Land Use Policy 103 (2021) 105322

Contents lists available at ScienceDirect

Land Use Policy

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

The spatial impact of rural economic change on river water quality[☆]

Cathal O'Donoghue[c,*], Cathal Buckley[a], Aksana Chyzheuskaya[c], Stuart Green[a], Peter Howley[b], Stephen Hynes[c], Vincent Upton[a], Mary Ryan[a]

[a] *Teagasc Rural Economy and Development Programme, Athenry, Galway, Ireland*
[b] *University of Leeds, UK*
[c] *National University of Ireland, Galway, Ireland*

ARTICLE INFO

Keywords:
Agricultural activity
Ecological water quality
Water Framework Directive

ABSTRACT

This paper, using Ireland as a case study, examines the relationship between rural economic activities and river water quality. The stipulation from the EU water framework directive (WFD) that all surface waters in the EU must be of 'good ecological status' necessitates a quantitative understanding of the major determinants of water quality. Within this context, this paper combines a number of spatial datasets relating to agricultural, land use, residential and industrial activities, to examine the major economic influences on the ecological quality of water resources. It is hoped that providing a comprehensive understanding of the effect of a variety of economic activities that influence the ecological quality of water will be an important tool in the management of risk and will allow for more appropriate land use planning aimed at restoring and maintaining water quality as required by the WFD. Results indicate that the level of forestry, industrial activity, the intensity and type of agricultural activity and the type of wastewater treatment in an area are all critical factors affecting the quality of water resources. The model finds that relationship between agriculture and water quality improved over time during a period where there was substantial legislative measures and financial support to facilitate improved water quality.

1. Introduction

In many countries, the biggest source of water pollution today is agriculture, not cities or industry (FAO, 2018). In targeting water pollution across the EU, the European Commission adopted the Water Framework Directive (WFD) in 2000. The WFD requires the integrated management of water resources throughout the European Union (EU) and commits Member States (MS) to ensure that all surface water bodies are of 'good' ecological[1] status by 2027 at the latest. One of the innovations of the WFD is that unlike preceding regulations, it requires MS to manage water bodies on a river-basin scale, a natural hydrological and geographical unit. In parallel, the Nitrates Directive[2] was introduced in 1991 to limit fertiliser usage with the aim of reducing the risk of loss of nutrients to water from agriculture. However while water quality has improved over the last 25 years, only 43 % of waterbodies were at good ecological status in 2009, with only a modest increase to 53 % by 2015.

Water quality depends largely on the local geology and ecosystem but human activities can also negatively affect water quality (Donohue et al., 2006). The primary means by which ecological water quality can be impacted are the loss of sediment to surface waters (which can interfere with the breeding process of indicator species on which ecological classifications are based) and the loss of nutrients nitrogen (N) and Phosphorus (P) to water, which can lead to eutrophication of waterbodies. While N and P are essential for plant growth, eutrophication arises from the oversupply of nutrients, which leads to excessive growth of plants and algae. The microbes that decompose these organisms consume large amounts of oxygen, leading to oxygen depletion and the reduction in ecological indicator species.

The definition of ecological status takes into account specific aspects of the biological quality elements, such as the composition and abundance of aquatic flora. Human activity can speed up natural

☆ This paper acknowledges funding from the Environmental Protection Agency (EPA) Research Programme 2014–2020.

* Corresponding author.
E-mail address: cathal.odonoghue@nuigalway.ie (C. O'Donoghue).

[1] Ecological status is measured using a 'Q-value', ranging from pristine (Q5) to bad (Q1) water quality.

[2] https://ec.europa.eu/environment/water/water-nitrates/index_en.html.

https://doi.org/10.1016/j.landusepol.2021.105322
Received 28 August 2019; Received in revised form 21 January 2021; Accepted 25 January 2021
Available online 5 February 2021

eutrophication, leading to cultural eutrophication. Due to the management of land and population settlement (with associated disposal of human waste), land runoff is accelerated and more nutrients are lost to lakes and rivers, and then to coastal estuaries and bays (Schindler et al., 2008).

There is a consensus in the scientific literature that land use (such as agriculture, forestry and human waste disposal), has a significant effect on river water quality (Woli et al., 2008; Varanka and Luoto, 2012), with much of the literature focused on the agricultural sector (Vatn et al., 1997, 2006; Lennox et al., 1998; Cuttle et al., 2006; Brouwer et al., 2008; Fezzi et al., 2008; Pulido-Velázquez et al., 2008; Volk et al., 2008). While research has demonstrated a link between agriculture and water quality, the diffuse nature of the loss of nutrients from soil to water is the focus of deterministic hydrological models (such as Withers and Haygarth (2007), that examine the apportionment of nutrient loads to chemical water quality as a function of inputs and hydrology at catchment or water-body level.

In addition to land use, landscape characteristics such as slope, soil and bedrock affect river water quality (Varanka and Luoto (2012; Mockler et al. (2017). In studying the effects of land-use and catchment characteristics on lake water quality, Curtis and Morgenroth (2013) used a statistical model to attribute the variation in water quality across Irish lakes to a range of contributory factors including septic tanks, urban waste-water treatment, phosphorus excreted by livestock as well as geomorphological and climatic variables. While a limited number of studies have examined the role of climatic variables such as precipitation and temperature (Larned et al., 2004), studies that simultaneously consider land use, soils/geology and climatic variables are rare. In the next section, we provide a short literature review on attempts to understand these relationships.

In facing the challenges related to sustainable agricultural intensification, climate change and the increasing demand for the delivery of ecosystem services from rural environments, there is a need for holistic analyses to examine the inter-temporal impact of rural land-uses on water quality. The objective of this study is to understand how land use changes have impacted on WFD objectives over time, in complex hydrological conditions.

The primary contribution of this paper, therefore is to both incorporate a broader set of explanatory factors of water quality and to capture changes in these explanatory factors over time in order to help to understand the high-level policy and industry impacts on water quality. This study focuses specifically on the effect of the nutrient load related to agriculture, forestry and human waste disposal over time. As WFD water-quality is determined on the basis of a qualitative ecological water quality measurement, a statistical model is required to indirectly examine the relationships between land use and qualitative water quality, taking hydrological interactions into account. A number of spatial scales are required as while many deterministic assessments against water quality objectives are undertaken at the rivers sub-catchment scale, many of the policy or socio-economic drivers are at national scale. As a result, this analysis requires a modelling framework that incorporates the local (sub-) catchment scale containing the water quality monitoring points, as well as the national scale, in order to be able to scale-up the analysis to ensure consistency with WFD reporting of the proportion of monitoring points that are 'good' status or better.

We focus on Ireland as a case study given (a) that in 2006 the Environmental Protection Agency (EPA) in Ireland reported that water quality was at a level below that required by the WFD (EPA, 2006a), (b) the heavy reliance on pastoral agricultural land use (c) the rural nature of the population with domestic waste treatment facilities and (d) major economic and population changes since 1995. As a pastoral, mainly animal-based system, where the temperate climate allows animals to graze outdoors for much of the year, we expect to observe a close link between agricultural intensity and water quality (Hennessy et al., 2005). Ireland as an export country, places a significant value on marketing the sustainability credentials of its food products (Dillon et al., 2010; Hennessy et al., 2013; Lynch et al., 2016). In such a context, it is of interest to know whether regulations such as the WFD and the Nitrates Directive, and voluntary agri-environmental scheme (AES) measures undertaken to improve compliance with the regulations, have had an impact over time.

In this paper, we examine long-run trends between economic activity and water quality, conditional on the local environmental/hydrological. This requires a number of spatial datasets to explore the effect of land use, soils/geology and climatic variables on river water quality in Ireland over a 20 year period. This period spans a volatile period economically, incorporating an economic boom from the mid-1990s to the mid-2000s followed by a recession until 2012. The period also saw major changes in relation to agricultural policy as a result of environmental regulation and and programmes of incentives to manage agricultural waste and to increase the level of private sector afforestation (new planting).

The analysis is centered around the development of a statistical model for water quality data, using a qualitative measure of ecological water quality (Q-value) recorded across EPA river monitoring stations as the dependent variable, to explain variations in river water quality. The use of economic data that vary both across space and time (i.e. panel data), allows us to examine whether the effect of various land use activities on river water quality has changed over time.

2. Literature review

There is a substantial literature internationally on the topic of understanding changes in water quality. A variety of analytical methodologies are used at varying spatial scales including Hierarchical Cluster Analysis (Wang et al., 2015; Garizi et al., 2011; Li et al., 2015; Daou et al., 2018); Pearson's correlation coefficient analysis to examine the relationship between the water quality parameters and impact factors (Mei et al., 2014; Ferrier et al., 2001); factor analysis (Rheuban et al., 2015);bivariate geographically weighted regression modelling approaches (Wang and Zhang, 2018); machine learning techniques (Elith et al., 2008; Breiman, 2001) and other regression based modelling approaches (Chang, 2008; Singh and Chang, 2014; Mainali and Chang, 2018). Li et al. (2015) supplement their cluster analysis with a Kendall test and the Moran's index to detect the seasonal and inter-annual variations in their dataset and the spatial autocorrelation of the explanatory variables in their model of variations of surface water quality while Jokinen et al. (2012) employ a descriptive approach to identify the main land use driver of a variety of water-based pathogens in Canadian waters.

Previous research attempting to explain the variation of water quality or the impact on societal welfare has also been focused at different spatial scales, from national (Ferrier et al., 2001; Donohue et al., 2006), to regional (Ahearn et al. 2005; Yan et al. 2015), to smaller spatial units, such as a water catchments (Murphy et al., 2015), or water management units (Curtis and Morgenroth, 2013; Hynes and O'Donoghue, 2020). Analyses at the lower spatial scales tend to capture greater detail in relation to drivers and pathways, but the results of such studies may not be relevant for different catchment types. There is clearly a trade-off between detail at a small spatial scale and the capacity to do a broader landscape or national scale analysis, which is useful in policy formation.

While many of the regression models that examine drivers of water quality contain only a limited number of explanatory variables (Singh and Chang, 2014; Liu et al., 2016) a number of others have included a comprehensive set. For example, in a study of the Sorraia Basin in Portugal and Spain, Segurado et al. (2018), estimated a multiple linear regression model, incorporating land use variables such as the share of urban, forestry, agriculture and irrigated crop upstream, as well as nutrient stressors (total N and total P), hydrological stressors (number and duration of events) and variables describing natural environmental variability (climate, slope, distance to waterbody and extent of

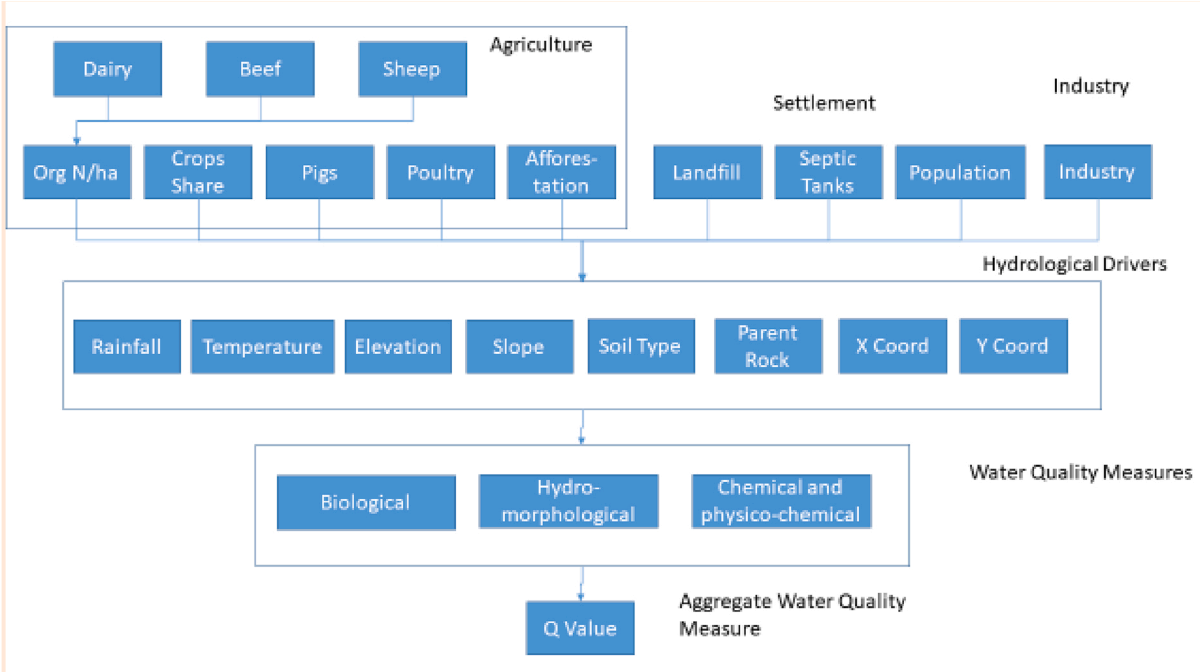


Fig. 1. Theoretical Framework.

catchment).

3. Theoretical framework

Water quality is defined by its physical, chemical and biological characteristics and is usually described in terms of certain criteria and standards. The quality of water in a particular water body determines the potential uses of water. The factors that affect water quality are complex and are assessed on a number of indicators ranging from water temperature, pH and chemical components, to biological and microbiological parameters (Novotny, 2003). This paper is concerned with ambient water quality and status as defined by the WFD which directs EU member states to conduct an assessment of water body status and to establish the classification schemes for 1) biological, 2) hydro-morphological (quantity and dynamics of flow, inputs to and from groundwater, presence or absence of impediments to fish movement, depth and width, structure of river substrate and riparian zone), 3) chemical and physico-chemical quality elements.

This paper examines trends in the relationship between economic activity on water quality over time. The overall theoretical framework is presented visually in Fig. 1, before discussing the range of variables required to model changes in water quality.

In relation to sources of pollution, a distinction is made between diffuse pollution and point source pollution. Where the specific source or location can be identified, pollution is described as point source. In the case of diffuse pollution, the source of pollution cannot be readily identified as it originates from air, land surface, subsurface zones and drainage systems and results from the interaction between weather events and the landscape.

Diffuse nutrient pollution has been shown to provide a substantial risk to water quality in rivers and lakes (Haygarth et al., 2003; Carpenter et al., 1998), but is difficult to monitor and control. However, the intensity of livestock agriculture, reflected in organic N load (the level of excretion of organic nitrogen (N) by grazing animals) is considered to be a crucial factor in N loss to groundwater (EPA, 2006b). Diffuse pollution is often a result of use and misuse of land, and the causes are mostly socioeconomic, encouraged by tradition, government subsidies, demands for cheap products, and lack of information about pollution and polluting behaviour (Novotny, 2003). Tietenberg and Folmer (2002) identify the following main sources of non-point (diffuse) pollution: runoff from streets, agriculture, atmospheric deposition, forest management and industrial (waste) disposal systems.

One of the most significant global sources of diffuse pollution is agriculture, arising from the use of land to produce crops, fibre and agri-food products from livestock (FAO, 2018). Pollutants that commonly originate from agricultural land include fertilisers such as nitrogen (N) and phosphorus (P) (which can derive from either inorganic fertilizer compounds or organic manure/animal slurry), or pesticides, pathogens and sediment (Novotny, 2003). However, while agriculture is a key generator of diffuse pollution, it is not the only source and other land use activities such as forestry, industrial land use, and human waste treatment among others, all contribute to the problem (Defra, 2002).

Nutrient and sediment losses from forests arise largely as a consequence of 'disturbances' such as forest establishment (drainage, cultivation, fertilisation, planting) and harvesting (removal of thinnings or clearfell of final crop). Historically in Ireland, much of the afforestation (new planting) was undertaken by the State, but since the advent of Irish and EU funding schemes for private growers in the late 1980s, virtually all afforestation is now undertaken (largely by farmers) on previously agricultural land. Thus, private planting increased from around 2,000 ha (ha) per year in the mid-1980s to a high of 17,000 ha in 1995, before dropping back to around 8000 ha per year from 2006 (Ryan et al., 2014). As a result, private forests are relatively young with more potential 'disturbances' arising from afforestation than from harvesting.

Human settlement, economic development and population growth also contribute to pollution in streams and the wider environment. The waste from domestic use, traffic and industry is washed off into storm gutters or streams during rainfall (Novotny, 2003), resulting in increased runoff volume to surface water bodies.

Another important source of diffuse water pollution is individual waste-water treatment (IWWT) systems. One of the features of the Irish landscape is a large number of dispersed farms and rural dwellings, resulting in a high proportion of rural households that are not connected to sewage treatment systems but have individual IWWTs. While many modern systems are more effective, traditional 'septic tanks' rely on anaerobic processes to reduce solids and organic components and the porosity of the soil to adequately act as a drain field. Ireland has over 400,000 septic tanks throughout the territory (EC Ireland, 2011). Poorly sited, installed and maintained septic tanks can affect ground water, allowing untreated chemicals and bacteria entering the environment.

Diffuse pollution has a significant stochastic component that depends on fluctuations in weather and other environmental factors (Olmstead, 2010). In studying interactions between agriculture, meteorology and water quality in Ireland, Schulte et al. (2006) found that there are regional differences in nutrient loss that should be taken into account. These are attributed to differences in soil characteristics and rain intensity, with substantial inter-temporal variation in agro-meteorological conditions.

4. Methodology

Based on the literature, we make a number of methodological choices:

- As a policy focused exercise, we choose water quality metrics that are of direct relevance to the WFD.
- As a national-focused assessment of drivers linked to improvement in the policy outcome, we choose to take a national perspective rather than a catchment perspective.
- Given the complexity of this approach, we utilise a multi-variate approach rather than a bivariate approach.
- Given the important changes in the trend in load rather than land use, we take a more detailed approach than that adopted by Donohue et al. (2006); Curtis and Morgenroth (2013) and Segurado et al. (2018) by incorporating more detail on nutrient load.

It could be argued that the observed effect of the explanatory variables that reflect different land use activities may be partly due to the correlation between these variables with other potentially more important explanatory geomorphological and climatic factors. To control for the effect of these potential confounding variables, a variety of geomorphological variables such as measures of elevation, slope, soil type as well as climatic variables such as rainfall and temperature, are included as covariates in our regression analysis. The inclusion of such variables is expected to highlight the complexity associated with environmental drives and provide additional information on differential impacts of organic N loads and also on pathways for diffuse pollution, given the constraints imposed on the source-pathway-receptor transport of nutrients to water by soils and landscape conditions (Micha et al., 2018). In addition, as hydrological impacts are not necessarily linear, it is important to include interaction terms as well as explanatory variables.

4.1. Dependent variable

In order to be consistent with the policy framework we are considering, this study utilises the WFD ecological (water) 'Q value' Quality Rating System[3] as the dependent variable, (denoted *Yit)*. This system uses a quality (Q) index from 1 to 5 to assess the ecological quality of water resources at monitoring points in each catchment at time *t*. The five water quality status classes are: high (Q5), good (Q4), moderate (Q3), poor (Q2) and bad (Q1), resulting in an ordinal dependent variable that takes on five discrete values (Q5 represents higher water quality status than 4, which means a higher status than 3, and so on). 'High status' (Q5) represents little or no human pressure and is the 'reference' or 'pristine' classification. Water quality assessment is based on the extent of deviation from these reference conditions, following the definitions in the WFD. 'Good status' means 'slight' deviation, 'moderate status' means 'moderate' deviation, and so on. As the ecological status is dependent on the presence and abundance of macro-invertebrate indicator species, the dependent variable provides a measure of the aggregate longer-term status of water-bodies.

Water quality is assessed at each monitoring point on a three-year cycle. The subscript *i* of *Y* indicates the *i*th water quality monitoring point, i = {1, …, n}, *t* represents years 1991, 2002 and 2011. Y_i is a scalar that takes the values of 1, 2, 3, 4 and 5. *Y* is an (*n x 1*) vector indicating the water quality level at each monitoring point. The *i*th element of the vector indicates the *i*th water quality monitoring point's level.

4.2. Spatial scale

This study is undertaken on a national scale, which is the scale of assessment for the WFD. However, the unit of analysis of the study is based around the dependent variable (Q value) which is collected at over 3000 sampling points across the river system. The explanatory variables relate to economic data for the sub-catchment immediately upstream of water quality monitoring points.

4.3. Explanatory variables

This study incorporates the main explanatory variables discussed in the literature including:

- land use and organic N-load in the sub-catchment
- population density and in particular septic tank density
- economic activity
- geomorphological and climatic variables such as soil, elevation, slope, rainfall and temperature.

Year dummy variables are included to capture time specific effects. Three agricultural census years of data were available so two year dummies are included. Characteristics such as physical land use and septic tank density of the river catchments are denoted by the vector X_{it} with k elements. The letter *k* indicates the *k*th independent variable, $k = \{1, \ldots, K\}$. *X* is an (*n* x *k*) matrix summarizing the economic and land use characteristics of each river catchment. The *n*th row indicates the characteristics of the nth catchment. Therefore, we can state that

$$Y_{it} = f(X_{it}) \ \forall \ i = 1, \cdots, n$$

4.4. Statistical method

Since the dependent variable is an ordered, qualitative variable, we estimate the relationship between Y and X with an ordinal response model. In this case as the difference between Q values is not the same, a continuous regression model is not appropriate, therefore an ordered probit model is estimated using the method of maximum likelihood via the Newton-Raphson algorithm (Long, 1997).

It is assumed that the level of water quality in a river catchment, denoted Y*it**, is a continuous function of catchment characteristics denoted by X_{it}, a vector of parameters of dimension (*k* x 1) denoted *β*, and a disturbance term *ε*, which is normally, identically, and independently distributed $\varepsilon \sim N(0, \sigma^2)$. Increasing values of Y*it** indicate an increasing level of water quality associated with that river system.

$$Y_{it}* = \beta' X_{it} + \varepsilon_{it}$$

4.5. Temporal variability

Due to the importance of inter-temporal trends, we use panel data with observations that are repeated over time in the same location. This requires the use of a panel estimator (see Arellano (2003) for a

[3] A method whereby an ecological Quality-index is assigned to a river or stream based on macro-invertebrate data, aquatic macrophytes and phytobenthos.

Table 1
Description and summary statistics of main explanatory variables.

Variable	Description	Mean	S.D.
Septic Tanks Density	Quantity of septic tanks per kilometer $(km)^2$ per ED	44.8	132.1
Organic Nitrogen	Quantity of organic N produced per hectare per ED (kgs)	105.0	30.8
Cereal production	Proportion of ED under arable crops	0.1	0.1
Pig rearing	Dummy variable indicating EDs in the top 25 % of pig rearing EDs based on density per km^2	0.04	0.19
Poultry rearing	Dummy variable indicating EDs above median of poultry production based on density per km^2	0.1	0.31
Forest cover	Hectares (ha) of Forestry per ED	44.9	76.2
Landfill within 3km	Dummy variable indicating a landfill is within 3 km of a Q-value monitoring point.	0.01	0.11
Area of monitoring ED	Size of ED in km^2 in which monitoring point is located	20.9	14.0
Rainfall	Rainfall in millimetres	1033.7	246.3
Temperature	Temperature in Celsius	10.3	0.7
Elevation	Elevation in metres	101.4	66.5
Slope	Slope in degrees	4.3	2.7

discussion on panel estimation).

However we find that our data sources have inconsistent measurement points:

- water quality is sampled every three years
- weather – weekly or daily measurements
- population, septic tank density and economic data - five year intervals
- Census of Agriculture data are collected every 10 years, with 2010 being the most recent Census year.

Given that the explanatory variables do not have the same temporal resolution as the environmental variables, we create a panel for the period of the explanatory variables rather than the dependent variables. An attempt to smooth the non-annual data to fit the point of collection of the water quality data, as in the case of Curtis and Morgenroth (2013), resulted in much poorer explanatory power than using the actual or close to actual point. To avoid smoothing, we use the longest period (where the data actually change) and use a period that is comparable across all three data sources. Therefore, we have chosen 1991, 2002 and 2011 Census of Agriculture data, in line with the closest years of the Census of Population, namely 2000/2002 and 2010/2011. This trade-off may have the effect of reducing our understanding of variability around the trend.

4.6. Data

In order to model the relationship between water quality and upstream economic activity, data are necessary in relation to water quality, agricultural intensity, forest intensity, settlement intensity (specifically the use of septic tank-based IWWT) and economic intensity, particularly in relation to the generation of industrial waste. In the appendix we outline the datasets required for this analysis. These datasets include the EPA water quality monitoring (Q-value) data, spatially referenced industrial activity and septic tank distribution data which are taken from the small area Census of Population levels of agricultural activity which are derived from the Census of Agriculture, and forest land cover data which are available from the Forest Service Further information on datasets and the calculation of nutrient loads from agriculture is described in Appendix A1.

Table 2
Average Organic N and Septic Tank Density (per ha) per Q-Value 1991-2011.

QV	1991	2002	2011
Organic N per Ha (a)			
1	118.1	104.9	108.0
2	113.7	108.4	90.7
3	118.4	113.4	93.6
4	111.0	103.9	91.4
5	106.0	96.1	81.6
Total	113.1	107.0	92.0
Septic Tank Density (b)			
Q Value	1991	2002	2011
1	84.8	147.8	173.3
2	57.4	43.5	78.6
3	47.8	58.2	46.3
4	35.1	40.5	47.2
5	29.1	23.7	27.3
Total	39.7	46.3	47.5

Table 3
Share of Q-Values per year.

QV	1991	2002	2011
1	1.9	0.2	0.7
2	7.1	4.0	2.0
3	16.4	16.07	14.2
4	37.7	56.7	68.4
5	26.9	23.0	16.6

Source: EPA, Ireland, 2007

4.7. Summary statistics

Next we describe the summary statistics from the combined datasets. Table 1 details the main explanatory variables. A number of additional variables reflecting environmental characteristics were also included.

Table 2(a) describes the average organic N per hectare (ha)[4] by Q-Value point. Generally we see an inverse monotonic relationship between organic N per ha and Q-value between Q5 and Q3. However the relationship is not clear between Q3 and Q1. At all Q-values, there is a general decline, reflecting an overall reduction in animal numbers over the period. In the earlier period, the number of cattle increases, but this is offset by a decline in the number of dairy animals,which have a higher organic N per ha coefficient, in addition to a substantial decline in the sheep population.

The average septic tank density over time per Q value is reported in Table 2(b). The relationship is more or less inversely monotonic. Overall the septic tank density has increased. This increase is largely in areas with water quality values of Q4 and Q1 and Q2. This growth has occurred differentially for the poorest water quality areas, which have seen the highest increases in septic tank density.

5. Results and discussion

In this section, results are presented and organised as follows. First, in examining water quality trends, we present the share of Q-value by year, the distribution of organic N load and septic tank density over time and Q-values coinciding with Agricultural Census years. Next the results of the statistical model are presented and discussed in terms of land use and environmental drivers. The results of the sensitivity analysis of the model to different specifications and statistical tests and finally, some limitations arediscussed.

[4] Includes manure deposited on land by grazing animals and animal dung/slurry from housed animals

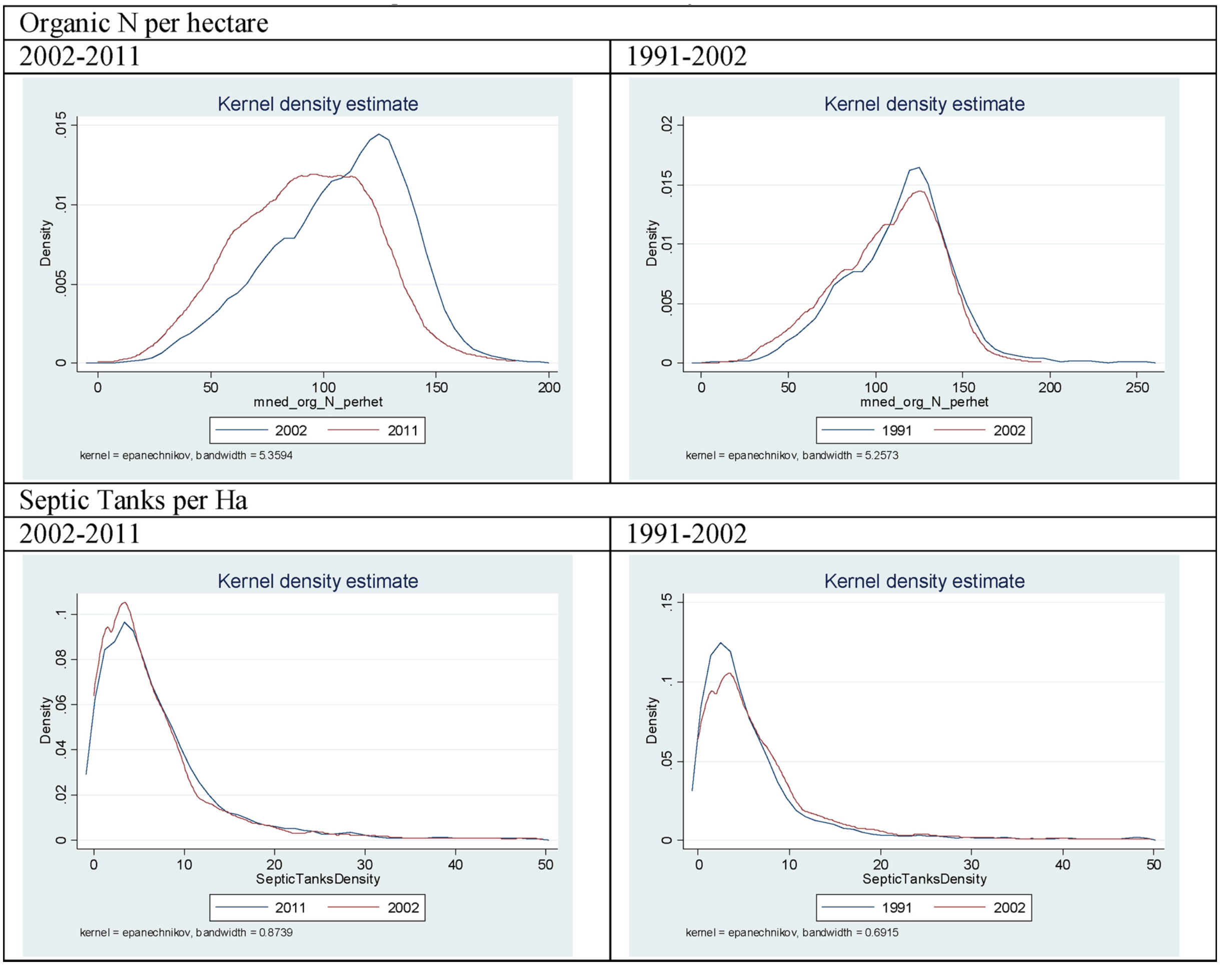


Fig. 2. Kernel Density Functions.

5.1. Trends in water quality

Table 3 describes trends in the distribution of water quality across the Quality Rating System (Q-value) for the three years considered in this study. The Q-value system has been used by the EPA to monitor the ecological quality of streams and rivers in Ireland since 1971 (McGarrigle et al., 2002). Over 3000 sites on some 13,200 km of main river channel are included in national monitoring to characterise water quality (EPA, 2006). An increase in the concentration of good (Q4) water quality is evident, increasing to 68.4 % in 2011, however there is a reduction in the percentage of Q5 pristine or reference sites. It is important to note that these values reflect the catchments that contain rural activity data. Q-value points that occur in urban only catchments are not included in this analysis.

In Fig. 2, we report the kernel density functions of the distribution of organic N per ha and the density of septic tanks. While the mean organic N per ha decreased between 1991 and 2002, the distribution did not change significantly with only a small increase in spread.[5] The change is more dramatic during the period 2002–2011, with a greater increase in spread and a clear reduction in mean. For septic tank density,[6] the change in the distribution is not significant over time, with most of the changes occurring at the extremes (this is not evident in the graphed kernel density as the top of the distribution is truncated to make it easier to see the mode).

Fig. 3 presents the spatial distribution of poor water quality over time, mapping spatial and inter-temporal trends in Q-values. We particularly note the incidence of water quality points with unsatisfactory values (red points denoting Q-value <4). Over time, the distribution of unsatisfactory points has become less dense.

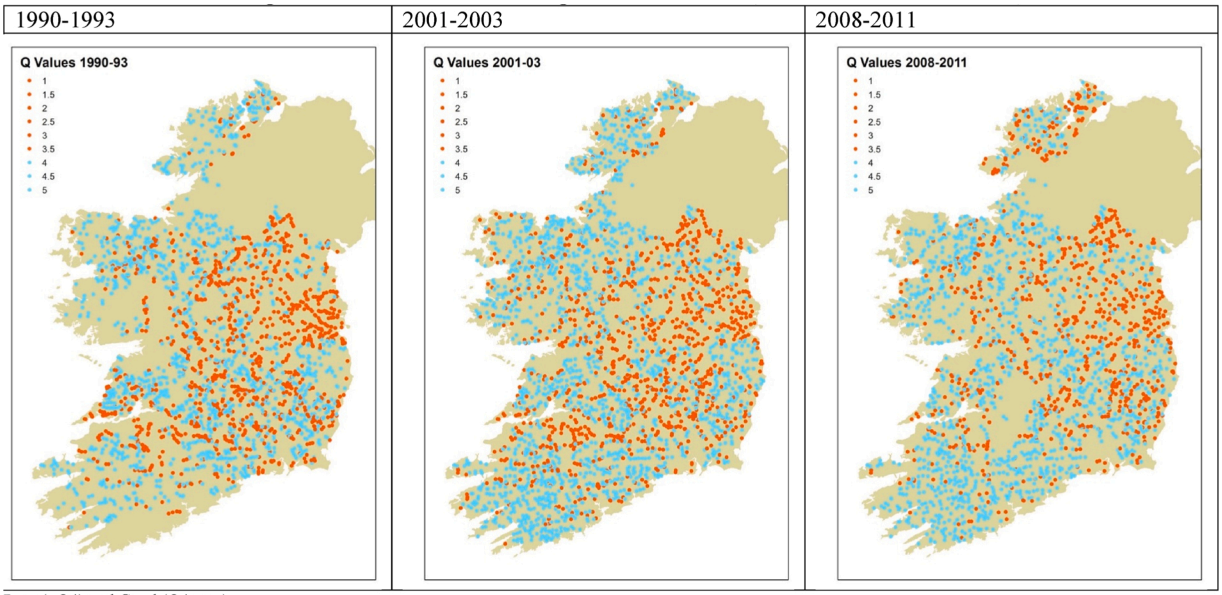
The distribution of unsatisfactory points has however shifted to the North and East over the period, (possibly reflecting intensive pig, poultry and other agricultural activities in combination with specific soil characteristics), with a reduction in the share of unsatisfactory points in the South and West. There is also an increase in unsatisfactory values in the commuting areas around Dublin and in coastal towns and estuaries, reflective of increases in population density, with consequent increases in industry and septic tank density.

5.2. Statistical model

As the dependent variable (Q-value) is categorically ordered, an ordered probit model was utilised, which takes the explanatory variables and estimates the probability of being in each category of water quality status (1–5). The model captures the underlying relationship between water quality and the main explanatory variables septic tank density and organic N per ha and incorporates time trends for the main explanatory variables and environmental variables. The model determines the major factors affecting the ecological quality of water sources measured at the EPA water quality monitoring points in each river catchment in 1991, 2002 and 2011. The model is estimated using the method of maximum likelihood via the Newton-Raphson algorithm

[5] It should be noted that when we refer to 2002 (2011), that Census of Agriculture numbers refer to 2000 (2010).

[6] Self-reported observations from Census of Population.



Poor (<Q4) and Good (Q4 or >)

Fig. 3. Q Values at times of Agricultural Census (1990-2; 2001-3; 2008-2011).

Table 4
Panel Ordered Probit Regression Model.

	Coefficient	S.E.
Septic Tank Density	−0.00027**	0.000102
Cereal Share of Land Use	−1.25738****	0.319362
Cereal Share of Land Use x 2002	0.549283	0.341399
Cereal Share of Land Use x 2011	0.684554**	0.340819
Organic N Density	−0.00265**	0.000908
Year Interaction		
Organic N Density x 1991	−0.00092	0.001067
Organic N Density x 2011	0.00271**	0.001133
Year Dummy		
2002	−0.27261**	0.129477
2011	−0.5252****	0.129326
Industry Share		
Industry	−0.07651	0.098335
Construction	0.045268	0.112441
Commerce	−0.24934**	0.1177
Transport, storage and communications	0.227862	0.183652
Public administration	0.009697	0.231106
Education, health and social	−0.09504	0.123385
Other	−0.2056	0.173893
Pigs per km2 (top quartile)	−0.14624**	0.173893
	0.000000711	0.000003
Cumulative Afforestation	0.000834****	0.000185
Landfill within 3km	−0.27005**	0.12948
Area of District	−0.0011266	0.000976
Rainfall	0.0000699	0.000163
Average Temperature	−0.11398**	0.0512
Median Elevation	0.00331****	0.00045
Mean Slope	0.1251****	0.01237
Time between River Quality Measurement and Census	−0.01099	0.00929
Coordinates		
x_coord	−0.000003****	0.000001
y_coord	0.000003**	0.000001
Environmental Variables	See Table 6	
Pseudo R^2	0.093	
N	7441	

**.

(Long, 1997). The coefficients of the ordered probit model indicate whether the explanatory variables are positively or negatively related to improved levels of water quality status (Long, 1997). The primary coefficient estimates and associated standard errors (SE) for the model specification and goodness of fit the fit of the model as measured by pseudo R^2 are presented in Table 4.

5.3. Land use drivers

5.3.1. Individual waste-water treatment

In Ireland and in the UK, wastewater from a significant proportion of the population (generally in rural areas) is treated by small-scale on-site individual waste-water treatment (IWWT) systems or septic tanks, where connection to a public sewer is not feasible. The design of many septic tanks reflects the historical legacy of the lack of wastewater disposal infrastructure. Due to problems in relation to poor design, management (leaks, lack of emptying) and being inappropriately sited (e.g. close to a watercourse), it is accepted that septic tanks can have a negative effect on water quality (Withers et al., 2011).

The results of the ordered probit show that the septic tanks variable is statistically significant at the one percent level with the anticipated sign i.e. the higher the density of septic tanks in the relevant EDs, the lower the value of the Q-value index at the monitoring point in the river catchment, reflecting the summary statistics presented earlier.

5.3.2. Agriculture

Pasture-based livestock enterprises, (beef, sheep and dairy) dominate Irish agriculture and account for approximately 80 % of overall agricultural output value. Livestock farming results in the deposition of organic N to land by grazing animals, as dung, urine or (liquid) slurry. Cereal growers generally use a range of inorganic fertilisers, herbicides and pesticides to increase productivity and to manage weeds, insects and diseases. Inappropriate or untimely applications of potential agricultural pollutants can increase the risk of run-off to water bodies.

In our analysis, the effects of the agricultural land use variables 'Cereal share of land use' and 'Organic N density' on water quality are similar. Both variables are significant and negative, indicating that water quality worsens as the share of cereals and organic N in EDs increases. These results are as expected and are consistent with other studies that examine the impact of agriculture on water quality (Monteny et al., 2001) and suggest that the higher the intensity of agriculture, the lower the likelihood of achieving a higher Q value.

In relation to the year dummy variables, the water quality in the 2002 and 2011 periods is significantly different to the base period of 1990. These time dummy variables are interacted with the organic N

Table 5
Panel Ordered Probit Regression Model with different specifications (excluding environmental variables).

	Model 1		Model 2		Model 3		Model 4		Model 5	
	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.	Coef.	S.E.
Septic Tank Density	−0.0004***	0.0001	−0.0003***	0.0001	−0.0004***	0.0001	−0.0003**	0.0001	−0.0002**	0.0001
Cereal Share of Land Use			−1.9121***	0.1187	−2.8962***	0.2877	−1.2574***	0.3194	−1.2075***	0.3206
Cereal Share of Land Use x 2002					0.7714**	0.3308	0.5493	0.3414	0.5401	0.3425
Cereal Share of Land Use x 2011					1.1499***	0.3349	0.6846**	0.3408	0.7394**	0.3427
Organic N Density (N per ha)	−0.0028***	0.0004	−0.0054***	0.0007	−0.0076***	0.0008	−0.0026**	0.0009	−0.0029**	0.0009
Year Interaction										
Organic N Density x 1991			0.001	0.0009	0.0015	0.0010	−0.0009	0.0011	−0.0009	0.0011
Organic N Density x 2011			0.0042***	0.0010	0.0036***	0.0011	0.0027**	0.0011	0.0032**	0.0011
Year Dummy										
2002			−0.026	0.1081	−0.0511	0.1167	−0.2726**	0.1295	−0.294**	0.1299
2011			−0.4608***	0.1083	−0.4635***	0.1213	−0.5252***	0.1293	−0.5814***	0.1299
Industry Share										
Industry					−0.2159**	0.0909	−0.0765	0.0983	−0.0433	0.0988
Construction					0.0811	0.1085	0.0453	0.1124	0.0765	0.1130
Commerce					−0.57***	0.1096	−0.2493**	0.1177	−0.2583**	0.1180
Transport, storage and communications					0.1238	0.1751	0.2279	0.1837	0.238	0.1842
Public administration					0.1457	0.2202	0.0097	0.2311	−0.0041	0.2314
Education, health and social					−0.4233***	0.1147	−0.095	0.1234	−0.0425	0.1240
Other					−0.3792**	0.1633	−0.2056	0.1739	−0.2327	0.1752
Pigs per km^2 (Top Quartile)					−0.0482	0.0646	−0.1462**	0.0711	−0.2554**	0.1152
Poutry per km^2					0	0.0000	0	0.000003	−0.0018*	0.0011
Cumulative Afforestation (ha)					0.001***	0.0002	0.0008***	0.0002	0.0906***	0.0187
Landfill within 3km					−0.2349*	0.1257	−0.2701**	0.1295	−0.2737**	0.1297
Soil Variables										
Pseudo R^2	0.0034		0.0176		0.0533		0.093		0.0983	
N	7954		7927		7471		7441		7411	

Note Statistically Significant * - 10 % level; ** - 5% level *** - 1% level.

and cereal share variables and show that relative to the base year of 1990, water quality is worse in 2002 but is better in 2011. Thus, notwithstanding the significant negative effect of agricultural activities on river water quality, it is important to note that our analysis indicates that this effect has significantly reduced over time. This could be a reflection of reduced levels of production over the period, as well as a variety of policy programmes, nutrient management measures and participation in agri-environment schemes, which have resulted in improved farm management practices. In addition, farmers invested €2.9 billion between 2005 and 2011 on improved farmyard management, animal housing and manure management facilities, and undertook more efficient use of both organic and inorganic fertilisers (Hennessy et al., 2013).

It has been widely reported that if statutory obligations in relation to water quality are to be met, then significant changes need to be undertaken by the agricultural sector throughout Europe (Haygarth et al., 2003; Bateman et al., 2006). As such, much recent research has investigated the effectiveness of various farm management mitigation measures in alleviating harmful impacts on water quality.

Within livestock enterprises, it has been found that N loss can be mitigated through changes in manure storage and manure application strategies (Chambers et al., 2000). For example, Lalor et al. (2011) report that 9% more N is available for plant uptake from manure if it is spread in spring (rather than in summer) and using targeted equipment i.e. using a trailing shoe as opposed to a splash plate. Thus better N uptake by grassland reduces the quantity available for loss to water.

Livestock dietary manipulation has also been shown to improve N use efficiency by animals, reducing N excretion and hence risk of N loss to water bodies (van Groenigen et al., 2006; Luo et al., 2008). Finally the use of cover crops has been shown to be very effective in terms of reducing N losses (Hooker et al., 2008). Additional conservation practices should be targeted where they can be most effective such as for cover crops, buffer strips and adaptive management of critical source areas (see Sharpley et al., 2000 for a review).

5.3.3. *Industry*

In examining the contribution of industry to water quality, we find that the variables 'Commerce', 'Pigs per km^2 (top quartile)' and 'Landfill within 3km' are significant and negative indicating that their presence has a negative impact on water quality. The commerce variable is indicative of the presence of a largely white-collar office-based sector. A negative relationship between population density and water quality has been widely reported. It is likely that as most commercial employment is located in urban areas, this variable is capturing the effect of population density. Thus thes results have more to do with the association with urban centres, than an explicit relationship with pollution associated with the sector. The relationship with intensive pig production and the presence of landfills close to EDs, is also consistent with *a priori* expectations and with the literature.

5.3.4. *Forest cover*

In terms of water quality, it might be expected that the level of forest cover in a catchment would impact on measured water quality either positively by acting as a filter, or negatively if there is ground disturbance associated with either forest establishment, management or timber harvesting operations, leading to potential sediment erosion and nutrient runoff. Interestingly, 'Cumulative Afforestation' is significant and positive, indicating that areas with greater forest cover tend to have better water quality. This could be due to a high incidence of historical planting on poorer more remote land. However, it could also be due to the nature of subsidies for afforestation during the period examined, which were weighted towards farmers and only payable on land that was formerly in agriculture, thus agriculture was often substituted with forestry.

This result is consistent with Novotny (2003) who suggests that undisturbed forests or woodland represent the best possible protection for land from sediment and pollutant losses. Woodlands and forests have low hydrological activity, due to high surface storage in leaves (interception), ground, mulch, and terrain roughness. Novotny (2003) also points out that even lowland forests with a high groundwater table, absorb large amounts of precipitation and actively retain water and

Table 6
Coefficients on Soil Variables in Model4 and interactions in Model 5.

	Model 4		Model 5	
	Coef.	S.E.	Coef.	S.E.
Soil Type				
Wind-blown sands undifferentiated Miscellaneous	0.0016**	0.000599	0.0017**	0.0040
Mineral alluvium Variable	0.0022***	0.000466	0.0021***	0.0000
Marl type soils Variable	0.0012*	0.0007018	0.0013*	0.0720
Derived from mainly non-calcareous parent materials Acid Brown Earths, Brown Podzolics	0.0009**	0.0004588	0.0009*	0.0540
Derived from mainly non-calcareous parent materials Surface water Gleys, Ground water Gleys	0.001**	0.0004591	0.001**	0.0360
Derived from mainly non-calcareous parent materials Peaty Gleys	0.0014**	0.0004669	0.0014**	0.0030
Derived from mainly non-calcareous parent materials Surface water Gleys (Shallow), Ground water Gleys (Shallow), Some outcropping rock	0.0015**	0.0006076	0.0014**	0.0230
Derived from mainly non-calcareous parent materials Peaty Gleys (Shallow)	0.0003	0.0006065	0.0001	0.8860
Predominantly shallow soils derived from non-calcareous rock or gravels with/without peaty surface horizon Podzols (Peaty), Lithosols, Peats, Shallow, lithosolic or podzolic type, Some outcropping rock	0.0008*	0.0004569	0.0008	0.1010
Derived from mainly non-calcareous parent materials Shallow Acid Brown, Earths/Brown Podzolics, Lithosols, Regosols, Some outcropping rock	0.0014**	0.0004672	0.0013**	0.0060
Blanket peat Blanket Peats	0.001**	0.0004598	0.001**	0.0330
Derived from mainly calcareous parent materials Grey Brown Podzolics	0.0009**	0.0004584	0.0009*	0.0560
Derived from mainly calcareous parent materials Surface water Gleys, Ground water Gleys	0.0008*	0.0004614	0.0007	0.1200
Derived from mainly calcareous parent materials Peaty Gleys	0.0012**	0.0004938	0.0011**	0.0240
Derived from mainly calcareous parent materials Surface water Gleys (Shallow), Ground water Gleys (Shallow), Some outcropping rock	0.001	0.0006034	0.0009	0.1190
Derived from mainly calcareous parent materials Peaty Gleys (Shallow)	−0.0009	0.0014254	−0.0005	0.7470
Predominantly shallow soils derived from calcareous rock or gravels with/without peaty surface horizon Lithosols, Peats, Some outcropping rock	0.0009	0.0007987	0.0006	0.4670
Derived from mainly calcareous parent materials Shallow Brown Earths/Grey Brown Podzolics, Rendzinas, Lithosols, Some outcropping rock	0.001**	0.0004639	0.001**	0.0370
Cutaway/cutover peat Basin Peats	0.001**	0.0004573	0.001**	0.0380
Fen peat Basin Peats	0.0007	0.0006233	0.0006	0.3290
Lacustrine–type soils Variable	−0.0001	0.000691	−0.0002	0.8260
Made/Built land Miscellaneous	−0.0005	0.0005611	−0.0007	0.2110
Beach sand and gravels Miscellaneous	0.001	0.0007624	0.0011	0.1610
Marine/ Estuarine sediments Miscellaneous	0.0012*	0.0006308	0.0011*	0.0700
Raised bog Basin Peats	0.028	0.0232984	0.0296	0.2050
Scree	0	0.0005135	0	0.9910
Water (including lakes, reservoirs and larger rivers) Miscellaneous	0.0002**	0.0000615	0.0002**	0.0060
Basalts & other Volcanic rocks	−0.0009*	0.0005022	−0.0009*	0.0900
Cambrian Metasediments	−0.0011**	0.0004636	−0.0011**	0.0220
Devonian Kiltorcan-type Sandstones	−0.001**	0.000469	−0.001**	0.0370
Devonian Old Red Sandstones	−0.001**	0.0004575	−0.001**	0.0370
Dinantian (early) Sandstones, Shales and Limestones	−0.001**	0.0004581	−0.0009**	0.0410
Dinantian Dolomitised Limestones	−0.001**	0.0004694	−0.001**	0.0360
Dinantian Lower Impure Limestones	−0.001**	0.0004599	−0.001**	0.0320
Dinantian Mixed Sandstones, Shales and Limestones	−0.0008*	0.0004665	−0.0008*	0.0950
Dinantian Pure Bedded Limestones	−0.0011**	0.0004627	−0.0011**	0.0210
Dinantian Pure Unbedded Limestones	−0.001**	0.0004567	−0.001**	0.0310
Dinantian Sandstones	−0.0009**	0.0004567	−0.0009*	0.0550
Dinantian Shales and Limestones	−0.0009**	0.0004627	−0.0009*	0.0510
Dinantian Upper Impure Limestones	−0.0009*	0.0004585	−0.0006	0.1810
Granites & other Igneous Intrusive rocks	−0.001**	0.0004594	−0.0008*	0.0870
Namurian Sandstones	−0.001**	0.0004577	−0.0009**	0.0460
Namurian Shales	−0.0011**	0.0004604	−0.0011**	0.0200
Namurian Undifferentiated	−0.0012**	0.0004632	−0.0013**	0.0050
Ordovician Metasediments	−0.0011**	0.0004584	−0.0011**	0.0160

(continued on next page)

Table 6 (*continued*)

	Model 4		Model 5	
	Coef.	S.E.	Coef.	S.E.
Ordovician Volcanics	−0.0011**	0.0004605	−0.0013**	0.0060
Permo-Triassic Mudstones and Gypsum	−0.0011**	0.0004646	−0.0011**	0.0220
Precambrian Marbles	0.0021	0.0026129	0.0021	0.4290
Precambrian Quartzites, Gneisses & Schists	0.0002	0.0011991	0.0004	0.7680
Silurian Metasediments and Volcanics	−0.0008*	0.0004704	−0.0009**	0.0490
Westphalian Sandstones	−0.0011**	0.0004572	−0.001**	0.0220
Westphalian Shales	−0.0011**	0.0004576	−0.001**	0.0240
Calcareous	−0.0014**	0.0005824	−0.0011*	0.0800
Non-calcareous (Siliceous)	−0.001**	0.0004834	−0.0012**	0.0140
Organic N Density x Dinantian Upper Impure Limestones			−0.0000018*	0.0570
Organic N Density x Ordovician Metasediments			0.0000018**	0.0050
Septic Tank Density x Dinantian (early) Sandstones, Shales and Limestones			−0.0000012	0.3850
Septic Tank Density x Dinantian Sandstones			0.0000017**	0.0180
Septic Tank Density x Dinantian Shales and Limestones			−0.0000055***	0.0000
Septic Tank Density x Namurian Shales			0.0000025***	0.0010
Septic Tank Density x Precambrian Marbles			0.0000072***	0.0010
Septic Tank Density x Precambrian Quartzites, Gneisses & Schists			−0.0000007***	0.0000
Septic Tank Density x Silurian Metasediments and Volcanics			0.0000004**	0.0130
Septic Tank Density x Westphalian Sandstones			−0.000024**	0.0020
Septic Tank Density x Westphalian Shales			0.0000208**	0.0060

Note Statistically Significant * - 10 % level; ** - 5% level *** - 1% level.

contaminants. In addition, environmental regulations around forest establishment and harvesting became more stringent over the period of this analysis and are enforced through requirements for subsidy payments and issuing of felling licences.

Even in managed forests, 'disturbances' that could potentially result in the loss of nutrients to water are infrequent, compared to agriculture. However while there are robust activity data in relation to afforestation and re-forestation (year of planting, species) for both public and private planting, forest harvesting events in private forests are not recorded officially. This paucity of forest harvesting activity data coupled with the sparse spatial and intertemporal nature of the water quality measurements, makes it difficult to pick up local impacts of forest disturbances and associated nutrient and sediment losses to water.

5.4. *Environmental drivers*

Due to the importance of the climatic and landscape context in relation to water quality outcomes, we also incorporate variables for rainfall, temperature, elevation and slope in the regression, where the coefficients for the range of soil (and underlying rock) variables presented in Table 6.

Table 7
Akaike and Bayesian Information Criteria for different models.

Model	Obs	ll(null)	ll(model)	df	AIC	BIC
1	7905	−11979.8	−11940.1	10	23900.29	23970.04
2	7905	−11979.8	−11787.7	15	23605.31	23709.93
3	7444	−11222	−10607.7	55	21325.43	21705.76
4	7411	−11165.8	−10113.1	116	20458.09	21259.74
5	7411	−11165.8	−10063.7	140	20407.44	21374.94

Overall, the fit of the model improves significantly through the addition of these contextual variables, as measured by the Pseudo R^2, which increases from 5 to 9%. (0.0533 to 0.093). While pseudo R^2 is typically small in non-linear models like the ordered probit, as opposed to continuous regression models, the pseudo R^2 value is not surprising in the water quality context, which is influenced by highly complex local hydrological systems. The purpose of this paper is to highlight the changing association over time in terms of economic and social drivers, captured by the statistical significance of the explanatory variables, albeit in a context where we cannot fully understand all influences on water quality outcomes.

Slope and elevation are both significant and positive, reflecting characteristics of remote areas where there are fewer pressures on water quality. Temperature is negatively significant, possibly reflecting poorer water quality in general, in the South, where temperatures are higher. The X (Y) coordinate coefficient is negative (positive), indicating poorer water quality in the East and South, where population centres and agricultural activity are more concentrated. Conditional on the XY coordinates, rainfall is not significant. We note that the time gap between water quality measurements (we use the most recent value after the relevant Census date) is not statistically significant, reflecting the strong inter-temporal auto-correlation in water quality monitoring. This means that water quality changes relatively slowly over time, so while there is some natural variation over time, it is not substantial. As a result using Census variables that differ by 0, 1 or 2 years do not affect the estimations of the models substantially.

5.5. *Sensitivity analysis*

In Table 5, we test the sensitivity of the model to different specifications, by building the model in stages. Model 1 is a simple model looking at the relationship between organic N density and septic tank density and water quality outcomes. In model 2, we time dummies are included to incorporate time trends. Model 3 incorporates other agricultural, industrial and population variables, while model 4 additionally incorporates environmental variables as in the case of the model already reported above. Model 5 introduces additional interaction terms between economic (organic N and septic tank density) and environmental (soils) variables, reporting only those interactions that have statistical significance.

In terms of sign and significance, each of the models reports the same qualitative story. Both septic tanks and organic N have a statistically significant negative coefficient indicating a positive correlation between these drivers and unsatisfactory water quality. When the time dummies are introduced, the coefficients remain positive and significant for organic N for 2011 and not significant for 1991, indicating an improvement between 2002 and 2011.

In terms of the level of the septic tank variable, there is no significant difference between models. In each case time dummies on the septic tank variable is not significant. The incorporation of time interactions, perhaps unsurprisingly, changes the average coefficient for the organic N. As expected, the inclusion of the environmental variables increases the explanatory power and also serves to visually highlight the complexity of the hydrological influences. Most importantly however, the qualitative conclusions are robust to their inclusion.

As expected, the addition of the interaction terms further improves the goodness of fit of the model. However we would like to test optimal model complexity. Table 7 reports the Akaike (AIC) and Bayesian (BIC) information criteria for different models. The AIC measure reduces as explanatory variables are added to each model, indicating growing efficiency of the model. For the BIC measure however, model 4 has the lowest value indicating a turning point and lower efficiency from model 5, evidence that the interaction terms may not add much extra information to the model.

5.6. Limitations

This model does not attempt to be a deterministic hydrological model, which would require different functional forms in different spatial areas. Such models would also require spatially lagged dependent variables, given the correlation across river flows. Rather the paper is a policy assessment, attempting to understand the statistical correlation between economic, social and land use drivers and their changes over time. We undertook sensitivity analysis in terms of a number of model estimations using the variables at water quality monitoring points, rather than using upstream data and found no substantial changes to the model estimates. Given this spatial autocorrelation, it is unsurprising that explanatory variables are related to each other in the immediate vicinity and thus the results are robust to this assumption. Incorporating spatial lag variables however, poses particular estimation problems. Because the lag variable (in the same river system) contains a lot of the same information as the other monitoring points in the river system, they radically increase the explanatory power of the model, but may affect our interpretation of the relationship with other explanatory variables. Incorporating lag dependent variables thus introduces important biases into the statistical models. Utilising GMM[7] estimators may assist in this, but such an analysis might be more suitable for a tudy examining spatial dynamics in water quality.

6. Conclusions

River water quality is affected by a combination of geomorphological (e.g. soil type, slope, elevation), climatic (e.g. precipitation) and anthropogenic factors (e.g. agricultural practices, forest cover, landfills and IWWT (septic tanks). The objective of this study was to examine long-term trends in the statistical relationships between anthropogenic and natural factors on water quality as measured for WFD monitoring. The primary contribution of this work is that it goes beyond previous studies that do not capture either the variety of economic drivers on the one hand or capture trends in water quality and their national level drivers on the other.

Although the main focus is the relationship between anthropogenic and water quality measures, we find that the addition of environmental variables adds to the explanatory power of the models, albeit with time-lags between losses of nutrients or pollutants to water bodies and any subsequent impact on water quality. The analysis shows that the rural land uses with the largest negative impacts on water quality over the period 1991–2011 are nutrient loads from agriculture and IWWT, consistent with *a priori* expectations and robust across all models.

While policy initiatives have been introduced addressing the design and management of septic tanks, with increasing rural populations, the design and siting of IWWT systems needs to continue to be improved. In relation to agriculture, there is widespread acceptance that measures to reduce the impact of agriculture on water will be strengthened in future iterations of EU legislation and agricultural policies, further enhancing the direction of travel in impact.

The period of this analysis (1991−2011) covers a period of lower agricultural productivity following limits imposed on dairy production, through decoupling of payments and through the introduction of regulation and incentives to improve agricultural management. While water quality improved and the coefficient on organic N (agriculture) fell significantly between 2002 and 2011 as the environmental production function of agriculture became more efficient (particularly in dairy systems) (Hennessy et al., 2013), the improvement in water quality may not continue, as a result of increases in dairy cow numbers post 2015, on foot of the removal of the dairy quota. This analysis could be extended to cover this period once the 2020 Census of Agriculture data are available.

To conclude, no one sector is responsible for adverse river water quality and as a result, solutions will depend on a multi-sectoral approach aimed at addressing the human and land-use factors affecting water quality. While significant measures to maintain 'good' water status under the WFD have been implemented across all the land-use sectors examined in this analysis, there is likely to be additional pressure on the agricultural sector in future due to increasing agri-food production to meet growing global demands and commitments for greenhouse gas reductions under the European Green Deal[8] climate change legislation. In this regard, it is hoped that the analytical framework developed in this paper will provide an important tool in the ongoing identification and management of risks and will allow for appropriate land use planning aimed at restoring and maintaining water quality as required by the WFD.

CRediT authorship contribution statement

Cathal O'Donoghue: Conceptualization, Formal analysis, Writing - original draft, Writing - review & editing, Supervision. **Cathal Buckley:** Writing - original draft. **Aksana Chyzheuskaya:** Writing - original draft. **Stuart Green:** Data curation. **Peter Howley:** Conceptualization, Writing - original draft, Funding acquisition. **Stephen Hynes:** Conceptualization, Methodology, Writing - original draft. **Vincent Upton:** Data curation. **Mary Ryan:** Writing - review & editing.

Appendix A

Agricultural Data - The Irish Census of Agriculture

A key requirement in a geographic assessment of the contribution of agriculture to water pollution is the availability and resolution of spatial data pertaining to individual sub-sectors. The main source of diffuse pollution from grass-based sectors such as livestock rearing, arises from the release of nitrous oxide. The main sources of nitrous oxide are nitrogen fertilisers and organic manure and urea deposited by grazing animals (Monteny et al., 2006). In contrast to extensive grass-based farm activities, cereal production requires much larger applications of chemical fertilisers.

The Irish Census of Agriculture collects data relating to agricultural activities on all farms within Ireland (CSO, 2002). The lowest level of spatial disaggregation for publicly available data is at the Electoral Division (ED)[9] level, of which 2850 out of 3440 EDs contain farms, with an average of 53 farms in each of these (min 10, max 320).

In order to represent agricultural activity, we utilise the proportion of farmland in each ED under crops, the number of pigs per hectare in each ED and the livestock density in each ED. The figures for livestock density were combined with livestock nutrient excretion factors (see Table 7) to produce estimates of organic N load per ED, which provides an indication of the intensity of agriculture in individual EDs. The

[7] Generalized method of moments (Lück and Wolf, 2016).

[8] https://ec.europa.eu/info/strategy/priorities-2019−2024/european-green-deal_en">https://ec.europa.eu/info/strategy/priorities-2019−2024/european-green-deal_en.

[9] Electoral Divisions (EDs) are the smallest legally defined administrative areas in the State for which Small Area Population Statistics (SAPS) are published from the Census.

Teagasc National Farm Survey (NFS, 2010) shows that dairy is the most intensive, with the majority of farms over 170 kg organic N per hectare while 80 % of cattle rearing farms are under 90 kg.

Forest Cover Data

To provide information on the level of forest cover within each ED, a land cover classification for Ireland developed by Teagasc under the Forest Inventory Planning System and Irish Forest Soils (FIPS–IFS)[10] project was used. The FIPS–IFS land-cover data set was developed using GIS and remote sensing, along with ground-truthing provided by field sampling. The mapping unit employed in the FIPS–IFS land-cover data set was one hectare. The main class in the FIPS–IFS land-cover data set included in this analysis is a combined variable for mature forestry, immature forestry and scrub. This forest cover GIS data was updated by Farrelly (2007) using annual afforestation statistics to reflect increasing private forest cover over time. As data are not collected on forest harvesting events in the private forest sector in Ireland, we focus here on cumulative forest cover over time.

Census of Population, Small Area Population Statistics (SAPS)

The Central Statistics Office (CSO) collects data in the Census of Population pertaining to the structure and services to residential dwellings in Ireland, including the number of rooms per house, toilet facilities and sewerage facilities in each ED. In relation to sewerage facilites, the EPA (2006) found that the presence of septic tanks, (the main method for wastewater treatment in rural households), has a significant negative impact on water quality. Therefore a variable representing the proportion of households in each ED that have septic tanks was included in the analysis. In the relevant Census years, we see that there are respectively 34,5900, 40,7768, 437,652 houses with septic tanks in 1991, 2002 and 2011, with the density increasing by 20 % over the period of the study, concentrated in the first 11 years. To examine the effect of septic tanks on river water quality using small area population statistics (SAPS), we derive a variable indicating the number of septic tanks per km^2 in each ED.

The SAPS dataset classifies all workers within each ED under eight industry types: Agriculture, Forestry and Fishing, Manufacturing, Construction, Commerce, Transport and Communications, Public Administration, Education, Health and Social work or Other Industry. We use these data as a proxy for economic activity. By combining the agricultural, forestry and census data described above with the associated Q values for the EPA monitoring stations, it is possible to examine the major economic factors affecting river water quality.

Other Environmental Data

Historically the Republic of Ireland has had a heavy reliance on landfills for the disposal of waste products. Landfills in Ireland were brought under the regulatory control of the EPA under the Environmental Protection Agency Act, 1992 and the Waste Management Act, 1996 (EPA, 2011a). The 1999 Landfill Directive (Council Directive 99/31/EC) was a major milestone in the regulation of landfills in the EU, as it specified the technical requirements for landfill design, operation, closure and aftercare. Leachate from landfills can pose a risk to water quality. Compliance with legislation is assessed by the EPA through the completion of site inspections, audits and monitoring of emissions and the quality of the environment. The location and co-ordinates of 77 landfill facilities in Ireland was secured from the EPA based on their licensing regime (EPA, 2011b), allowing us to create an indicator of whether a landfill facility was within 3 km upstream of a Q value monitoring site.

[10] https://t-stor.teagasc.ie/handle/11019/543.

Table A1
Annual nutrient excretion rates for livestock.

Livestock type	Total Nitrogen kg/year	Total Phosphorus kg/year
Dairy cow	85	13
Suckler cow	65	10
Cattle (0–1 year old)	24	3
Cattle (1–2 years old)	57	8
Cattle > 2 years	65	10
Lowland ewe & lambs	13	2

Source: ISB (2014): Table 6- Good Agricultural Practice for Protection of Waters Regulations 2014.

Table A2
Distribution of Organic N by Farm System, 2010.

Organic N (kg per ha)	0–89	90–169	170–209	> = 210	Total
System					
Dairy	0.044	0.659	0.230	0.067	1.000
Cattle Rearing	0.801	0.195	0.004	0.000	1.000
Cattle Other	0.683	0.310	0.003	0.003	1.000
Sheep	0.989	0.011	0.000	0.000	1.000
Mixed	0.222	0.653	0.099	0.308	1.000
Total	0.542	0.351	0.066	0.027	1.000

Source: Teagasc National Farm Survey (2010).

Environmental and physical characteristics of watersheds can also play an important role in water quality (Donohue et al., 2005, 2006). Bedrock data from the Geological Survey of Ireland (GSI) (1:100,000 bedrock shapefile (GSI, 2013)) and soil data from the Teagasc EPA soil and subsoil map (Fealy et al., 2009) provided data on geological and soil characteristics. A digital elevation model (DEM) at a 25 m resolution was employed to derive a series of elevation-related variables. A slope map was generated from the DEM at the same resolution. Climactic data were derived from models developed by Sweeney and Fealy (2003). Polygon-based data were intersected with the ED shapefiles to derive the area of soil and bedrock categories in each ED. For raster data, the average and median values were calculated across each ED.

The relevant upstream area for each water quality monitoring point was derived from river sub-basins maps. establish. Each point was joined to the river sub-basin that it fell within. Shapefiles describing the monitoring stations and the river sub-basins were attained from the EPA (http://gis.epa.ie/GetData/Download). Thus, a dataset relating Q-values from individual monitoring stations to the characteristics of the relevant river sub-basin and related upstream EDs was created. Upstream was determined on the basis of elevation. EDs with centroid elevation greater than the centroid elevation of the ED in which the Q-value monitoring point was recorded, were deemed to be upstream of the monitoring point and average upstream values were derived for the analysis.

See Tables A1 and A2.

Appendix B. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.landusepol.2021.105322.

References

Arellano, M., 2003. Panel Data Econometrics. Oxford University Press.
Bateman, I.J., Brouwer, R., Davies, H., Day, B.H., Deflandre, A., Di Falco, S., Georgiou, S., Hadley, D., Hutchins, M., Jones, A.P., Kay, D., Leeks, G., Lewis, M., Lovett, A.A., Neal, C., Posen, P., Rigby, D., Turner, R.K., 2006. Analysing the agricultural costs and non-market benefits of implementing the water framework directive. J. Agric. Econ. 57 (2), 221–237.
Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Brouwer, R., Hofkes, M., Linderhof, V., 2008. General equilibrium modelling of the direct and indirect economic impacts of water quality improvements in the Netherlands at national and river basin scale. Ecol. Econ. 66 (1), 127–140.

Chambers, B.J., Smith, K.A., Pain, B.F., 2000. Strategies to encourage better use of nitrogen in animal manures. Soil Use Manag. 16, 157–161.

Chang, H., 2008. Spatial analysis of water quality trends in the Han River basin, South Korea. Water Res. 42 (13), 3285–3304.

CSO, 2002. Census of Agriculture Main Results, 2000. Central Statistics Office, Dublin.

Curtis, J., Morgenroth, E., 2013. Estimating the effects of land-use and catchment characteristics on lake water quality: irish lakes 2004-2009. J. Stat. Soc. Inquiry Soc. Ireland 42.

Cuttle, S.P., Haygarth, P.M., Chadwick, D.R., Newell-Price, P., Harris, D., Shepherd, M. A., Chambers, B.J., Humphrey, R., 2006. An Inventory of Measures to Control Diffuse Water Pollution from Agriculture. Report to Defra.

Daou, C., Salloum, M., Legube, B., Kassouf, A., Ouaini, N., 2018. Characterization of spatial and temporal patterns in surface water quality: a case study of four major Lebanese rivers. Environ. Monit. Assess. 190 (8), 485.

Dillon, E.J., Hennessy, T., Hynes, S., 2010. Assessing the sustainability of Irish agriculture. Int. J. Agric. Sustain. 8 (3), 131–147.

Donohue, I., Styles, D., Coxon, C., Irvine, K., 2005. Importance of spatial and temporal patterns for assessment of risk of diffuse nutrient emissions to surface waters. J. Hydrol. 304, 183–192.

Donohue, I., McGarrigle, M.L., Mills, P., 2006. Linking catchment characteristics and water chemistry with the ecological status of Irish rivers. Water Quality 40, 91–98.

Elith, J., Leathwick, J.R., Hastie, T., 2008. A working guide to boosted regression trees. J. Anim. Ecol. 77, 802–813.

Environmental Protection Agency, 2006a. Water Quality in Ireland. Aquatic Environment, Office of Environmental Assessment.

Environmental Protection Agency, 2006b. Eutrophication From Agricultural Sources - Effects of Agricultural Practices on Nitrate Leaching. https://www.epa.ie/pubs/reports/research/water/ERTDI-58-web-ready1.pdf.

Farrelly, N., 2007. The farm forest resource in Ireland: opportunities and challenges for rural development in Ireland. Small-scale For. 6, 49–64.

Fealy, R.M., Green, S., Loftus, M., Meehan, R., Radford, T., Cronin, C., Bulfin, M., 2009. Teagasc EPA Soil and Subsoils Mapping Project-Final Report, vol. II. Teagasc, Dublin.

Ferrier, Robert C., Edwards, Anthony C., Hirst, David, Littlewood, IanG., Carol, D., Bob Morris, R., 2001. Water quality of Scottish rivers: spatial and temporal trends. Sci. Total Environ. 265 (1–3), 327–342.

Fezzi, C., Hutchins, M., Rigby, D., Bateman, I.J., Posen, P., Hadley, D., Deflandre-Vlandas, A., 2008. Integrated Assessment of Water Framework Directive Nitrate Reduction Measures. CSERGE Working Paper ECM 08-02.

Garizi, A.Z., Sheikh, V., Sadoddin, A., 2011. Assessment of seasonal variations of chemical characteristics in surface water using multivariate statistical methods. Int. J. Environ. Sci. Technol. 8 (3), 581–592.

Geological Society of Ireland, 2013. Bedrock Geology Shapefile. http://www.dcenr.gov.ie/Spatial+Data/, Geological+Survey+of+Ireland/ GSI+Spatial+Data+Downloads.htm. Accessed 10:03:13. 1:100,000.

Haygarth, P., Johnes, P., Butterfield, D., Foy, B., Withers, P., 2003. Land Use for Achieving 'Good Ecological Status' of Waterbodies in England and Wales: A Theoretical Exploration for Nitrogen and Phosphorus. Report to DEFRA, IGER, Oakhampton.

Hennessy, T., Shalloo, L., Dillon, P., 2005. The economic implications of complying with a limit on organic nitrogen in a decoupled policy environment–an Irish case study. J. Farm Manage. 12 (6), 297–311.

Hennessy, T., Buckley, C., Dillon, E., Donnellan, T., Hanrahan, K., Moran, B., Ryan, M., 2013. Measuring Farm Level Sustainability with the Teagasc National Farm Survey. Teagasc, Oak Park, Carlow. https://www.teagasc.ie/media/website/publications/2013/SustainabilityReport.pdf.

Hooker, K.V., Coxon, C.E., Hackett, R., Kirwan, L.E., O'Keeffe, E., Richards, K.G., 2008. Evaluation of cover crop and reduced cultivation for reducing nitrate leaching in Ireland. J. Environ. Qual. 37, 138–145.

ISB, 2014. S.I. No. 31/2014 - European Union (Good Agricultural Practice for Protection of Waters) Regulations 2014. Irish Statute Book. http://www.irishstatutebook.ie/eli/2014/si/31/made/en/print.

Jokinen, C.C., Edge, T.A., Koning, W., Laing, C.R., Lapen, D.R., Miller, J., Mutschall, S., Scott, A., Taboada, E.N., Thomas, J.E., Topp, E., 2012. Spatial and temporal drivers of zoonotic pathogen contamination of an agricultural watershed. J. Environ. Qual. 41 (1), 242–252, 1.

Lalor, S.T.J., Shröder, J.J., Lantinga, E.A., Oenema, O., Kirwan, L., Schulte, R.P.O., 2011. N fertilizer replacement value of cattle slurry in grassland as affected by method and timing of application. J. Environ. Qual. 40 (2), 362–373.

Larned, S.T., Scarsbrook, M.R., Snelder, T.H., Norton, N.J., Biggs, B.J.F., 2004. Water quality in low-elevation streams and rivers of New Zealand: recent state and trends in contrasting land-cover classes. N. Z. J. Mar. Freshwater Res. 38, 347–366.

Lennox, S., Foy, R., Smith, R., Unsworth, E., Smyth, D., 1998. A comparison of agricultural water pollution incidents in Northern Ireland with those in England and Wales. Water Res. 32, 649–656.

Li, J., Meng, X., Zhang, Y., Li, J., Xia, L., Zheng, H., 2015. Analysis of the temporal and spatial distribution of water quality in China's major river basins, and trends between 2005 and 2010. Front. Earth Sci. 9 (3), 463–472.

Liu, J., Zhang, X., Xia, J., Wu, S., She, D., Zou, L., 2016. Characterizing and explaining spatio-temporal variation of water quality in a highly disturbed river by multi-statistical techniques. Springer Plus 5, 1171.

Long, J.S., 1997. Regression Models for Categorical and Limited Dependent Variables. CA: Sage, Thousand Oaks.

Lück, A., Wolf, V., 2016. Generalized method of moments for estimating parameters of stochastic reaction networks. BMC Syst. Biol. 10, 98. https://doi.org/10.1186/s12918-016-0342-8.

Luo, J., Ledgard, S.F., Lindsey, S.B., 2008. A test of a winter farm management option for mitigating nitrous oxide emissions from a dairy farm. Soil Use Manag. 24, 121–130.

Lynch, John, Hennessy, Thia, Buckley, Cathal, Dillon, Emma, 2016. Teagasc National Farm Survey 2015 Sustainability Report. Teagasc, Athenry, Co. Galway.

Mainali, J., Chang, H., 2018. Landscape and anthropogenic factors affecting spatial patterns of water quality trends in a large river basin, South Korea. J. Hydrol. 564, 26–40.

McGarrigle, M.L., Bowman, J.J., Clabby, K.J., Lucey, J., Cunningham, P., MacCarthaigh, M., Keegan, M., Cantrell, B., Lehane, M., Clenaghan, C., Toner, P., 2002. Water Quality in Ireland 1998–2000. Environmental Protection Agency Report, Ireland.

Mei, K., Liao, L., Zhu, Y., Lu, P., Wang, Z., Dahlgren, R.A., Zhang, M., 2014. Evaluation of spatial-temporal variations and trends in surface water quality across a rural-suburban-urban interface. Environ. Sci. Pollut. Res. - Int. 21 (13), 8036–8051.

Micha, E., Roberts, W., Ryan, M., O'Donoghue, C., Daly, K., 2018. A participatory approach for comparing stakeholders' evaluation of P loss mitigation options in a high ecological status river catchment. Environ. Sci. Policy 84, 41–51.

Mockler, E.M., Deakin, J., Archbold, M., Gill, L., Daly, D., Bruen, M., 2017. Sources of nitrogen and phosphorus emissions to Irish rivers and coastal waters: estimates from a nutrient load apportionment framework. Sci. Total Environ. 601, 326–339.

Murphy, P.N., Mellander, P.E., Melland, A.R., Buckley, C., Shore, M., Shortle, G., Wall, D. P., Treacy, M., Shine, O., Mechan, S., Jordan, P., 2015. Variable response to phosphorus mitigation measures across the nutrient transfer continuum in a dairy grassland catchment. Agric. Ecosyst. Environ. 1 (207), 192–202.

Novotny, V., 2003. Water Quality. Diffuse Pollution and Watershed Management. John Willey & Sons, Inc., Hoboken, New Jersey.

Rheuban, J.E., Williamson, S.C., Costa, J.E., Glover, D.M., Jakuba, R.W., McCorkle, D.C., et al., 2015. Spatial and temporal trends in summertime climate and water quality indicators in the coastal embayments of Buzzards Bay, Massachusetts. Biogeosciences Discuss. 12 (15).

Segurado, P., Almeida, C., Neves, R., Ferreira, M.T., Branco, P., 2018. Understanding multiple stressors in a Mediterranean basin: combined effects of land use, water scarcity and nutrient enrichment. Sci. Total Environ. 624, 1221–1233.

Sharpley, A., Foy, B., Withers, P., 2000. Practical and innovative measures for the control of agricultural phosphorus losses to water: an overview. J. Environ. Qual. 29 (1), 1–9.

Singh, S., Chang, H., 2014. Effects of land cover change on water quality in urban streams at two spatial scales. Int. J. Geospatial Environ. Res. 1 (1), 8.

Sweeney, J., Fealy, R., 2003. Establishing Reference Climate Scenarios. Chapter 2. In Climate Change Scenarios and Impacts for Ireland, Final report. Environmental Protection Agency. Johnstown Castle Estate, Co. Wexford, Ireland, pp. 5–31.

van Groenigen, J.W., Koola, D.M., Velthofa, G.L., Oenemaa, O., Kuikmana, P.J., 2006. Int. Congr. Ser. 1293, 347–350.

Varanka, S., Luoto, M., 2012. Environmental determinants of water quality in boreal rivers based on partitioning methods. River Res. Appl. 28, 1034–1036.

Vatn, A., Bakken, L.R., Botterweg, P., Lundeby, H., Romstad, E., Rørstad, P.K., Vold, A., 1997. Regulating nonpoint source pollution from agriculture: an integrated modelling analysis. Eur. Rev. Agric. Econ. 24, 207–229.

Vatn, A., Bakken, L.R., Bleken, M., Baadshaug, O.H., Fykse, H., Haugen, L.E., Lundekvam, H., Morken, J., Romstad, E., Rørstad, P.K., Skjelvåg, A.O., Sogn, T., 2006. A methodology for integrated economic and environmental analysis of pollution from agriculture. Agric. Syst. 88, 270–293.

Volk, M., Hirschfeld, J., Dehnhardt, A., Schmidt, G., Bohn, C., Liersch, S., Gassman, P.W., 2008. Integrated ecological-economic modelling of water pollution abatement management options in the Upper Ems River Basin. Ecol. Econ. 66, 66–76.

Wang, Q., Wu, X., Zhao, B., Qin, J., Peng, T., 2015. Combined multivariate statistical techniques, water pollution index (WPI) and Daniel trend test methods to evaluate temporal and spatial variations and trends of water quality at Shanchong River in the Northwest Basin of Lake Fuxian, China. PLoS One 10 (4), e0118590.

Wang, X., Zhang, F., 2018. Multi-scale analysis of the relationship between landscape patterns and a water quality index (WQI) based on a stepwise linear regression (SLR) and geographically weighted regression (GWR) in the Ebinur Lake oasis. Environ. Sci. Pollut. Res. - Int. 25 (7), 7033–7048.

Withers, P.J., Haygarth, P., 2007. Agriculture, phosphorus and eutrophication: a European perspective. Soil Use Manag. 23 (1), 1–4.

Withers, P.J.A., Jarvie, H.P., Stoate, C., 2011. Quantifying the impact of septic tank systems on eutrophication risk in rural headwaters. Environ. Int. 37, 644–653.

Woli, K.P., Hayakawa, A., Kuramochi, K., Hatano, R., 2008. Assessment of river water quality during snowmelt and base flow periods in two catchment areas with different land use. Environ. Monit. Assess. 137, 251–260.