% of values of the network index (always connected, i.e. connection probability at location *i* p_i^c= 1) and the smallest 5% of values of the network index (never connected, i.e. connection probability p_i^c= 0). The connection probability is taken as our delivery index for eroded material. We have tested this for a small catchment (52.1 km^2) by comparison to a physically-based distributed hydrological model and have shown that our delivery index contains a significant amount of information in relation to both the probability and duration of hydrological connection in upland environments with shallow soils (Lane et al., 2009).

Locational risk

We now combine the generation and delivery risks to determine the locational risk of delivery of generated material to the drainage network (p_i^{gc}):

$$p_i^{gc} = p_i^g \cdot p_i^c \qquad [3]$$

Routing, Accumulating and Dilution of locational risk

We route and accumulate the locational risk under the assumption that this is controlled by the topographically-driven accumulating area: i.e. the risk at a point is the sum of all locational risks upstream of that point. This leads to the risk loading to a point in the drainage network (L_j) with *j* upslope contributing cells which will increase monotonically with distance down through the drainage network:

$$L_j = \sum_{i=1}^{j} p_i^g \cdot p_i^c$$

[4]

The risk loading takes no account of: (1) the propensity for dilution, where a high loading from a small upstream contributing area will have a more serious environmental effect that a high loading from a high upstream contributing area; or (2) loss of risk (e.g. due to deposition or chemical transformation). In this paper, we assume that although deposition results in the local degradation of habitat, for fine sediment this deposition is relatively small as compared to that which is delivered, meaning that there is no need to correct for the loss of risk. This assumption is commonly made in sediment delivery models for large river basins (e.g. Naden and Cooper, 1999) and is supported by sediment budget studies. For instance, Owens *et al.* (1999) showed that only 4% of the fine sediment delivered was deposited in the bed of the River Tweed, although the loss may be greater once the transition of gravel to sand has passed (as in lowland systems, e.g. Collins and Walling, 2007). As the focus of this work is habitat where the bed is still predominantly gravel, we believe this is an acceptable assumption to make. However, as Figure 3 shows, dilution is a property of drainage networks that cannot be overlooked. The simplest way to deal with dilution is to scale the loading by the upslope contributing area to give a risk loading per unit area, akin to a concentration (C_j);

$$C_j = \frac{\sum_{i=1}^{j} p_i^g \cdot p_i^c}{\sum_{i=1}^{j} a_i \cdot r_i}$$

[5]

where: a_i is the cell size and r_i is the rainfall weighting factor. This equation takes account of possible rainfall variations between sub catchments and the propensity for such variation will increase with basin size. This is represented by weighting upslope contributing areas by the amount of upstream contributed precipitation, using temporal averages that reflect the time-integration of the study. However, such an analysis is complicated by the fact that spatial variability in precipitation should also result in spatial variability in connectivity. Hence, the predicted relative long term average wetness, calculated using the topographic wetness index, also utilises the rainfall weighting factor.

Bayesian Analysis with Respect to Generation Risk

In the model described above, there will be uncertainty associated with: (1) the determination of the hydrological risk of erosion; (2) the relationship between land cover and soil erodibility; (3) the relationship between topographic data uncertainty and the network index; (4) the scaling between the network index and the delivery index; (5) the impacts of topographic uncertainty upon flow paths and hence both flow and risk accumulation; (6) the simple manner in which the risk loading is transformed into a risk concentration using the rainfall weighted upslope contributing area; and (7) the meaning of our estimates of risk to instream organisms. In our definition of risk, high levels of generation and high levels of connection are assumed to produce a higher level of risk. However, it is possible that for a particular organism, this risk may be good or bad. For instance, O'Grady (2003) notes that Atlantic salmon are particularly dependent upon invertebrates supported by autochthonous production of organic matter, something that is more likely to be sustained by higher levels of nutrient input. Brown trout will also feed on invertebrates supported by allocthonous production of organic matter. Thus, a land cover such as deciduous woodland that simultaneously produces organic matter whilst reducing the erosion of sediment-bound nutrients such as phosphorus, will produce 'good' risk for trout and 'bad' risk for Atlantic salmon. This is just one example of the confounding influence of land use on the net ecosystem productivity, which manifests itself, in our case, as uncertainty as to exactly what our risk estimates mean for instream organisms. In addition there will be interaction between these uncertainties, notably arising from the time integration and the possibility that particular land management practices are coincident with particular periods of higher or lower connectivity. The causes of uncertainties (1), (3) through (6) and the interaction uncertainties are well known and hence potentially modelled. However, the uncertainties associated with (2) and (7) are more acute and so we compliment our use of logical reasoning with a Bayesian method in which we infer erosion weights, and hence our risk estimates, with reference to the spatially-distributed data described above, a form of inverse modelling (Lane, 2008).

The Bayesian method that we use is effectively a likelihood estimation procedure (e.g. Beven and Binley, 1992) in which we infer the range of plausible p_i^e values in a Monte Carlo sampling framework from the spatially-distributed fry data that we have. This approach reduces the uncertainty that derives from the imposition of our own assumptions as to what constitutes a 'good' or 'bad' land cover, as happens with our logical estimates. We undertook 30,000 model simulations, randomly selecting values in the range $0 \le p_i^e \le 1$ for each land cover for each simulation. We then determined the values of an objective function, to describe the level of association between risk estimates and fry abundance appropriate to the nature of the abundance data (Equation 7), for each simulation. Simulations were ranked according to the estimated value of the objective function for each simulation from most likely to least likely. Starting with the *x* most likely simulations, we determined the mean and standard deviation of the parameter values associated with those *x* simulations, then progressively increased *x*, each time recalculating the mean and the standard deviation. The maximum value of *x* considered was 200. The plot of rank against mean and standard deviation allows us to see two things: (1) those land covers whose p_i^e values matter will be characterised by a narrow range of values, or a small standard deviation; and (2) whether or not the p_i^e values themselves match our *a priori* expectations of the relative erodibility associated with different land covers. In the final stage of analysis, for each value of *x*, we compared the mean and standard deviation of parameter values for each value of *x* greater than 10 with the mean and standard deviation of parameter values for the 30,000 simulations. Where the difference was significant at the 95% level (Student's t, two-tailed) we concluded that weighting of the source risk by a particular land cover mattered.

Case study and data sources

In this paper, we develop and assess the SCIMAP (Sensitive Catchment Integrated Modelling and Planning) framework for the 2310 km^2 River Eden catchment in northern England (Figure 4). The Eden catchment comprises a range of land covers, with four dominant: arable; intensive or improved pasture; extensive pasture; and moorland. There are a range of physical, ecological and topographic conditions, with the geology ranging from sandstone to limestone and instream ecological conditions ranging from oligotrophic to mesotrophic (Parsons et al. 2001). The large spatial area, variety of land cover classes and differing ecological environments make the Eden catchment a useful place to test the prediction of the risk model.

In the form that the risk management framework is applied here, it requires the following data sources: (1) topographic data of appropriate spatial resolution and vertical precision; (2) land cover data; and (3) rainfall data to determine potential dilution effects. For the topographic data, we use Interferometric Synthetic Aperture Radar data produced by InterMap. This comprises 5 m resolution digital terrain data (i.e. after trees and buildings have been removed) that is estimated to be precise to ± 1.0m. These data required pre-processing in two steps. First, although topographic depressions or pits in the landscape are genuine features, they are often errors in the DEM surface. In particular, the determination of the upslope contributing area and flow routing requires information on the direction in which the water would flow once a pit has been flooded. Therefore, to calculate the upslope contributing area the pits in the digital elevation model (DEM) are filled using the Planchon and Darboux (2001) algorithm. This pre-processing is only utilised for the calculation of the upslope contributing area and other terrain derivates, such as slope, are calculated with the original DEM. This processing enables the capturing of the impact of depressions and their role in landscape disconnection while accurately representing the upslope contributing area. The second step is the calculation of the upslope contributing area for each point in the landscape. We use the D-infinity ($D\infty$) method of Tarboton (1997). This method utilises multiple flow paths to give an accurate representation of the flow direction and avoids the straight line artefacts of simpler approaches such as D8 (Gallant and Wilson 1996).

We use the land cover map of Great Britain for 2000 (Centre for Ecology and Hydrology, 2000) to give land cover estimates at a 30 m resolution, which were then interpolated from 30 m onto the finer scale data using a nearest neighbour algorithm. As the land cover dataset is synoptic and dated, it is probable that the actual nature of land management is misrepresented. However, as data from the annual Agricultural Census remains confidential at resolutions higher than the parish scale, this was judged as the best alternative.

The spatial pattern of rainfall is derived from the UK Metrological Office long term annual average rainfall dataset (Perry and Hollis 2005). This dataset was then interpolated onto the topographic data resolution used herein via a nearest neighbour algorithm.

We use salmonid fry data from the Eden Rivers Trust, who have conducted catchment-wide surveys of brown trout and Atlantic salmon fry (0+ year class / subyearling) for the years 2002 – 2005 at 200-300 sites per year (Maltby, 2002, Townsend-Cartwright, 2004, Dickson, 2004, Dickson et al., 2005). These surveys have been carried out using semi-quantitative electrofishing following the approach of Crozier and Kennedy (1994). The method is semi-quantitative as it is based upon single-pass, rather than repeat-pass electrofishing. The focus on salmon fry is based upon observations that suggest that for these species, fry tend to remain within 100s of metres of their spawning site. For example, Einum and Nislow (2005) observed fry to remain within 644m and 884m downstream, and 1,500m and 642m upstream of their redd in two years of observations respectively, with the median dispersal range being 92m and 41m in Year 1 and Year 2 respectively. Similarly, Kennedy (1982, cited in Crisp, 1996) found over 70% of fry to be within 100m downstream of their stocking point. There are always exceptions and Beall et al., (1994, cited in Crisp, 1996) found that salmon had dispersed over 2000m downstream, with a substantial number moving between 1000m and 1500m by the October of their first year. It has been suggested that dispersal of fry is constrained by: (1) energetic costs and the lack of feeding opportunities during dispersal, which may lead to starvation; and (2) increased exposure to predators (Einum and Nislow, 2005). This lack of mobility within the first few weeks and months of life means that fry are highly susceptible to density-dependent mortality and, as a result, to local habitat conditions which regulate the local carrying-capacity. Parr and adult populations can be more mobile, with dispersal ranging from 10s m-1000s m. Dispersal downstream is typically greater than the degree of dispersal upstream. Some dominant fish may aggressively defend a small localised territory which is profitable for food, whilst other, more subordinate fish may be more mobile and 'float' between territories (Suter and Huntingford, 2002).

These observations inform the sampling strategy adopted in the Eden Rivers Trust sampling strategy. Typically, these species spawn in the Northern hemisphere winter and the alevin emerge in late winter / early spring. Thus, the semi-quantitative electrofishing focused upon late spring and summer, corresponding to the fry life stage. Here, we use data from 2002 and 2003. In both years, sampling was stratified to sites where trout fry were expected to be found (suitable riffle habitat). Suitable trout fry locations were then sampled randomly, with no repetition. Catch efficiency was recorded at each site. This was defined as the number of caught fish as a percentage of the total number of fish that could be caught (i.e. caught, plus visually observed but not caught) (Hilborn and Walters, 1992). To focus on sites where we had a reasonable confidence in the salmonid fry estimates, we followed Crozier and Kennedy's (1984) recommendation and only sites where catch efficiency exceeded 60% were used in all of the subsequent analysis. Thus, 17.5% of sites were rejected by this criterion in 2002 and 17.6% in 2003. This has the potential to introduce some bias into our dataset as the sites that were excluded were ones where fish were present, but unidentifiable. However, the sites with low catch efficiency were randomly distributed in space and the proportions were relatively small. The two years of data were not pooled as interannual variability in catchment scale abundance did occur, which we assume is related to catchment-scale exogenous factors (such as broad climatic variability). All samples were transformed to abundance, defined as the number of fry sampled per 5 minute interval. We checked for spatial autocorrelation in our analysis. Ideally, each sample location should contain a brown trout population independent of all other locations: i.e. it should be associated with a distinct set of redds which may or may not have been impacted upon by the upstream catchment characteristics. This independence was easier to achieve given the restricted range of dispersal of brown trout. However, to check for this independence, we calculated the distance between adjacent sites and thresholded this to determine an index of the number of non-independent sites. We could not identify significant spatial autocorrelation in the datasets provided sites were set at > 500 m apart. Where this was not the case, one of the two sites causing the spatial autocorrelation was randomly sampled and removed. Finally, a subset of semiquantitative sites were compared with the results of fully quantitative fishing as a final check on the quality of the semiquantitative estimates.

The focus on salmonid fry is valuable in this application for three reasons. First, salmonid fry are the least mobile life stage and supporting the assumption that their abundance is strongly influenced by exposure to local conditions, and hence contain a spatial signal that reflects the spatial variability in those local conditions. Second, the recruitment of salmonid fry is though to be impacted by processes that are associated with surface transport of eroded material (e.g. fine sediment, pollutants bound to fine sediment). For instance, low fry counts have been attributed to fine sediment infiltration into spawning gravels (Soulsby et al., 2001; Greig et al., 2005), which reduces the supply of the oxygen required for effective incubation and alevin emergence. There are two main disadvantages of using such validation data in this application: the abundance of fry will reflect other potential limits upon fry recruitment. Therefore the data may contain a spatial signal that is driven by other factors and hence be noisy with respect to the fine sediment signal that is a focus of this research. The second is that the method is only semi-quantitative: in order to allow a larger number of sites to be sampled, it is based upon single pass rather than multiple pass. In order to assess the general reliability of the semi-quantitative data, fully quantitative electro-fishing was undertaken for a subsample of sites in each sample year.

Comparison of risk predictions with the fry abundance data considered risk predictions classified by the presence/absence of fry but also the relationship between fry abundance and associated predicted risk. For the presence/absence analysis, generated risk predictions were extracted at sites in the river Eden catchment with and without fry present. Mann Whitney tests were used to examine whether there was a statistically significant difference in the central tendency of the risk predictions. Two sample Kolmogorov-Smirnov tests were also used to test whether there was a difference in the distribution (location and shape) of the risk predictions between sites with and without salmonid fry. Nonparametric tests were selected due to the non-normal distribution of the data which could not be corrected for using standard transformation procedures.

In order to understand the relationship between individual fry abundance values and SCIMAP predictions of risk concentration we developed an objective function. Semi-quantitative fry abundance data are commonly severely skewed, with many sites have low or zero abundance and a small number of sites having a high abundance. As a result, most descriptors of fry abundance use classifications. Here, we adopt the classification of Crozier and Kennedy (1994) which has five classes: Excellent; Very Good; Average; Poor; and Very Poor. We use these data in two ways. First, we classify the predicted risk concentration values for each fry sampling site in each year into five categories: very low risk); low risk; average risk; high risk; and very high risk. We use this to determine a contingency tabulation of risk concentration class (*p*) with fry abundance class (*q*). This is used to determine an objective function based upon the orthogonal distance (d_{pq}) of each combination of risk concentration class and fry class from the diagonal of equality (*p*=*q*), weighted by the number of sites classed into *pq*:

$$OF = \frac{\sum_{q=1}^{5}\sum_{p=1}^{5} d_{pq} n_{pq}}{\sum_{q=1}^{5}\sum_{p=1}^{5} n_{pq}} = \frac{\sum_{q=1}^{5}\sum_{p=1}^{5} 0.5\left[2(p-q)^2\right]^{0.5} n_{pq}}{\sum_{q=1}^{5}\sum_{p=1}^{5} n_{pq}} \qquad [6]$$

Using [6], and with the five classes used here, a perfect level of agreement should be with an $OF = 0$ and a perfect level of disagreement should be with an $OF = 2^{1.5}$. Thus, we rescale [6] to vary from 1 (perfect agreement) to 0 (perfect disagreement):

$$OF = 1 - \frac{\sum_{q=1}^{5}\sum_{p=1}^{5} 0.5\left[2(p-q)^2\right]^{0.5} n_{pq}}{2^{1.5}\sum_{q=1}^{5}\sum_{p=1}^{5} n_{pq}} \qquad [7]$$

The objective functions were used in three ways. First, they were used in a conventional validation, in which the reasoned erodibility weights were used in the risk estimation. Second, they were used for the Bayesian modelling as a means of determining the values of the erodibility weights on each land cover required to maximise the objective function. The appeal of this approach is that it makes no prior assumptions about the land cover weightings required in relation to salmonid populations. Third, the OF [7] was applied to test the level of agreement between the quantitative and the semi-quantitative fry data. The OF value for salmon fry was 0.874 and that for trout fry was 0.842. In contingency terms, the accuracy of the comparison (i.e. the percentage of sites for which $p = q$) was 66.7% for salmon and 61.3% for trout. The accuracy measure does not take into account the percentage of sites for which p would equal q under random sampling of all p,q combinations. Given the class memberships of salmon and trout, we would expect the percentage level of agreement under random assignment to be 14.5% and 16.8% for salmon and trout respectively. Thus, we conclude that the semi-quantitative data contain a significant signal for the purposes of this analysis. We note that the OF values calculated from [7] (i.e. 0.874 and 0.872) represent the upper limits of the possible OF values that might be achievable during inverse modelling due to uncertainties in the semi-quantitative data. Only semi-quantitative data are used in the subsequent analysis.

Results

Model application using logical erodibility weights

Error! Reference source not found. shows the main derivates of the DEM that are used in the calculation of the point scale soil erodibility risk. **Error! Reference source not found.**a shows the network index which is used to determine the surface flow connection risk. The map shows that the areas of the Eden catchment that are predicted to be the most highly connected are located in the western section of the lowlands. The Pennine hillslopes on the eastern side of the catchment and the Lake District hillslopes in the south west are predicted to be the least connected areas. **Error! Reference source not found.**b shows the predicted spatial pattern of the soil erodibility as determined by the surface land cover. This map shows that the greatest risk of soil erosion occurs in the main lowland plain of the Eden catchment, especially towards the north-west. **Error! Reference source not found.**c shows the distribution of the stream power index which represents the energy available to erode the surface. The highest values of the stream power index occur on the steep slopes located both on the Pennine and Lake District hillslopes. In the areas of the catchment which have high soil erodibility risk (**Error! Reference source not found.**b) tend to be negatively correlated with the high stream power index (**Error! Reference source not found.**c).

Figure 5d and Figure 5e show the convolution of the source area analysis with the connectivity analysis. The effect of this calculation (Figure 5d) is to highlight areas predominantly in the catchment headwaters and certain areas in the main valley as being at risk of erosion due to hydrological processes. This does not take into account the mitigating effect of land cover, which is introduced via the erodibility treatment in Figure 5e. This removes much of the area classified as having a high risk of erosion and focuses the risk in the main valley of the catchment, largely where arable land is well-connected to the drainage network. Note the implication here is both curative and preventative. An environmental restoration strategy aiming to reduce risk (i.e., curative) would focus upon Figure 5e. However, strategies aimed at preventing further environmental degradation would evaluate the locations in Figure 5d are where a change land management activities should be evaluated carefully in order to prevent future environmental degradation..

In the final stage of the analysis, we integrate through to the drainage network. **Error! Reference source not found.** shows the accumulated risk weighted by the dilution potential, essentially the risk concentration. If the convolution of connectivity and locational risk were everywhere uniform, and the rainfall field homogeneous, then the risk should accrue linearly as the potential for dilution accrues. Where the risk is some multiple of the standard deviation greater than the mean, then the risk is increasing disproportionately faster than the increase in dilution potential: i.e. a particular risky input to the drainage network has been identified. Where the risk is some multiple of the standard deviation less than the mean, the risk is increasing disproportionately more slowly than the increase in dilution potential and the drainage network is benefiting from low risk inputs. If we consider hydrological risk without land cover weighting (**Error! Reference source not

found.a), this would focus diffuse pollution mitigation activities on the catchment headwaters, where the risk of erosion and the risk of connectivity combine to cause an accumulated risk that is not well balanced by accumulating dilution potential. These are sensitive areas of the catchment where land covers or land use practices that allow more erosion (e.g. temporary change in land cover due to heather burning) might have a major water quality impact due to increased export of fine sediment. However, **Error! Reference source not found.**b shows that when the land cover weighting is introduced, concern switches to areas lower in the drainage basin, where the presence of arable cropping, and hence the risk of the land surface being bare, results in a different identification of risky sub catchments.

Model testing using logical risk estimates

Table 1 shows that the risk predictions discriminate extremely effectively between the presence and absence of trout fry in both 2002 and 2003, and the Mann Whitney and Kolmogorov-Smirnov tests reveal that these differences are statistically significant ($p<0.05$). The risk predictions are less effective at discriminating between sites where salmon fry are present and absent. Differences in the mean risk are slight and Mann Whitney tests are not significant ($p>0.05$) in both years. However, the Kolmogorov-Smirnov tests for sites with and without salmon fry do report a statistically significant difference in the distribution of risk predictions for all years excepting 2005. This suggests that there may still be a relationship between risk predictions and salmon fry but that it is non-linear. To investigate the ecological significance of the risk predictions further, salmonid abundance data were also examined. The risk predictions were classified into five equal membership risk classes ranging from 1 (the 20% of sample locations with the lowest risk) and 5 (the 20% of sample locations with the highest risk).

Model application using the Bayesian-based inverse modelling

Figure 7 shows the results of the inverse modelling for the years 2002 (Figure 7a and Figure 7c) and 2003 (Figure 7b and Figure 7d) for both trout (Figure 7a and Figure 7b) and salmon (Figure

7c and Figure 7d). In these plots, the solid line plots the mean erodibility against the associated objective function value. As the objective function value is defined as a perfect association with a value of 1.0, the mean weighting for the best 10 simulations is the farthest right of each plot and the mean weighting for the best 200 simulations is the farthest left on each plot. The dashed line is the standard deviation of the erodibilities assigned to each land cover. These results can be interpreted as follows. First, as we randomly sampled erodibility weights (p_i^g) between zero and one, the mean of all 30,000 simulations was 0.50 and the standard deviation ±0.29. The arable land cover plots in Figure 7 have mean and standard deviations almost identical to these values for almost all values of the objective function. Thus, for this analysis, the erodibility weighting given to arable land cover in the risk analysis could be sampled randomly between 0 and 1, and there would be no impact upon the value of the objective function achieved. Thus, the extent to which there are arable cover classes upstream of a fry sample point does not seem to explain the spatial variability in fry populations. Second, the results are very different for improved pasture. Here, the standard deviations widen as the magnitude of the objective function (the level of association between salmonid fry abundance and the predicted risk) is reduced, suggesting that a narrower range of erodibility weights needs to be applied to the improved pasture to optimise the association between risk and fry. Similarly, the erodibility weights that produce higher levels of association between predicted risk and salmonid populations are generally greater than 0.75, although in two cases (Figures 7a and 7c), the weightings trend towards 0.5 at lower values of the objective function. For the 200 best simulations shown in the improved pasture plots in Figure 7, the mean erodibility is still statistically distinguishable from 0.5 (at $p = 0.05$). Thus, to get good levels of agreement with observed data, improved pasture should not be assigned randomly. Following from the ordinal form of the objective functions ([7]), this can be interpreted symmetrically: the extent to which there are improved pasture land uses, as filtered by the propensity to connect in the risk analysis, upstream of a fry sample point is important in explaining the spatial variability in fry populations; or where improved pasture is hydrologically connected, it produces higher risk, in relation to the spatial variability in fry populations. The reverse is true for extensive grazing: in all cases, well-connected extensive grazing sites produce low risk from the perspective of fry populations. The third, and perhaps most interesting observation relates to moorland. For trout, in both 2002 and 2003, the required erodibility weighting must be low: the standard deviation increases rapidly as the values of the objective function fall (Figure 7a and 7b) suggesting that only low weightings of moorland give the best values of the objective function. For salmon, also in both 2002 and 2003, the required land cover risk weighting must be high, and again, the weighting seems to matter, albeit only for the best 150 or so simulations (e.g. values of the objective function less than c. 0.575 in Figure 7d). There appears to be a functional difference between what salmon fry and trout fry view as creating risk: well-connected moorland is risky for salmon but not for trout.

Discussion and conclusions

The primary findings of this work are four fold. First, combination of hydrological connection and erodibility into a risk model produces patterns of risk that vary spatially in ways that distinguish both salmonid and trout fry populations. Second, there are functional differences between salmon and trout as to what constitutes a risky land cover. The ecological interpretation of these differences is currently under way. For instance, it is well-known that trout can display strongly territorial behaviour (e.g. Elliot, 1994). If trout have a preference for low levels of risk associated with a combination of poor hydrological connection and large amounts of upstream moorland, then this may be to the exclusion of salmonid fry in those locations. Thus, the higher optimised erodibility weight of moorland for salmon may not be because moorland is bad, but simply sites with large amounts of well-connected moorland upstream are preferential to trout.

Third, and related to this moorland weighting finding, the work demonstrates the importance of inverse modelling (Lane, 2008), certainly in studies of fish populations, but also diffuse pollution more widely. The logical assignments of erodibility originally identified were very different to those identified using inverse modelling. Notably, we identified that the erodibility weightings assigned to arable land covers were unimportant, that those associated with improved pasture needed to be much higher (~0.75) than expected (0.2) and that those assigned to moorland were species dependent. It is worth reflecting on why our logical assignments were incorrect. This was largely because we had a perceptual model that emphasised the delivery of fine sediment to the water course as being instrumental to habitat degradation. Hence, we weighted our erosion potential by a land cover defined erodibility ([1]). There is good support for this, as whilst female salmon are commonly able to clear out fines from river gravels during construction of a redd, high rates of delivery of fine sediment to the redd (e.g. Soulsby et al., 2001; Greig et al., 2005) can prevent fry emergence. However, this is only one of a number of possible impacts of fine sediment erosion. The inverse modelling results question at least two possible dimensions of the logical analysis: (i) that the association between erodibility and land cover was as we hypothesised (i.e. arable is the most erodible); and/or (ii) that the eroded material that causes habitat degradation is associated with particular land covers because it carries other potentially problematic material such as animal waste. The observation that improved pasture requires a higher erodibility risk factor may not reflect the fact that improved pasture is more erodible, but rather that material eroded from improved pasture carries risks that impact disproportionately upon salmonids. Although this is only one of a number of possible explanations for the weightings arrived at during inverse modelling, and the account of the importance of and dynamics of pesticides is only one part of a more complex process of pesticide behaviour, it emphasises the advantage of spatially rich data in guiding the model building process, notably that associated with what we perceive is important. We cannot be certain about the pesticide explanation but, in management terms, and notably in the context of a precautionary approach, the formulation of the risk analysis is advantageous. The high risks are not only caused by particular land covers but also the process of delivery. Thus, where the science is uncertain (exactly why improved pasture seems to be responsible for degradation of salmonid fry), and may not be sufficient to justify landscape scale changes in agricultural practice, our approach focuses analyses upon upon a sub-set of fields that are both potentially risky in terms of the generation of material (e.g. improved pasture) and hydrologically well-connected.

Fourth, this analysis is based upon demonstrating that hydrologically-connected risks impact the spatial structure of fry. This is likely to be only one factor, as part of a multitude of different factors, including local riparian scale factors, such as canopy cover and in stream barriers, that need to be considered. We are currently exploring combining the risk predictions with information on other possible causes of salmonid habitat degradation in a multivariate framework in order to address the relative importance of the methodological approach reported here.

This analysis does make a number of important hydrological assumptions regarding the ways in which the landscape mediates patterns of water flow and hence the transfer of material from the landscape to rivers and streams. In this case, our parameterisation of delivery, including the simple linear transformation of the Network Index into a delivery index, was tested upon a hydrologically-similar upland catchment by comparison with a distributed and physically-based hydrological model (Lane *et al.*, 2009). This does not take into account the effects of soil depth, vegetation cover etc. A more computationally intensive approach would use a time-dependent numerical model to simulate the actual percentage of time that a point in a catchment is saturated and able to export, and use this information to assign a relative risk to each location. However, such an approach requires use of additional hydrological theory. The simplicity of the risk-based approach becomes undermined as attention has to be given to the well-established problems of calibrating hydrological models. However, such an approach opens up the possibility of simulating how the relative risk of connection changes due to the effect of climate change upon rainfall and evapotranspiration, and hence upon the amount of time that each location is surface connected. In relation to the focus of this manuscript, salmonid fry, thinking through how to parameterise the delivery index may provide a means of exploring assumptions regarding the impacts of climate and landscape change upon fry habitat.

Acknowledgements

This paper is based upon work funded by NERC CONNECT grant NE/C508850/1 and Environment Agency award SC070014. The authors are grateful to a series of Eden Rivers Trust Fisheries Officers for managing the fisheries data collection and analysis. The paper benefited from exceptionally helpful and constructive comments from two anonymous reviewers.

References

Adams, R., Dunn, S.M., Lunn, R., Mackay, R. and O'Callaghan, J.R., 1995. Assessing the performance of the NELUP hyodrological models for river basin planning. *Journal of Environmental Planning and Management*, **38**, 53-76.

Allan, J.D., and Johnson, L.B. 1997. Catchment-scale analysis of aquatic ecosystems. *Freshwater Biol*ogy, **37**, 107–111.

Armstrong, J.D., Grant, J.W.A., Forsgren, H.L., Fausch, K.D., DeGraaf, R.M., Fleming, I.A., Prowse, T.D and Schlosser, I.J., 1998. The application of science to the management of Atlantic salmon (Salmo salar): integration across scales. *Canadian Journal of Fisheries and Aquatic Science*, **55**(Suppl. 1), 303–11.

Armstrong, J.D., Kemp, P.S., Kennedy, G.J.A., Ladle, M. and Milner, N.J., 2003. Habitat requirements of Atlantic salmon and brown trout in rivers and streams. *Fisheries Research*, **62**, 143-170.

Benda, L., Poff, N.L., Miller, D., Dunne, T., Reeves, G., Pess, G. and Pollock, M., 2004. The network dynamics hypothesis: how channel networks structure riverine habitats. *Bioscience*, **54**, 413-427.

Beven, K.J. and Binley, A., 1992. The Future of Distributed Models: Model Calibration and Uncertainty Prediction *Hydrological Processes*, **6**, 279-298

Beven, K.J. and Kirkby, M.J. 1979. A Physically Based, Variable Contributing Area Model of Basin Hydrology; *Hydrological Sciences Bulletin* **24**, 43-6

Beven, KJ. and Wood, E.F., 1983. Catchment geomorphology and the dynamics of runoff contributing areas. *Journal of Hydrology*, **65**, 139–158.

Beven, KJ., Heathwaite, AL., Haygarth, PM., Walling, DE., Brazier, RE. and Withers, P., 2005. On the concept of delivery of sediment and nutrients to stream channels. *Hydrological Processes*, **19**, 551-556.

Binder, C., Boumans, R.M. and Costanza, R., 2003. Applying the Patuxent Landscape Unit Model to human dominated ecosystems: the case of agriculture. *Ecological Modelling*, **159**, 161-77.

Bishop, K., Seibert, J., Kohler, S. and Laudon, H., 2004. Resolving the double paradox of rapidly mobilised old water with highly variable responses in runoff chemistry. *Hydrological Processes*, **18**, 185–189.

Blackwell, M.S.A., Hogan, D.V. and Maltby E., 1999. The use of conventionally and alternatively located buffer zones for the removal of nitrate from diffuse agricultural run-off. *Water Science and Technology*, **39**, 157-64.

Borah, D.K. and Bera, M., 2003. Watershed-scale hydrologic and nonpoint-source pollution models: Review of mathematical bases. *Transactions of the American Society of Agricultural Engineers*, **46**, 1553-66.

Borah, D.K. and Bera, M., 2004. Watershed-scale hydrologic and nonpoint-source pollution models: Review of applications. *Transactions of the American Society of Agricultural Engineers*, **47**, 789-803.

Bowes, M.J., Hilton, J., Irons. G.P. and Hornby, D.D. 2005: The relative contribution of sewage and diffuse phosphorus sources in the River Avon catchment, southern England: Implications for nutrient management. *Science of the Total Environment*, **344**, 67-81.

Burt, T.P. and Pinay, G., 2005. Linking hydrology and biogeochemistry in complex landscapes. *Progress in Physical Geography*, **29**, 297-316.

Burt, T.P., Matchett, L.S., Goulding, K.W.T., Webster, C.P. and Haycock, N.E., 1999. Denitrification in riparian buffer zones: the role of floodplain hydrology. *Hydrological Processes*, **13**, 1451-63.

Centre for Ecology and Hydrology, 2000. *Land Cover Map of Great Britain*, CEH, Wallingford, http://www.ceh.ac.uk/data/lcm/.

Collins, A.L. and Walling, D.E., 2007. Fine-grained bed sediment storage within the main channel systems of the Frome and Piddle catchments, Dorset, UK. *River Research and Applications*, **23**, 429-450.

Crisp, 1996. Environmental requirements of common riverine European salmonid fish species in fresh water with particular reference to physical and chemical aspects. *Hydrobiologia*, 323, 201-21.

Crozier, W.W. and Kennedy, G.J.A. 1994. Application of semi-quantitative electrofishing to juvenile salmonid stock surveys*. Journal of Fish Biology,* **45**, 159-164.

De Roo, A.P.J. and Jetten, V.G., 1999. Calibrating and validating the LISEM model for two data sets from the Netherlands and South Africa. *Catena*, **37**, 477–493.

Dickson, J. 2004. Semi-quantitative electrofishing survey in support of the development of catchment-scale restoration targeting – 2004. Eden Rivers Trust.

Dillon, P.J. and Molot, L.A., 1997. Effect of landscape form on export of dissolved organic carbon, iron and phosphorus from forested stream catchments. *Water Resources Research*, **33**, 2591–2600.

Durance, I., Lepichon, C., Ormerod, S.J., 2006. Recognizing the Importance of Scale in the Ecology and Management of Riverine Fish. *River Research and Applications,* **22**, 1143-1152.

Einum, S. and Nislow, K.H., 2005. Local-scale density-dependent survival of mobile organisms in continuous habitats: an experimental test using Atlantic salmon. Oecologia, 143, 203-10.

Ekholm, P., Turtola, E., Juha Grönroos, J., Seuri, P. and Ylivainio, K., 2005. Phosphorus loss from different farming systems estimated from soil surface phosphorus balance. *Agriculture, Ecosystems and Environment*, **110**, 266–278.

Gallant, J. C. and Wilson, J. P. 1996. TAPES-G: A grid based terrain analysis program for the environmental sciences. *Computers and Geosciences*, **22**, 713 - 722

Gburek, W.J., Sharpley, A.N., Heathwaite, A.L. and Folmar, G., 2000. Phosphorus management at the watershed scale. *Journal of Environmental* Quality, **29**, 130–144.

Greig, S.M., Sear, D.A., Smallman, D. and Carling, P.A., 2005. Impact of clay particles on the cutaneous exchange of oxygen across the chorion of Atlantic salmon eggs. *Journal of Fish Biology*, **66**, 1681-91.

Harding, J.S., Benfield, E.F., Bolstad, P.V., Helfman, G.S., and Jones, E.B.D., III. 1998. Stream biodiversity: The ghost of land use past. *Proc. Natl. Acad. Sci. U.S.A.*, **95**, 14,843-7.

Harris, G. and Heathwaite, A.L., 2005. Inadmissible evidence: knowledge and prediction in land and riverscapes. *Journal of Hydrology*, **304**, 3-19.

Haycock, N.E. and Muscutt, A.D., 1995. Landscape management strategies for the control of diffuse pollution. *Landscape and Urban Planning*, **31**, 313-21.

Heathwaite, A.L., 2003. Making process-based knowledge useable at the operational level: a framework for modelling diffuse pollution from agricultural land. *Environmental Modelling and Software*, **18**, 753–760.

Heathwaite, A.L., Fraser, A.I., Johnes, P.J., Hutchins, M., Lord, E. and Butterfield, D., 2003a. The phosphorus indicators tool: a simple model of diffuse P loss from agricultural land to water. *Soil Use and Management*, **19**, 1–11.

Heathwaite, A.L., Johnes, P.J. and Peters, N.E., 1996. Trends in nutrients. *Hydrological Processes*, **10**, 263–293.

Heathwaite, A.L., Quinn, P.F. and Hewett, C.J.M., 2005. Modelling and managing critical source areas of diffuse pollution from agricultural land using flow connectivity simulation. *Journal of Hydrology*, **304**, 446-61.

Heathwaite, A.L., Sharpley, A.N. and Gburek, W.J., 2000. A conceptual approach for integrating phosphorus and nitrogen management at catchment scales. *Journal of Environmental Quality*, **29**, 158–166.

Heathwaite, A.L., Sharpley, A.N., Beckmann, M. and Rekolainen, S., 2003b. Assessing the risk of agricultural nonpoint source phosphorus pollution. In: Sims, J.T., Sharpley, A.N. (Eds.), Phosphorus: Agriculture and the Environment. ASA/CSSA/SSSA Monograph.

Herrmann, S., Dabbert, S. and Raumerb, H.S., 2003. Threshold values for nature protection areas as indicators for bio-diversity—a regional evaluation of economic and ecological consequences. *Agriculture, Ecosystems and Environment*, **98**, 493–506.

Hooda, P.S., Moynagh, M., Svoboda, I.F., Thurlow, M., Stewart, M., Thomson, M. and Anderson, H.A., 1997. Soil and land use effects on phosphorus in six streams draining small agricultural catchments in Scotland. *Soil Use and Management* , **13**, 196–204.

Hunsaker, C.T. and Levine, D.A., 1995. Hierarchical approaches to the study of water-quality in rivers. *Bioscience*, **45**, 193-203.

Hynes, H.B.N., 1975. The Stream and its Valley. *Verhandlungen der Internationalen Vereinigung fuer Theoretische und Angewandte Limnologie*, **19**, 1-15.

Johnes, P.J. and Heathwaite, A.L., 1997. Modelling the impact on water quality of land use change in agricultural catchments. *Hydrological Processes*, **11**, 269–286.

Johnes, P.J., 1996. Evaluation and management of the impact of land use change on the nitrogen and phosphorus load delivered to surface waters: the export coefficient modelling approach. *Journal of Hydrology*, **183**, 323–349.

Johnson, L.B. and Gage, S.H., 1997. Landscape approaches to the analysis of aquatic ecosystems *Freshwater Biology*, **37**, 113–132.

Johnson, L.B., Richards, C., Host, G.E., and Arthur, J.W. 1997. Landscape influences on water chemistry in Midwestern stream ecosystems. *Freshwater Biology,* **37**, 193–208.

Jordan, C. and Smith, R.V., 2005. Methods to predict the agricultural contribution to catchment nitrate loads: designation of nitrate vulnerable zones in Northern Ireland. Journal of Hydrology, **304**, 316–329

Jordan, C., Mihalyvalvy, E., Garrett, M.K. and Smith, R.V., 1994. Modelling of nitrate leaching on a regional scale using a GIS. *Journal of Environmental Management*, **42**, 279–298.

Kirchner, J.W., Feng, X. and Neal, C., 2000. Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature*, **403**, 524-7.

Kuusemets, V. and Mander, U., 1999. Ecotechnological measures to control nutrient losses from catchments. *Water Science and Technology*, **40**, 195-202.

Lammert, M. and Allan, J.D., 1999. Assessing biotic integrity of streams: Effects of scale in measuring the influence of land use/cover and habitat structure on fish and macroinvertebrates. *Environmental Management*, **23**, 257-70.

Lane, S.N., 2008. What makes a fish (hydrologically) happy? A case for inverse modelling. *Hydrological Processes*, **22,** 22 pp 4493-4495

Lane, S.N., Brookes, C.J., Kirkby, M.J. and Holden, J. 2004. A network-index based version of TOPMODEL for use with high-resolution digital topographic data. *Hydrological Processes,* **18**, 191-201.

Lane, S.N., Morris. J., O'Connell. P.E. and Quinn, P.F., 2006. Managing the Rural Landscape. Chapter 17 in Thorne, C.R., Evans, E.P. and Penning-Rowsell, E. *Future Flood and Coastal Erosion Risk*, Thomas Telford.

Lane, S.N., Reaney, S.M. and Heathwaite, A.L., 2009. Representation of landscape hydrological connectivity using a topographically driven surface flow index. *Water Resources Research*, 45, W08423, doi:10.1029/2008WR007336.

Maltby, A. 2002. Application of semi-quantitative electrofishing to juvenile salmonid densities in the river Eden catchment in support of river regeneration activities. Eden Rivers Trust.

Matthews, R., 2006. The People and Landscape Model (PALM): Towards full integration of human decision-making and biophysical simulation models. *Ecological Modelling*, **194**, 329-43.

McKergow, L.A., Weaver, D.M., Prosser, I.P., Grayson, R.B. and Reed, A.E.G., 2003. Before and after riparian management: sediment and nutrient exports from a small agricultural catchment, Western Australia. *Journal of Hydrology*, **270**, 253-72.

Meador, M.R. and Goldstein, R.M., 2003. Assessing water quality at large geographic scales: Relations among land use, water physicochemistry, riparian condition, and fish community structure. *Environmental Management*, **31**, 504-17.

Munafo, M., Cecchi, G., Baiocco, F. and Mancini, L., 2005. River pollution from non-point sources: a new simplified method of assessment. *Journal of Environmental Management*, **77**, 93–98.

Muscutt, A.D., Harris, G.L., Bailey, S.W. and Davies, D.B., 1993. Buffer zones to improve water quality – a review of their potential use in UK agriculture. *Agriculture, Ecosystems and Environment*, **45**, 59-77.

Naden, P.S. and Cooper, D.M., 1999. Development of a sediment delivery model for application in large river basins. *Hydrological Processes*, 13, 1011-1034.

Newson, M.D., 1997. *Land, Water and Development: River Basin Systems and their Sustainable Management*, Routledge, London.

O'Grady, M.F., 1993. Initial observations on the effects of varying levels of deciduous vegetation on salmon stocks in Irish waters. *Aquaculture Research*, **24**, 563-73.

Owens, P.N., Walling, D.E. and Leeks, G.J.L., 1999. Deposition and storage of fine-grained sediment within the main channel system of the River Tweed, Scotland. *Earth Surface Processes and Landforms*, 24, 1061-76.

Parsons, H., Walker, J., Scarlet, P. and Hornby, D. 2001. River Eden RHS and geomorphology evaluation. Final Report 2001. Environment Agency.

Perry M.C., Hollis D.M. 2005. The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology,* **25,** 1041-1054

Pess, G.R., Montgomery, D.R., Steel, E.A., Bilby, R.E., Feist, B.E. and Greenberg, H.M., 2002. Landscape characteristics, land use, and coho salmon (Oncorhynchus kisutch) abundance, Snohomish River, Wash., USA. *Canadian Journal of Fisheries and Aquatic Science*, **59**, 613–623.

Pionke, H.B., Gburek, W.J. and Sharpley, A.N., 2000. Critical source area controls on water quality in an agricultural watershed located in the Chesapeake Basin. *Ecological Engineering*, **14**, 325–335.

Planchon, O., and F. Darboux. 2001. A fast, simple and versatile algorithm to fill the depressions of digital elevation models. *Catena*, **146**, 159–176.

Poff, N.L. 1997. Landscape filters and species traits: Towards mechanistic understanding and prediction in stream ecology. *Journal of the North American Benthological Society,* **16**, 391–409.

Priess, J.A., de Koning, G.H.J., Veldkamp, A., 2001. Assessment of interactions between land use change and carbon and nutrient fluxes in Ecuador. *Agriculture, Ecosystems and Environment*, **85**, 269–279.

Quinn, P., 2004. Scale appropriate modelling: representing cause-and-effect relationships in nitrate pollution at the catchment scale for the purpose of catchment scale planning. *Journal of Hydrology*, **291**, 197–217.

Quinn, P.F. and Anthony, S., 1999. The use of high resolution digital terrain models to represent wash-off in natural and man-influenced environments. *IAHS Publication*, **257,** pp. 255–264.

Reaney, S.M., Bracken (nee Bull), L.J. and Kirkby, M.J. 2007. The use of the Connectivity of Runoff Model (CRUM) to investigate the Influence of Storm Characteristics on Runoff Generation and Connectivity in Semi-Arid Areas. *Hydrological Processes*, **21**, 894-906.

Rich, C.F., McMahon, T.E., Rieman, B.E., and Thompson W.L., 2003. Local-habitat, watershed, and biotic features associated with bull trout occurrence in Montana streams. *Transactions of the American Fisheries Society*, **132**, 1053-64.

Richards, C., Haro, R.J., Johnson, L.B. and Host, G.E., 1997. Catchment and reach-scale properties as indicators of macroinvertebrate species traits. *Freshwater Biology*, **37**, 219-30.

Roth, N.E., Allan J.D. and Erikson, D.L., 1996. Landscape influences on stream biotic integrity assessed at multiple spatial scales. *Landscape Ecology*, **11**, 141-56.

Smith, R.V., Lennox, S.D., Jordan, C., Foy, R.H. and McHale, E., 1995. Increase in soluble phosphorus transported in drainflow from a grassland catchment in response to soil phosphorus accumulation. *Soil Use and Management,* **11**, 204–209.

Soulsby, C., Malcolm, R., Gibbins, C., and Dilks, C., 2001. Seasonality, water quality trends and biological responses in four streams in the Cairngorm Mountains, Scotland. *Hydrology and Earth Systems Science*, **5**, 433-50.

Stauffer, J.C., Goldstein, R.M. and Newman, R.M., 2000. Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. *Canadian Journal of Fisheries and Aquatic Science*, **57**, 307–316.

Suter, H.C. and Huntingford, F.A., 2002. Eye colour in juvenile Atlantic salmon: effects of social status, aggression and foraging success. *Journal of Fish Biology*, **61**, 606-14.

Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. *Water Resources Research*, **33**, 309-319.

Townsend, C.R., Arbuckle, C.J., Crowl, T.A., and Scarsbrook, M.R., 1997. The relationship between land use and physicochemistry, food resources and macroinvertebrate communities in tributaries of the Taieri River, New Zealand: A hierarchically scaled approach. *Freshwater Biology*, **37**, 177-191.

Townsend-Cartwright, S. 2004. Semi-quantitative electrofishing to assess the status of salmonid fry populations within the Eden catchment in 2002 and 2003 in support of river restoration work. The Eden Rivers Trust.

Vatn, A., Bakken, L., Bleken, M.A., Baadshaug, O.H., Fykse, H., Haugen, L.E., Lundekvam, H., Morken, J., Romstad, E., Rørstad, P.K., Skjelvåg, A.O. and Sogn, T., 2006. A methodology for integrated economic and environmental analysis of pollution from agriculture. *Agricultural Systems*, **88**, 270-93.

Wang, L., Lyons, J., Kanehl, P., and Gatti, R. 1997. Influences of watershed land use on habitat quality and biotic integrity in Wisconsin streams. *Fisheries (Bethesda)*, **22**, 6–12.

Wang, L., Lyons, J., Rasmussen, P., Seelbach, P., Simon, T., Wiley, M., Kanehl, P., Baker, E., Niemela, S. and Stewart, P.M., 2003. Watershed, reach, and riparian influences on stream fish assemblages in the Northern Lakes and Forest Ecoregion, USA. *Canadian Journal of Fisheries and Aquatic Science*, **60**, 491–505.

Weber, A., Fohrer, N. and Möller, D., 2001. Long-term land use changes in a mesoscale watershed due to socio-economic factors - effects on landscape structures and functions. *Ecological Modelling*, **140**, 125–140.

Wiley, M.J., Kohler, S.L., and Seelbach, P.W. 1997. Reconciling landscape and local views of aquatic communities: Lessons from Michigan trout streams. *Freshwater Biology*, **37**, 133–148.

Wolf, J., Hackten Broeke, M.J.D., and Rötter, R., 2005. Simulation of nitrogen leaching in sandy soils in the Netherlands with the ANIMO model and the integrated modelling system STONE. *Agriculture, Ecosystems and Environment*, **105**, 523–540.

Wood, F.L., Heathwaite, A.L. and Haygarth, P.M., 2005. Evaluating diffuse and point phosphorus contributions to river transfers at different scales in the Taw catchment, Devon, UK. *Journal of Hydrology*, **304,** 118–138.

Young, P.C., Parkinson, S. and Lees, M.J., 1996. Simplicity out of complexity: Occam's razor revisited. *Journal of Applied Statistics*, **23**,165–210.

Figure 1. A hydrologically disconnected fine sediment source: surface erosion by overland flow and its accumulation in a surface hollow. Source, Eden Rivers Trust.

Figure 2. Dilution effects where a high risk suspended sediment tributary meets a low risk suspended sediment tributary. Source, Eden Rivers Trust.

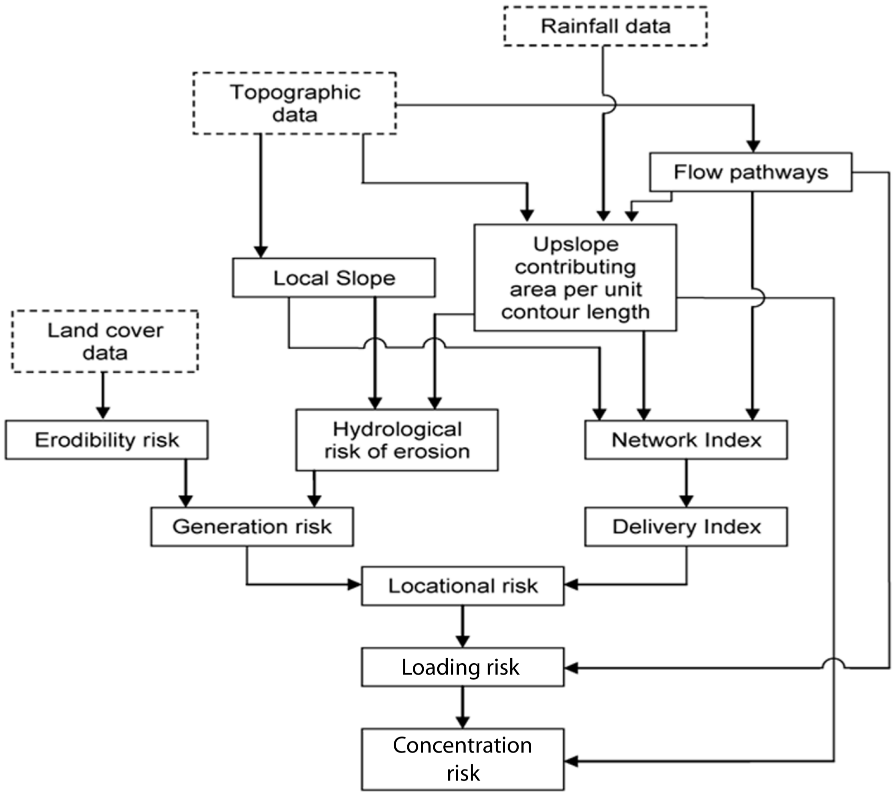


Figure 3. Overview of the SCIMAP model for fine sediment risk

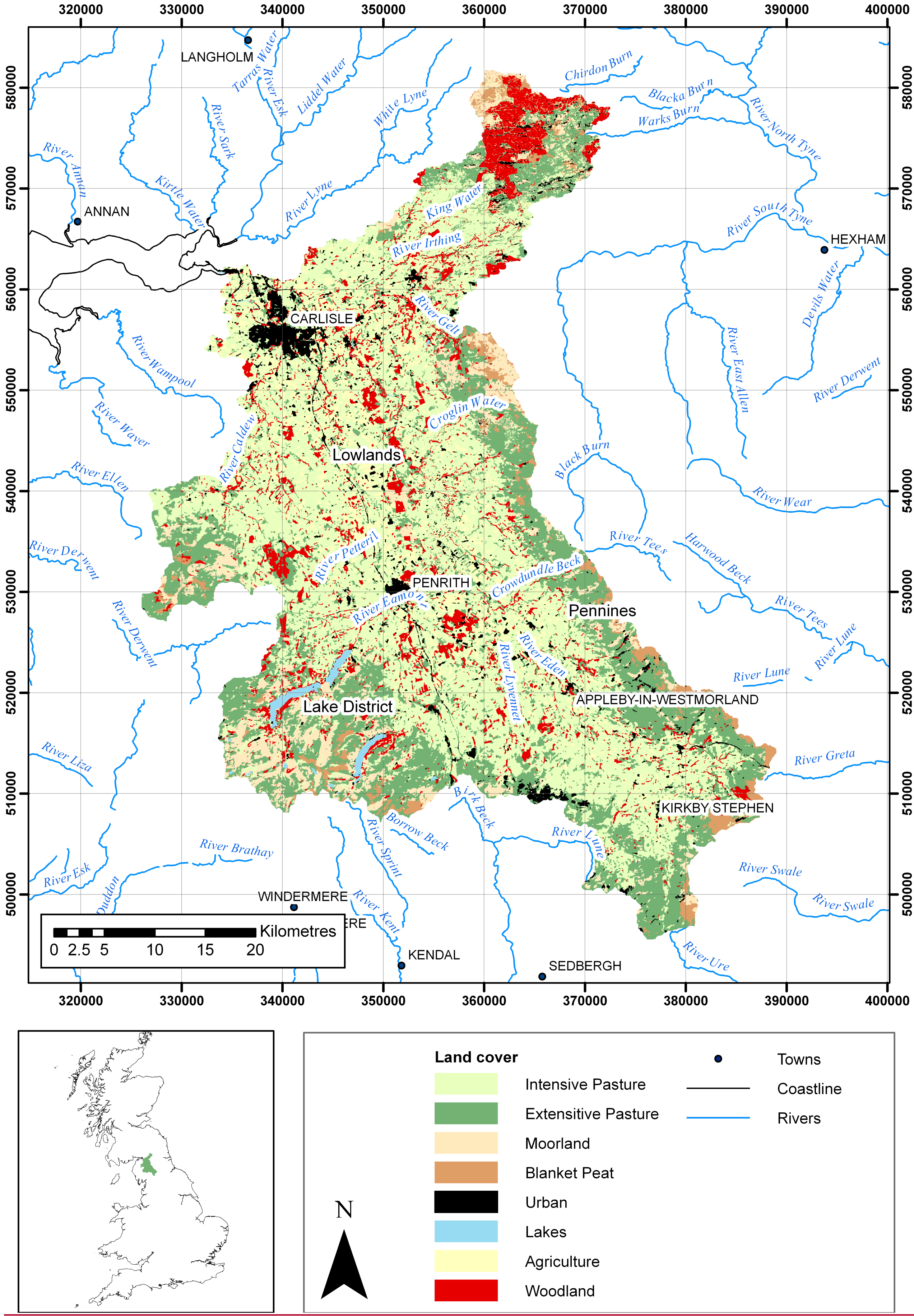


Figure 4. The River Eden catchment, north England, UK. Crown Copyright Ordnance Survey. An EDINA Digimap/JISC supplied service.

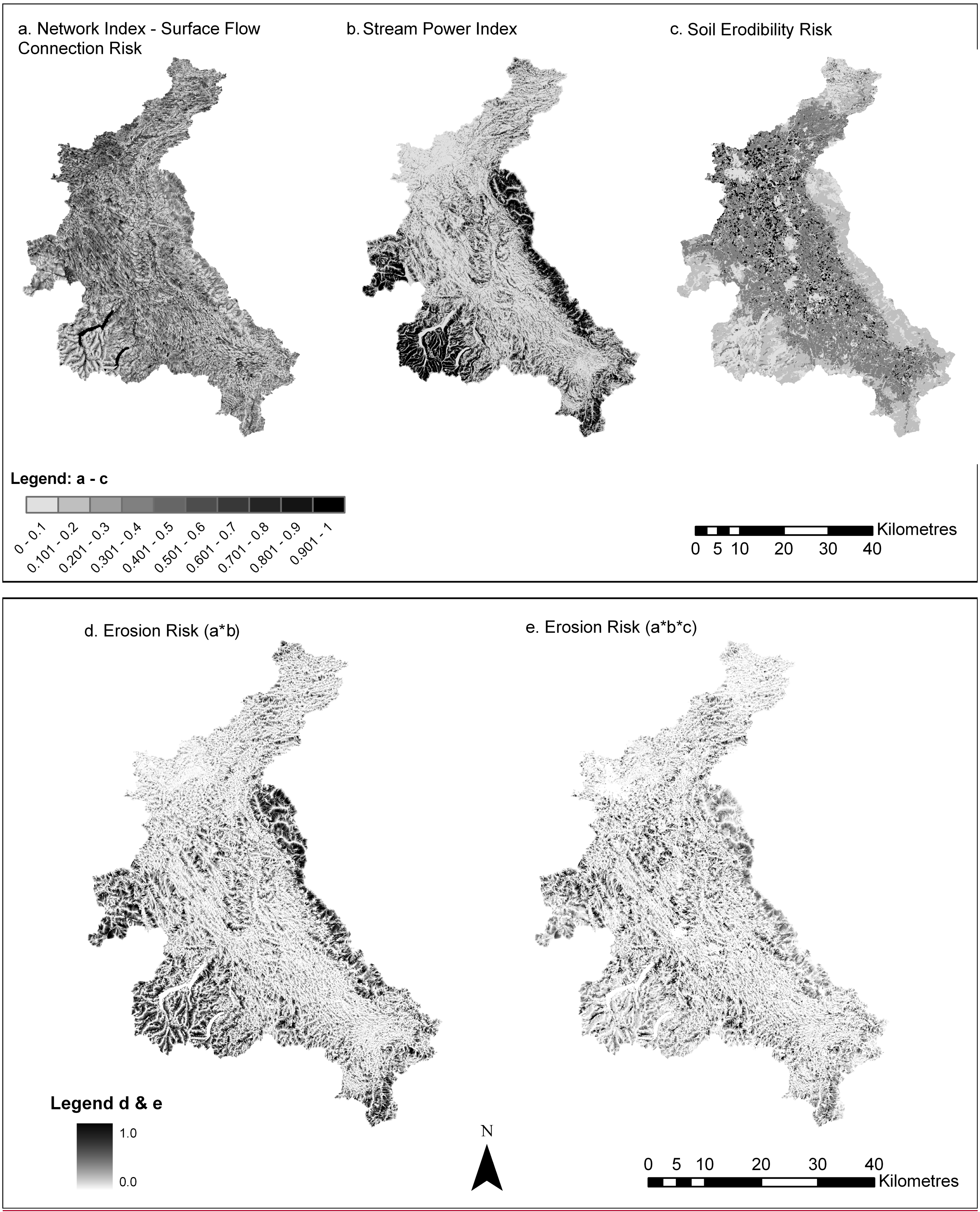


Figure 5 shows the main derivates of the DEM that are used in the calculation of the point scale soil erodibility risk. Figure 5a shows the network index which is used to determine the surfaces flow connection risk. Figure 5b shows the predicted spatial pattern of the soil erodibility as determined by the surface land cover Figure 5c shows the distribution of the stream power index which represents the energy available to erode the surfaces. Figure 5d and Figure 5e show the convolution of the source area analysis with the connectivity analysis.

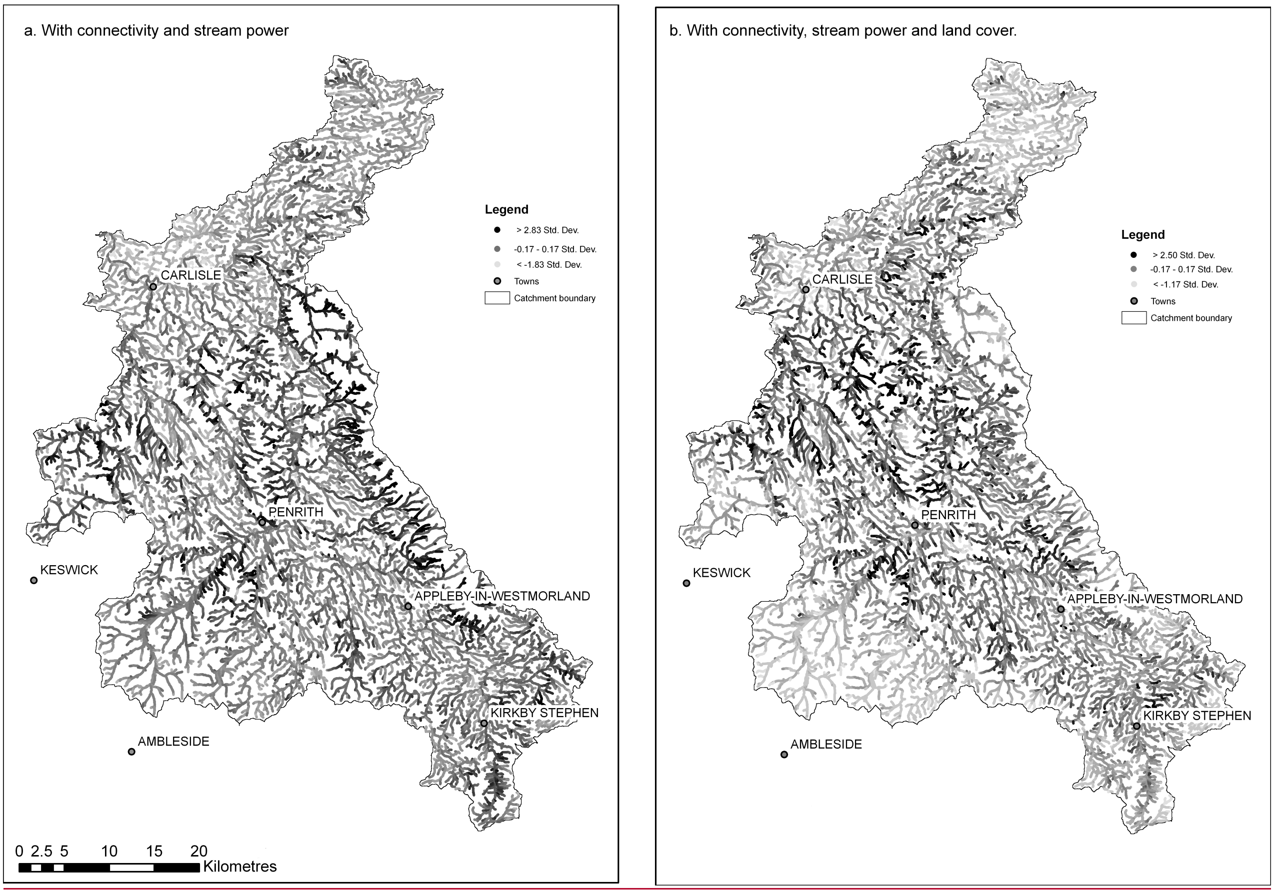
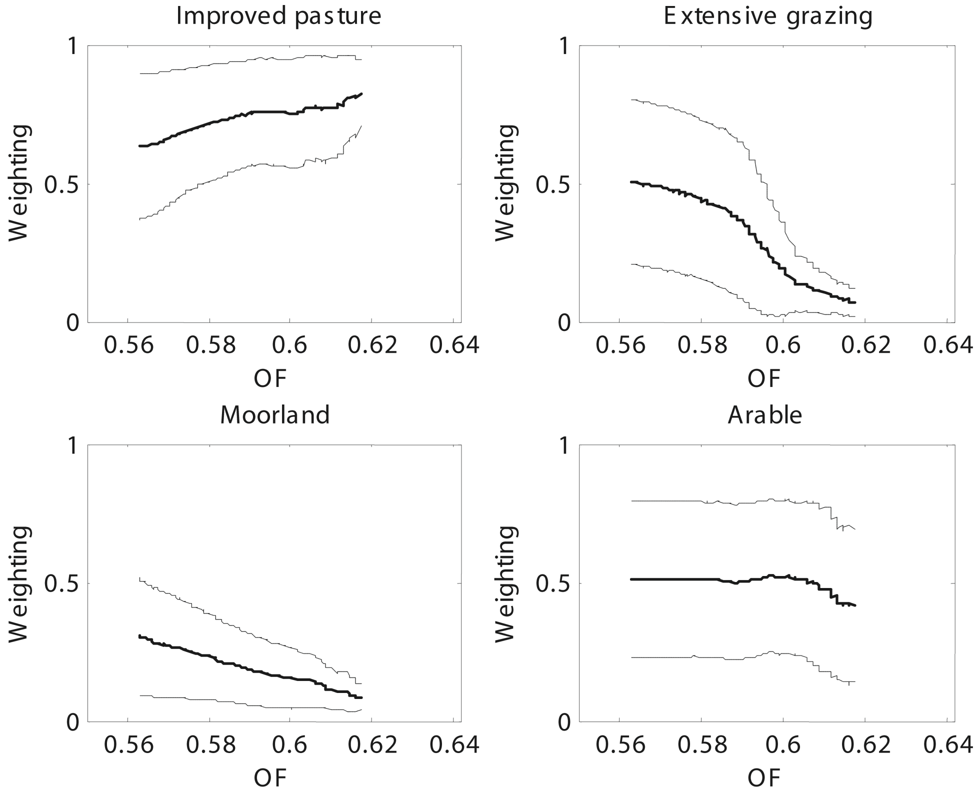
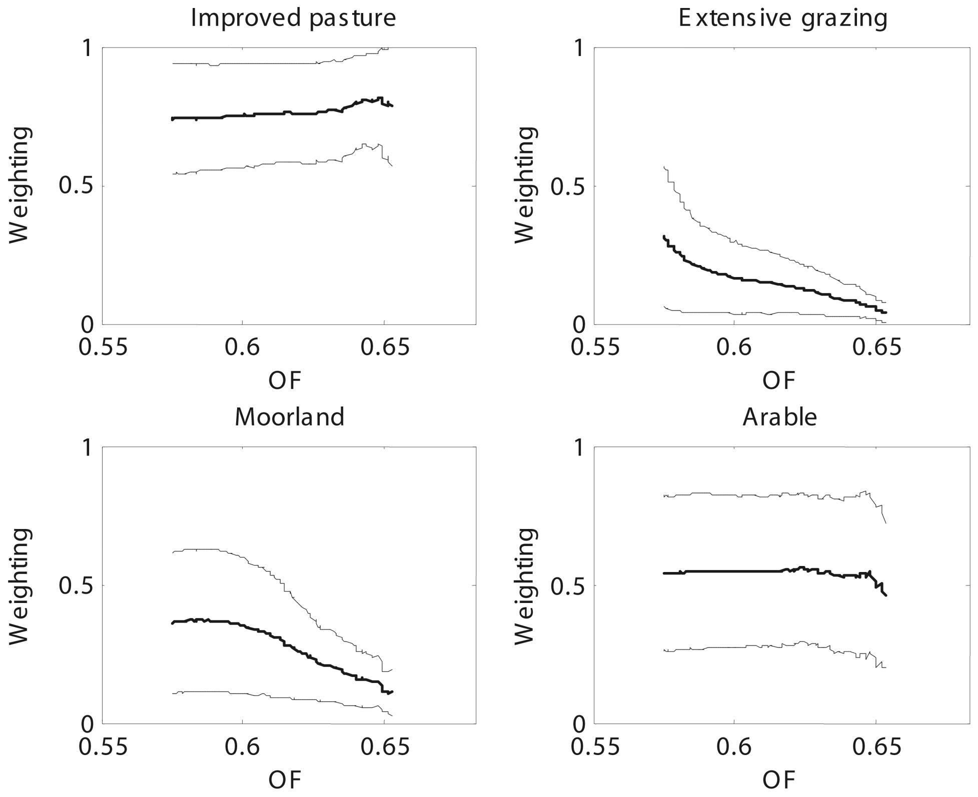
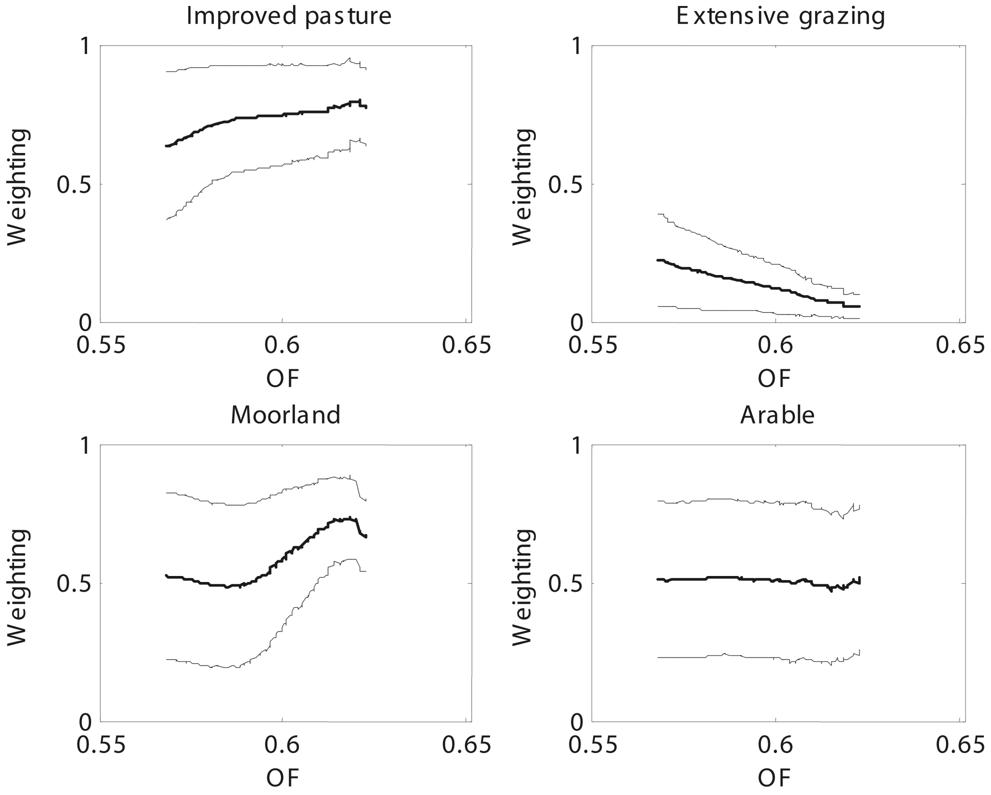


Figure 6. The accumulated risk weighted by the dilution potential, essentially the risk concentration


Improved pasture
Extensive grazing
Moorland
Arable
Weighting
OF
1
0.5
0
0.56
0.58
0.6
0.62
0.64


Improved pasture
Extensive grazing
Moorland
Arable
Weighting
OF
1
0.5
0
0.55
0.6
0.65



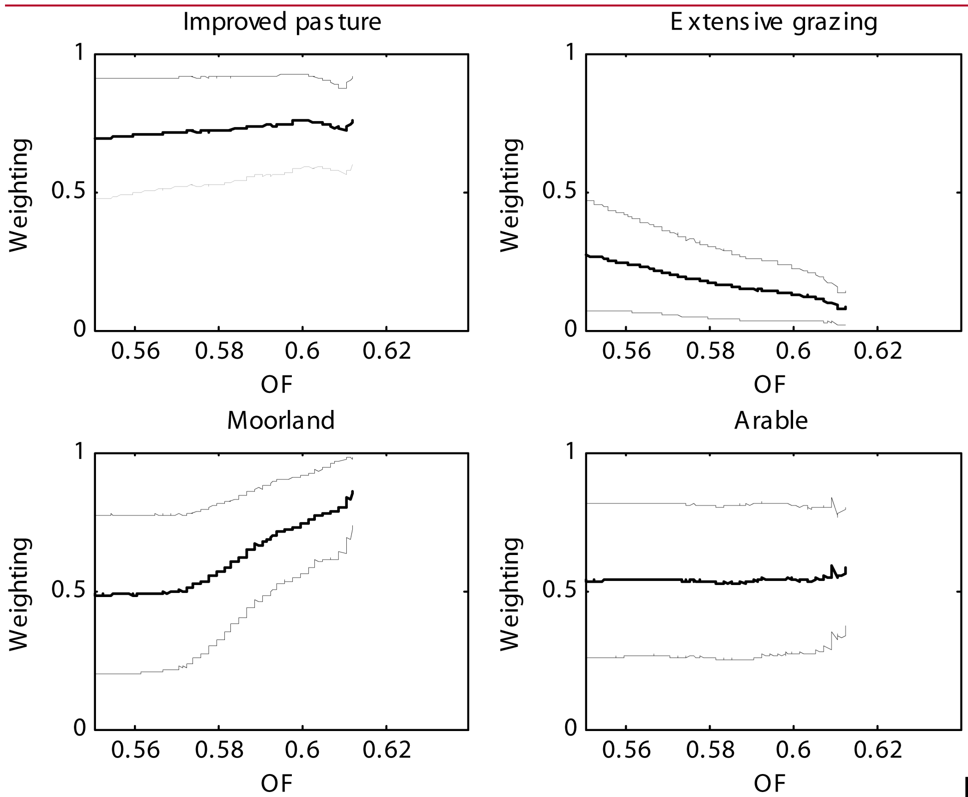


Figure 7. The results of the Bayesian-based inverse modelling for the years 2002 (Figure 7a and Figure 7c)

and 2003 (Figure 7b and Figure 7d) for both trout (Figure 7a and Figure 7b) and salmon (Figure 7c and Figure 7d). In these plots, the heavy solid line plots the mean erodibility against the associated objective function value. As the objective function is defined as a perfect association with a value of 1.0, the mean weighting for the best 10 simulations is the farthest right of each plot and mean weighting for the best 200 simulations is the farthest left on each plot. The light solid line isthe standard deviation of the erodibilities assigned to each land cover.

7a

7b

7c

7d

Table 1. Statistical comparison of risk estimates with and without trout and salmonid fry present.

Year	Species	Fry present	Fry absent	*p* value	
				Mann Whitney	Kolmogorov-Smirnov
2002	Trout	0.0450 ± 0.0024, n = 167	0.0587 ± 0.0025, n = 184	<0.0001	0.0003
	Salmon	0.0506 ± 0.0021, n = 204	0.0543 ± 0.0031, n = 147	0.937	0.050
2003	Trout	0.0521 ± 0.0022, n = 192	0.0762 ± 0.0046, n = 83	<0.0001	<0.0001
	Salmon	0.0575 ± 0.0024, n = 154	0.0619 ± 0.0040, n = 121	0.773	0.018