CREW CENTRE OF EXPERTISE FOR WATERS

Lags in water quality response to diffuse pollution control measures: a review

CENTRE OF EXPERTISE FOR WATERS

Lags in water quality response to diffuse pollution control measures: a review

Ioanna Akoumianaki

Hydro Nation International Centre
The James Hutton Institute
Scottish Government
Riaghaltas na h-Alba
gov.scot

CENTRE OF
EXPERTISE
FOR WATERS

Published by CREW – Scotland's Centre of Expertise for Waters. CREW connects research and policy, delivering objective and robust research and expert opinion to support the development and implementation of water policy in Scotland. CREW is a partnership between the James Hutton Institute and all Scottish Higher Education Institutes and Research Institutes supported by MASTS. The Centre is funded by the Scottish Government.

Authors: Ioanna Akoumianaki
James Hutton Institute, Invergowrie, Dundee, Scotland UK

CREW Project Manager: Nikki Dodd

Please reference this report as follows: I. Akoumianaki. Lags in water quality response to diffuse pollution control measures: a review. Project Code: CRW2018_19. Available online at: crew.ac.uk/publications.

ISBN: 978-0-902701-82-3

Dissemination status: Unrestricted

Review: Marc Stutter

Acknowledgements: The project lead wishes to acknowledge the constructive ideas in the delivery of the project provided by the steering group: Brian McCreadie, Fiona Napier, Adam Carson, Darrell Crothers, Susan Campbell, Alison Bell (SEPA) and David Lister (Scottish Government). Many thanks to Marc Stutter (James Hutton Institute) who reviewed the final draft report and participated in the final meeting on 1st September 2020.

Cover photos (left - right): https://extension.usu.edu/waterquality/powerpoint/; https://en.wikipedia.org/wiki/Riparian_buffer licensed under CC BY-SA

Contents

List of figures

List of tables

List of boxes

List of Tables in APPENDIX

Executive Summary

Research questions

1. What key catchment processes influence lags in water quality response to diffuse pollution control measures (hereafter the measures)?
2. What (inter)national evidence base is available on lags in water quality response to measures for each type of measure and pollutant, i.e. *total phosphorus (P), soluble reactive inorganic phosphorus (SRP), total nitrogen (TN), nitrate, faecal indicator organisms (FIO), and sediment*?
3. Is it possible to define/identify catchment typologies in Scotland to estimate lags in water quality response for each pollutant and type of measure? If not, why not?

Background

SEPA implement regulatory and incentivised measures to protect and improve water quality. The intended effects of the various measures implemented are to: (i) avoid or reduce inputs of pollutants at source (*Source control*); (ii) control / delay transport of pollutants in-field (in field *Transport control*); and (iii) trap pollutants before they reach waterbodies (riparian *Trapping*). Estimating lags in water quality response to measures (Definition 1) based on catchment typologies (Definition 2) could help SEPA to improve diffuse pollution control and communicate to stakeholders the causes of the perceived lack of response to measures in waterbodies that have not improved yet.

Research undertaken

This project undertook a systematic review of the literature on water quality response and lags in response to the measures implemented. Overall findings were discussed in a sense-checking workshop. This Executive Summary (ES) presents the key findings and policy recommendations.

Key Findings

1. There is paucity of empirical evidence and lack of understanding about precisely how long it takes a water quality response to measures to occur, whether it be the first detectable improvement or the trajectory to the first response or to the end-point of compliance with water quality standards. There is no evidence that fixed timeframes for a water quality response to measures can be set. Predicted lags in water quality response based on a catchment typology were not found in the literature. Long-term water quality and catchment data are key to quantifying lags.
2. Lack of water quality response to measures was attributed to combinations of the reasons below:
   - Uncertainties about the effectiveness of measures and the level of implementation required for a water quality response;
   - Low efficiencies of the measures in the context of background catchment variability and pressures such as climate change, which translate to small, undetectable improvements;
   - Lack of effectiveness due to non-optimal implementation of the measures;
   - Longer time required for the measures to become fully effective;
   - Variable function and performance of the measures in response to environmental conditions;
   - Lack of appropriate long-term water quality and catchment data to account for catchment factors;
   - Poor understanding of the start time of the post-implementation period at the catchment scale, which affects statistical analyses and study designs;
   - Monitoring design, which may be introducing a statistical lag, or is unable to detect the magnitude of improvement that has occurred or can occur under site-dependent circumstances.
3. Studies that observed a water quality response to measures, found that lags broadly increase with:
   - Catchment size (Fig. ES1) but differ for the same pollutant in catchments with similar size/ measures;
   - Legacy effects from past pollutant inputs stored in the soils and in-stream;
   - Travel time from sources to receptors, e.g. when groundwater hydrologic pathways predominate;
   - Residence time in-field, in-stream and in the aquifer, which is generally enhanced by storage;
   - Presence of multiple, non-agricultural sources (e.g. wastewater discharges) of that pollutant.
4. Lags reported in temperate regions were in the range of 1-25 years for river waterbodies (Fig. ES1. A) and potentially longer than 20 years for a groundwater nitrate response. Studies based on a Before-After/ Control-Impact (BACI) design used 2-4 years of baseline data. The Trend design involved long-term (over ten years) monitoring and comparisons between catchments and was used in cases of gradual implementation of the measures across medium-sized catchments (size: 20-300km^2). Based on long-term data, river and groundwater water quality trajectories of response to measures are subject to site-dependent seasonal, interannual and decadal, climate-related, variation.

5. Studies that reported a water quality response within five years post-implementation of the measures (Fig.ES1) attributed the relatively fast response to optimal implementation, i.e. extensive and spatially integrated implementation, targeting to match pressures to biogeochemical and hydrologic processes at farm scale and application of a combination of *Source control*, *Transport control* and *Trapping* across the landscape.
6. No catchment typologies for lags in water quality response were identified because of:
   - Complexity: Multiple interacting catchment factors involved in diffuse pollution control;
   - Paucity of long-term water quality and catchment datasets, which are required to quantify lags;
   - Localised, case-study nature of most studies on effectiveness of measures and related response;
   - Inconsistent reporting of catchment factors in studies of water quality response to measures;
   - Lack of knowledge about the function of measures across a range of conditions and environments;
   - Inconsistencies in available evidence per type of pollutant;
   - Difficulty in quantifying the complex processes determining catchment typologies.

Definitions: (1) Any statistically significant improvement in water quality in the waterbody downstream of the catchment where the measures are implemented at or above the level projected to deliver a water quality improvement at a catchment scale. (2) Characteristics such as waterbody type, catchment size, land use, precipitation, pollutant retention and travel / residence time, legacy effects and implementation of measures.

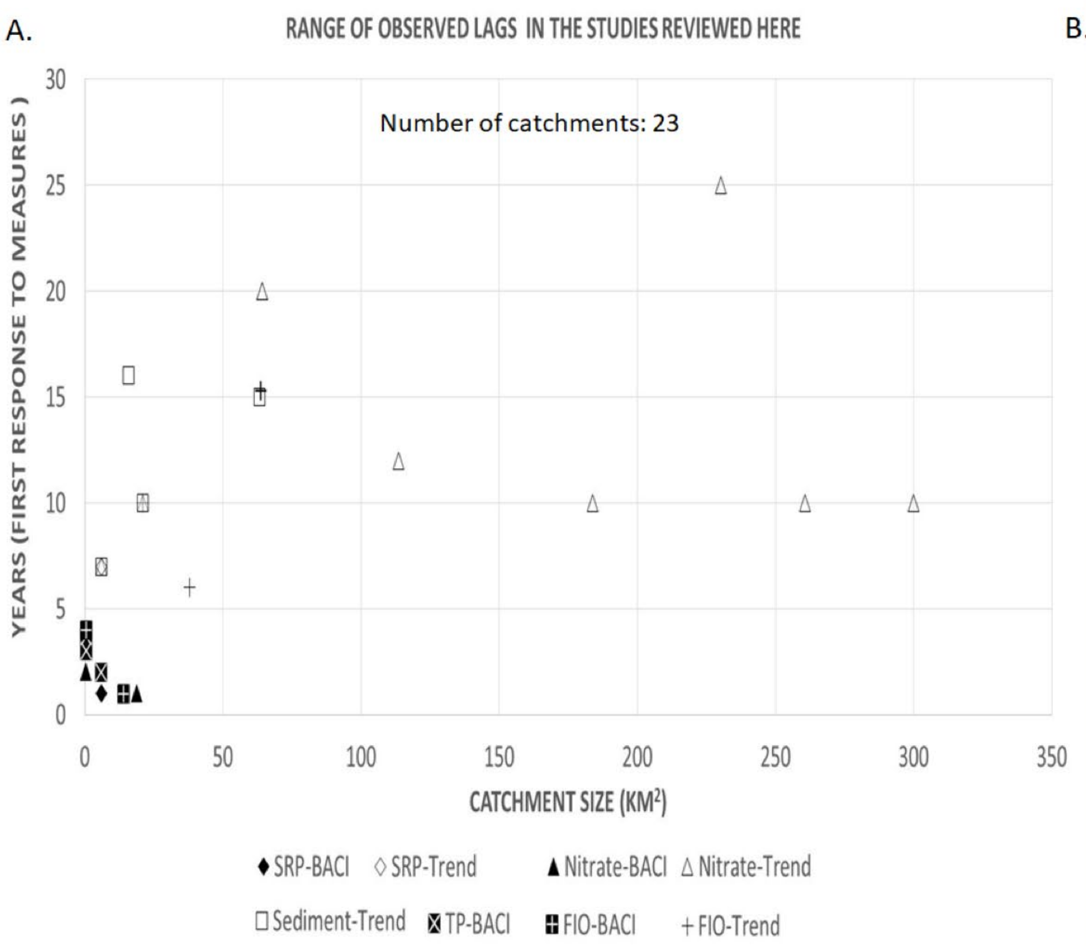
Literature-based recommendations to improve understanding of lags in water quality response

1. Account for dominant legacy effects and hydrologic paths when targeting the measures to address pressures.
2. Promote spatially integrated implementation of a combination of different types of measures.
3. Include *Source control*, as the intended effect (i.e. reducing inputs at source) is independent of legacy effects and hydrologic paths.
4. Collect long-term monitoring data from catchments where the measures are implemented and from control catchments (pristine, or without measures); control data are key to separating effect of measures from the effects of other factors on water quality.
5. Apply a BACI or Trend monitoring design using long-term data depending on availability of pre-implementation data or on mode of uptake of measures (gradual or not).
6. Account for catchment-scale influences on water quality of factors such as rainfall, land use, application of fertiliser, livestock numbers, streamflow, discharges from point-sources, and data on soil sorption capacity and rates of biogeochemical processes.
7. Model water quality responses to catchment processes to derive catchment-specific typologies and understand sensitivity to measures over time and guide further action.

Practical implications - recommendations for Scotland

1. Keep monitoring water quality to help understand lags and inform further action.
2. Adjust expectations for water quality response and recovery, i.e. there is no evidence supporting fixed timeframes for waterbody improvement.
3. Plan for lags in water quality response. This may involve:
   - Planning for longer-term monitoring and flexible objectives as in "learning by doing";
   - Prioritising measures that deliver immediate results by accounting for hydrologic paths;
   - Targeting sources nearest to receiving waters for faster improvements;
   - Demonstrating results to the public in areas delivering immediate water quality responses.
4. Account for dominant hydrologic paths at farm scale during catchment characterisation surveys and when targeting; this means collecting evidence on soil properties, soil sorption capacity, legacy nutrients, geology, streamflow and precipitation along with evidence on land use and pressures.
5. Match the measures to the pollutant(s), pollutant source(s), and hydrologic transport pathways.
6. Promote spatially integrated implementation of a combination of types of measures.
7. Avoid inputs at source (*Source control*).
8. Consider soil pore-water nutrient measurements to demonstrate effectiveness of *Source control*.
9. Consider retro-fitting the correct measure(s) to site-specific losses when assessment of the measures in place shows that the predominant sources of pollutants have not been addressed.
10. Develop modelling approaches (e.g. a decision support tool) examining the effect of a suite of catchment factors on water quality using readily available desk-based GIS data.
11. Investigate strategies for the effective communication of scientific evidence on lags and adaptive management approaches in the context of cost-effectiveness of the measures.

A.



B.

Nutrients

<u>*Source control* measures</u>
- Reduction in N and P fertilisers
- Optimisation of N
- Manure storage
- Livestock restrictions (fencing)
- Reduction of stocking rates
- Farmyard runoff control
- Removal from crop production
- Irrigation management

<u>In-field *Transport control* measures</u>
- Catch crops

<u>Riparian *Trapping* measures</u>
- Constructed wetland
- >8 m Riparian buffers
- Riparian restoration with woody vegetation

FIO

<u>*Source control* measures</u>
- Livestock exclusion with off-site watering, bridges, crossings
- Improved farm infrastructure

<u>Riparian *Trapping* measures</u>
- Streambank stabilisation (incl. revetments)
- Riparian restoration (2-8m up to 30m) with woody vegetation

Sediment

<u>*Source control* measures</u>
- Livestock exclusion

<u>In-field *Transport control* measures</u>
- In-field erosion control

<u>Riparian *Trapping* measures</u>
- Streambank stabilisation
- Riparian restoration

Figure ES1. A. Relationship between time (years) elapsed between the start of implementation of measures and the first detectable water quality response to measures in river waterbodies in temperate regions based on BACI and Trend monitoring designs. No compliance with water quality standards was observed in any of these studies. B. Combinations of measures implemented per catchment per pollutant in the studies presented in A.

1.0 Introduction

1.1 Aim

The aim of this project was to critically review the evidence on lags between the implementation of diffuse agricultural pollution control measures (hereafter the measures) and a catchment-scale water quality response to measures in rivers and groundwater. This report presents evidence on: (i) the processes causing lags along the source-mobilisation-delivery continuum from sources to receiving waters for the key pollutants targeted by the measures in Scotland, i.e. *total phosphorus (TP) and soluble reactive inorganic phosphorus (SRP)*[1], *total nitrogen (TN)*[2], *nitrate (nitrate*$^-$*), faecal indicator organisms (FIO), and sediment;* (ii) current understanding of the factors influencing lags water quality response[3] to measures; (iii) observed lags in the response of TP, SRP, nitrate$^-$, FIO and sediment to measures in rivers and groundwater in temperate regions. This evidence was summarised in a table to explore typecasting of likely water quality response and recovery timescales to measures by waterbody/catchment type, measure, and pollutant in Scotland at a sense-checking workshop.

The knowledge gained from this project will enable Scottish Environment Protection Agency (SEPA) to better understand the current state of knowledge on trajectories of waterbody status improvement in response to the measures. It will also help SEPA set realistic timescales for the achievement of water quality objectives, adjust action and stakeholder *expectations* from the implementation of *measures in Scotland, and* prioritise action.

APPENDIX I.1 provides literature-based definitions on key concepts to *assist* you in understanding research and technical terms commonly used throughout the report.

1 The sum of all phosphorus components in natural waters, total phosphorus (TP), is made up of phosphorus in particulate and soluble forms. Particulate phosphorus (PP) is a combination of organic and inorganic filtrate (>0.45 μm); soluble phosphorus is made up of soluble unreactive phosphorus and soluble reactive inorganic phosphorus (SRP). SRP responds to colorimetric tests (molybdate reactive), is usually considered as bioavailable and is also known as dissolved reactive inorganic phosphorus or orthophosphate, hereafter reported as SRP.

2 Total nitrogen (TN) is the sum of nitrate-nitrite and total Kjeldahl nitrogen (TKN), i.e. nitrogen in organic substances (living and dead organic matter), ammonia and ammonium.

3 At the start-up workshop, the steering group agreed to examine "secondary response times for ecology (*diatoms, invertebrates, and estuarine macroalgae*)", which was initially requested by SEPA, once the potential for delivering catchment typologies for pollutants has been discussed. The reason for that decision was that ecological response times depend on complex ecological processes independent of or confounded by catchment processes.

1.2 Policy context

The River Basin Management Plans (RBMP) developed under the European Union (EU) Water Framework Directive (WFD; 2000/60/EC, OJEC, 2000) set out the requirements for the necessary Programmes of Measures (POMs; see APPENDIX I.1) to achieve time-scaled environmental objectives for surface water and groundwater waterbodies and protected areas.

Article 11 of the WFD prescribes the implementation of "basic" (regulatory) measures, and where necessary, "supplementary" (incentivised) measures to achieve the objectives set in RBMP. 'Basic' measures are described as minimum requirements including relevant existing EU legislation (e.g. the Nitrate Directive), designed to control practices resulting in point (e.g. farmyards) and diffuse (e.g. cropland) pollution sources.

In Scotland, basic measures are implemented as a mandatory set of requirements known as Diffuse Pollution General Binding Rules (DP GBR) and are outlined in the Controlled Activities Regulations (CAR 2019). On compliance with DP GBR, SEPA deliver guidance to farmers and land managers on the uptake of supplementary measures funded by the Agri-environment Climate Scheme of the Scotland Rural Development Programme (SRDP)[4] to help improve and protect water quality beyond compliance with regulations. SEPA are currently reviewing rural diffuse pollution pressures and WFD objectives for the third six-yearly RBMP cycle.

SEPA need to communicate to stakeholders the causes of the perceived lack of water quality response to measures. As specified in the project request: "Understanding the catchment processes determining the timescales required for water quality response to measures is key to developing realistic water quality objectives in the third RBMP cycle. This will help to gauge where further actions may be targeted in agricultural catchments and minimise the risk of assuming that the measures have been ineffective and inefficient".[5]

1.3 Project objectives

SEPA asked CREW to systematise the evidence-base on the catchment typology determining lags between the implementation of measures and water quality response for each pollutant.

The specific objectives of the project were to:

Objective 1: Review and identify key catchment processes influencing lags in water quality response to common implemented measures in Scotland.

4 For an overview of schemes see: Scottish Government (no date).

5 For the terms effectiveness and efficiency of measures see APPENDIX I.1.

Objective 2: Review the (inter)national evidence base on observed lags in water quality response to measures for each type of measure and key pollutant, *i.e.* TP, SRP), TN, nitrate, FIO, and sediment.

Objective 3: Assess the potential for identifying catchment typologies for each pollutant's broad response timescales to each implemented measure (or combination of measures).

Objective 4: Discuss a table of lags in water quality response per pollutant per measure in a sense-checking workshop.

A clarification on terminology is also provided.

1. The term "water quality response times", which was initially mentioned in the project's spec, is hereafter reported as "lags in water quality response to measures" or briefly as "lags" (see APPENDIX I.1). Lags may refer to: (i) the time required to detect any measurable (significant) water quality improvement; or (ii) the time required to achieve the objectives set for a waterbody, which is more general as a concept and may include a range of different end-points from detecting any measurable improvement (as in (i)) to achieving the best possible status for a particular waterbody. Following on from this, the meaning of "waterbody recovery" is conditional on specific policy objectives determined by site-specific circumstances.
2. The term "catchment typology" was initially mentioned in the project request but was not found in the literature on lags in water quality response to measures. It refers to a suite of characteristics such as waterbody type, catchment size, land use, precipitation, pollutant retention and travel / residence time, legacy effects from historic inputs and implementation of measures in terms of types of measures (Figure 1).
3. Waterbody type: refers to streams, rivers, and groundwater waterbodies but not to loch, transitional or coastal waterbodies.
4. The term "common implemented measures in Scotland" refers to GBR and SRDP measures currently implemented in Scotland's agricultural catchments (Figure 1), hereafter reported as SEPA's measures. SEPA's measures were divided into three categories in terms of their intended effect (see APPENDIX I.1.) to enable comparisons with the terminology and combinations of measures implemented elsewhere (Biddulph et al., 2017; Lintern et al., 2018; Rittenburg et al., 2015; Schoumans et al., 2015). In brief:
   - Source control measures reduce or avoid inputs of pollutants.
   - In-field Transport control measures delay transport of pollutants through the soil to enable their uptake by crops and removal (e.g. denitrification and FIO die-off), and soil stabilisation.
   - Riparian Trapping measures enhance trapping of pollutants through plant uptake, retention, and streambank stabilisation to enable removal before delivery to waterbodies.
5. The term "pollutants" refers, hereafter, collectively to TP, SRP, TN, nitrate, FIO, and sediment.

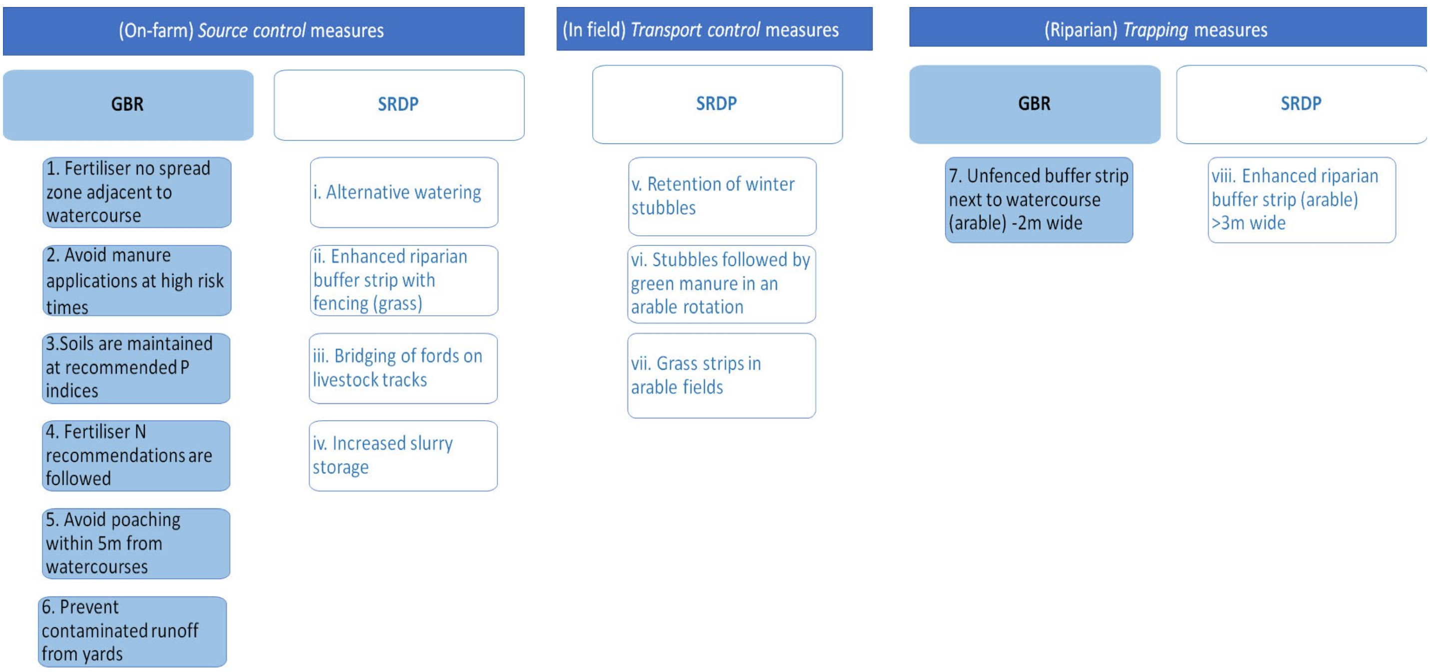


Figure 1. List of SEPA's measures in relation to their intended effect. Measures 1-4: GBR 18 (Storage and application of fertiliser); Measure 5: GBR 19 (Keeping of livestock); Measure 6: GBR 10 (Preventing pollution from yard runoff); Measure 7: GBR 20 (Cultivation of land); Measures i-viii: SRDP measures. See also APPENDIX I.1 for measures implemented elsewhere under each category of intended effect of diffuse pollution control measures.

1.4 Outline of the report

The report includes the following sections:

- Section 2 gives an overview of the literature review approach (detailed in APPENDIX I.2).
- Section 3.1 (Obj. 1) outlines causes of lags (processes leading to lags are detailed in APPENDIX II).
- Section 3.2 (Obj. 1) summarises catchment factors influencing lags (relationships between catchment factors are described in APPENDIX III).
- Section 4 (Obj. 2) explores patterns in observed lags per pollutant in relation to catchment factors (metadata reported in APPENDIX IV) and discusses policy challenges in understanding lags.
- Section 5 (Obj. 3) lists the reasons why lags based on catchment typologies could not be identified.
- Section 6 presents literature-based recommendations to better understand lags and lists practical implications of available evidence on lags for Scotland.

2.0 Approach

Addressing Objective 1 involved undertaking a systematic review of the literature (i.e. published online by April 2020, unless otherwise stated). Review for Objective 2 involved critical appraisal of the evidence on lags from 16 peer-reviewed articles and a report on water quality response to measures (APPENDIX II). Objective 2 focused on evidence from small to medium sub-catchments (1-300 km^2) in temperate regions to allow parallels to be drawn with the Scottish context. The evidence collected informed Objective 3, a list of recommendations and a sense-checking workshop, which was held in September 2020. APPENDIX I.2 details the literature review approach.

3.0. Obj. 1: Causes of lags and catchment factors influencing processes leading to lags

3.1 Causes of lags in water quality response to measures

Water quality response is understood as a pollutant transfer continuum, whereby nutrient sources as inputs at the farm scale and field soils can be exposed to a mobilisation mechanism through biogeochemical processes and, via hydrologic pathways, delivered to streams or other water bodies where an impact may be observed (Granger et al., 2010; Rittenburg et al., 2015). Figure 2 gives schematic representation of this continuum in the context of all types of pressures in a catchment.

Lags in water quality response to measures are a "fact of life" in catchment management because the processes causing these lags are ubiquitous (Chen et al; 2018 Meals et al., 2010). The **processes causing lags** in water quality response to measures are (Chen et al., 2018):

1. **Legacy effects from past (historic) inputs.** Pollutants from past inputs can persist long after inputs have ceased as a result of the implementation of the diffuse and point source control measures (past legacy effects). Past legacy effects on lags are poorly quantified internationally (Chen et al., 2018). Gregory et al. (2007) observed that given the storage capacity of soils and sediments for nutrient and other pollutants, it is unrealistic to expect *Source control* will have an immediate impact on water quality. Growing evidence from long term studies (i.e. in the range of 50 years) shows that past legacy effects delay response to Source control measures from years to decades (Dupas et al., 2019; Van Meter and Basu, 2017).

2. **Biogeochemical legacy effects.** Biogeochemical transformations and transport through soil, bacteria, crops, crop residues, livestock waste, vegetated buffers, wetlands, streambed and stream water increase residence time before mobilisation and removal or delivery to receiving waters. Biogeochemical legacy effects lead to accumulation of pollutants ("sinks" or storage) which are known to delay phosphorus and nitrogen response to measures, especially when this happens at sites in-field, in streambank or streambed prone to occasional disturbance due to fast runoff or streamflow, livestock, or erosion (Agouridis 2005; Chen et al., 2018). Degradation of vegetated buffers, which serve as nutrient storage sites, will also delay response to measures (Stutter et al., 2019). Poor matching of pressures with mobilisation processes at farm-scale has been noted as a problem leading to low efficiencies of the measures and lack of response to measures[6] (Biddulph et al., 2017; Chen et al., 2019; Meals et al., 2010; Rittenburg et al., 2015; Wilcock et al., 2013).

6 i.e. lack of accounting of the biogeochemical processes enabling mobilisation and removal of pollutants (including legacy pollutants) through crop uptake of nutrients, denitrification and FIO die-off as well as immobilisation through sorption and soil retention.

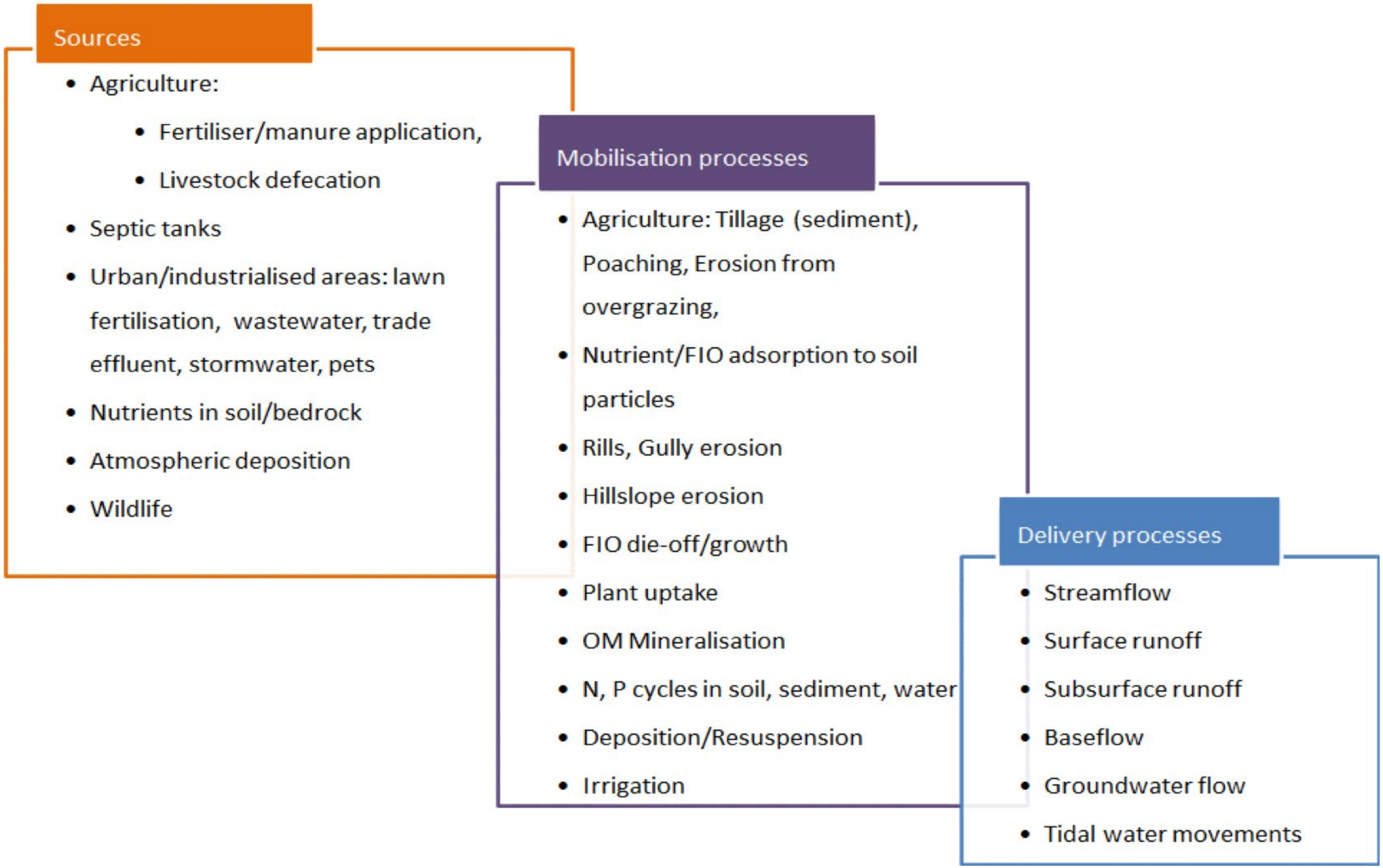


Figure 2. Processes involved in the pollutants' source-mobilisation-delivery continuum.

3. **Hydrologic legacy effects.** Slow water travel time from the farm to the receiving waters through the catchment can delay transport of pollutants, thus masking the effects of pollution control measures. Key surface delivery pathways are surface runoff (a.k.a. overland flow); streamflow; and tidal water movements (Figure 3). Key sub-surface delivery pathways include subsurface runoff (a.k.a. throughflow); preferential flow (i.e. fast, vertical or lateral, crack- or macropore-dominated flow in soils); soil matrix flow (i.e. slow uniform flow in soil without macropores); water infiltration; artificial drainage; lateral groundwater flow; and baseflow (Figure 3). Travel times of these transport pathways are summarised for different pollutants and surface and subsurface pathways in Table 1.

4. **Current inputs from multiple non-agricultural sources.** Non-agricultural sources can mask the effect of the measures at catchment scale. For example, non-agricultural point and diffuse sources in urban, rural and industrial sectors can increase pollutants at catchment-scale, particularly nutrients (e.g. Stets et al., 2020), sediment (Biddulph et al., 2017; Collins and Zhang 2016) and FIO (Kay et al., 2012); see also Figure 2. In addition, anthropogenic and wildlife FIO can counteract the effects of the measures as shown by microbial source tracking in agricultural catchments with livestock exclusion measures in place (Kay et al., 2012). In Scotland, key sources of pollutants potentially confounding response to measures include: (i) private sewerage systems, which serve approximately 10% of the population in rural areas; and (ii) wastewater and stormwater discharges (combined sewage overflows) from the public sewerage network (SEPA 2020).

The hydrologic processes underpinning hydrologic legacy effects are explained in APPENDIX II.1 and II.2. The biogeochemical processes leading to storage of pollutants in the catchment for each form of pollutant are explained in APPENDIX II.3 to II.5.

Based on understanding of these processes, Meals et al. (2010) broke down lag time in water quality response to measures into three interrelated and **temporally overlapping** components:

1. The amount of time taken for the measures to produce their intended effect at the farm scale. This refers to reducing past and biogeochemical effects at the farm scale, i.e. reduced pollutant levels in the soil and removal of pollutants from the soil through biogeochemical transformations.
2. The amount of time taken for the intended effect of the measure(s) to be delivered from each farm to the adjacent waterbody. This refers to reducing hydrologic legacies (travel time of in-field hydrologic paths) via delivery of mobilised pollutants.
3. The amount of time it takes for a water quality response at the catchment scale, i.e. the travel time of pollutants through the river network and groundwater.

It is difficult to estimate the time required for each lag component in water quality response to the measures. Table 1 summarises the evidence detailed in APPENDIX II on travel times of pollutants and fate of each form of pollutant along the pollutant transfer continuum. As shown in Table 1, there is a wide range of travel times of pollutants and different risks and benefits in the context of lag times in response to measures. The next section describes the current state of understanding of the catchment factors influencing the processes leading to lags and discusses knowledge gaps.

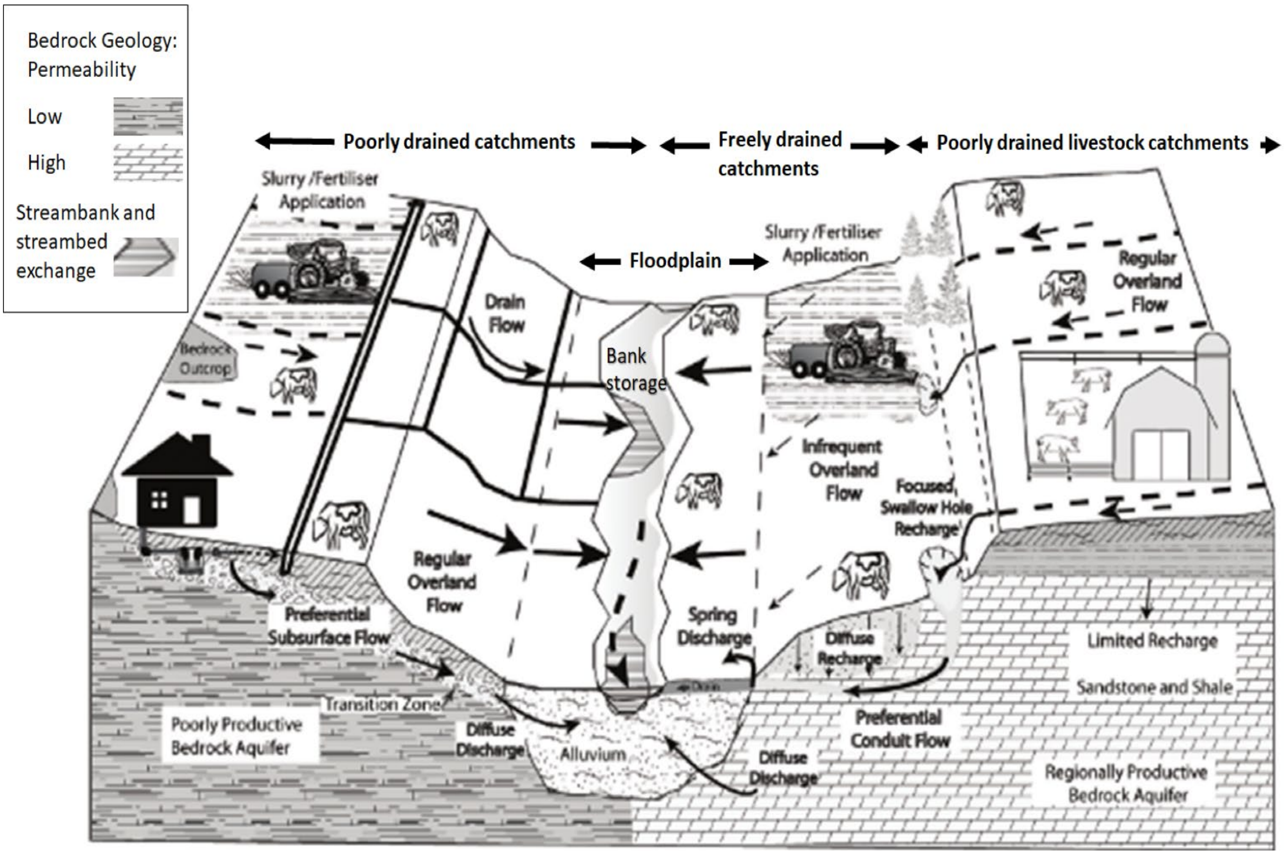


Figure 3. Schematic representation of delivery pathways in flashy (poorly drained) catchment and freely drained catchments showing surface and subsurface/groundwater hydrologic paths. The thicker arrows represent a larger relative flow component than the thinner arrows. Dashed arrows represent intermittent flow. Adapted from Deakin et al., 2016.

Precipitation regime: High **Soil type:** Freely or poorly drained **Hydrologic path:** Overland flow, low infiltration rates, soil matrix flow	**Greatest risk of pollutant loss:** • Particulate-Adsorbed forms (Phosphorus, NH4+, FIO) • Sediment • Dissolved forms: nitrate, orthophosphate	**Measures accounting for hydrologic paths** • *Source Control* • *Transport control*: taller winter stubble for sediment loss control • Enhanced buffers strips: taller vegetation within in-field or riparian (>3m wide) buffer strips to enable deposition of sediment, removal of N and P through plant uptake and infiltration, allowing for immobilisation of pollutants
Precipitation regime: Low to moderate **Soil type:** poorly drained – shallow water table **Hydrologic paths: S**hallow throughflow, lateral preferential flow, overland flow during heavy, srain rain flow	**Greatest risk of pollutant loss:** • Dissolved: Nitrate, Orthophosphate • Particulate-Adsorbed forms (Phosphorus, NH4+, FIO) • Sediment	**Measures accounting for hydrologic paths** • *Source Control* • *Transport control* : winter stubble and green manure to enhance immobilisation and removal (e.g. denitrification) in-field
Precipitation regime: low to moderate **Soil type:** freely drained, low water table **Hydrologic path:** percolation, leaching, preferential pathways, soil matrix flow	**Greatest risk of pollutant loss:** • Leached: Nitrate, orthophosphate, FIO	**Measures accounting for hydrologic paths** • *Source Control* • *Transport control* and *Trapping*: ensure soil root zone of in-field or riparian buffers is deep enough to intercept subsurface flow and allow for pollutant immobilisation

Figure 4. "Land type framework" (see text). Modified from: Rittenburg et al. (2015).

Table 1. Estimates of travel time through different hydrologic pathways and hydrologic compartments. Based on a synthesis of the evidence reviewed in APPENDIX II.

Field to catchment outlet through overland surface runoff and streamflow	**Field to stream through subsurface runoff in the soil matrix)**	**Field to stream through subsurface preferential pathways/ drainage**	**Transport below the water table through groundwater** **(residence in groundwater)**	**From groundwater to streamflow via baseflow**
Relevant to: TN, Nitrate, TP, SRP, free-FIO, Particulate P, Organic N, sediment-bound ammonium and FIO, and sediment	Relevant to: TN, Nitrate, TP, SRP, Organic N, sediment-bound ammonium and FIO, and sediment	Relevant to: TN, Nitrate, TP, SRP, Organic N, sediment-bound ammonium and FIO, fine sediment	Relevant to: TN, Nitrate, TP, SRP, leached FIO	Relevant to: TN, Nitrate, TP, SRP
Days to months Longer travel times occur: -If overland runoff is slow (flatter slopes) -In the presence of in-field and riparian vegetation -Soil infiltration is high (sandy soils) -Streamflow is low -Deposition in streambed (but still less than a year) -In main stem than tributaries (for sediment bound pollutants) -In the floodplain of flood-prone areas (by years or decades) -In the case of longer and frequent dry spells Risks of prolonged travel time for lags: -Built up of legacy nutrients and FIO in-field -Built up of streambed legacy nutrients, FIO and sediment in-stream	Months to years or decades -P (dissolved and particulate): 5-30 years -N (dissolved and particulate): in the range of decades Longer times are favoured by: -Past legacy effects -Small soil pore spaces -Low groundwater recharge / water table -High soil clay or mineral content Benefits of prolonged residence in the soil for lags (depending on pH, mineral content, oxygen and precipitation): -Sufficient time for N removal (denitrification) -Higher chance for P immobilisation -FIO immobilisation Risks of prolonged residence in the soil for lags (depending on pH, mineral content, oxygen and precipitation): -Built up of legacy nutrients and FIO	Days to months -Dissolved pollutants: Same movement rate as that of water infiltration Risks of prolonged travel times for lags: -Bypassing of in-field Transport control and riparian Trapping measures.	Shallow aquifers: Months to years Deep aquifers: years to decades (e.g. for nitrate) Travel time depends on: -Bedrock geology -Depth / size of the aquifer Longer times found in: -Sandstone type of aquifers (e.g. Scotland: Fife, Strathmore, Moray, Central belt, Southern Scotland) Benefits of prolonged travel time for lags: -N removal (denitrification) -P sorption on aquifer matrix Risks of prolonged travel time for lags: -Past legacy effects	-Important in baseflow-dominated catchments and during the low flow season. -Travel time and exchange between streambed and streamflow depends on redox and sediment properties. Risks for lags: -Increasing in-stream concentration long after the implementation of measures.

3.2 Catchment factors influencing the processes leading to lags

APPENDIX III describes the relationships between catchment factors and pollutants along the pollutant transfer continuum.

The catchment factors influencing the processes leading to lags are[7]:

1. **Implementation of measures.** This refers collectively to: intended effect of the measures (APPENDIX I.1); extent; distribution; distance of sites of implementation from receiving waters, which is related to catchment size, artificial drainage and hydrologic connectivity; targeting pressures, hydrologic paths and biogeochemical processes; maintenance; function under varying environmental conditions such as precipitation and temperature; and performance over time. Rittenburg et al. (2015) described **how the measures influence the processes leading to lags** by combining catchment factors and these processes into a "land type framework" (Figure 4). This a simplified conceptual typology of hydrologic paths predicting which measures will deliver improvements for a form of pollutant. For example, precipitation effects and predominant hydrologic paths may counteract or confound the intended effects of the measures, if the measures are not properly targeted or maintained to address hydrologic processes. According to this framework, the effectiveness of *Source control* is the least affected by hydrologic paths. On the other hand, lack of appropriate targeting of Transport control and Trapping measures can lead to fast transfer of pollutants to receiving waters before past legacies can be addressed or biogeochemical transformations (e.g. plant uptake, FIO die-off or denitrification[8]) can take place. In this context, riparian buffers are ineffective at sites dominated by preferential pathways (Stutter et al., 2019).

   In the same line, a more recent study suggested that c a t c h m e n t typologies that classify catchments based on their concentration-discharge (C-Q) combinations could be used as part of a decision support system for the improvement of monitoring design and for spatially targeting catchment scale-measures (Hashemi et al., 2020). For example, they suggested that application of Trapping measures (e.g., constructed wetlands) may be appropriate for catchments showing increase of C with Q, while Source control measures may be more useful for catchments showing decrease of C with Q (Hashemi et al., 2020).

2. **Landscape features such as land use/land cover, geology, soil type, and topography.** Each of these features has a different relationship with water quality (see Box 1A and B). There is no clear reporting of the effect of these factors on lags. However, it is understood that overland runoff and streamflow are slower in flatter slopes, and this may prolong travel time for pollutants thus, increasing the risk of built up pollutants in riparian areas and in streambed (APPENDIX II.1 and APPENDIX III.2: Catchment slopes). It is important to remember that seasonal changes in vegetation cover are important drivers of change in nutrient uptake and riverine nutrient and sediment (Guo et al., 2019). There is also relatively good understanding of the combined influence of geology and hydrology on groundwater nitrate but much less is known on the effects of land use, the measures and climate on groundwater water quality, which are influenced by site-dependent hydro-geological gradients (Fenton et l., 2011; McDowell et al., 2020; Vero et al., 2018, Wang et al., 2013).

3. **Catchment size.** A larger catchment size comes with longer hydrologic paths as a result of longer distance from field to stream and through potentially denser and more complex river networks, which increase the risk of built up of pollutants at various in filed sinks and streambed storage sites (Section 3.1; APPENDIX II.1). It can be assumed that a larger catchment size may translate into prolonged travel times, and therefore into prolonged lags in water quality response to measures. However, this assumption fails to account for legacy effects from past inputs, the different transformation and transport processes for each form of pollutant (APPENDIX II.3-5), type of soils, bedrock geology and slopes, presence of point sources and factors related to the implementation of measures, e.g. extent and distribution of measures (Lintern et al., 2018; Meals et al., 2010; Melland et al., 2018; Rittenburg et al., 2015).

7 Source of evidence, unless otherwise stated for specific information: Agouridis et al., 2005; Chen et al., 2013; Heathwaite 2010; Guo et al., 2018; Lintern et al., 2018; Liu et al., 2017; Melland et al., 2018; Osmond et al. 2019; Rittenburg et al., 2015; Stutter et al., 2019; Van Meter and Basu 2018.

8 APPENDIX II.

4. **Hydrology.** This includes features related to hydrologic paths discussed as part of hydrologic legacy effects in Section 3.1 (see also Figure 3) as well as to catchment response to rain (flashy vs slow) and hydrologic connectivity. It could be assumed that flashy small catchments will respond faster to measures because of faster export of pollutants from the catchment, and therefore smaller past legacy effects and shorter hydrologic legacy effects. However, this assumption fails to account for the effect of factors such as type of farming, biogeochemical legacies, which are different for each form of pollutant, and precipitation regime (Chen et al., 2018; Rittenburg et al., 2015); see also Figure 3 and 4. It must be also noted that changes in hydrology and especially in streamflow (e.g. same-day streamflow) are among the key drivers of temporal variability in riverine nutrients and sediment at catchment scales (Guo et al., 2019). Additional important drivers of temporal variability in water quality are also related to hydrology and include recent (less than a month) streamflow and soil moisture (Guo et al., 2019).

5. **Climate.** This refers to weather characteristics (precipitation and temperature); seasonal or periodic, decadal scale climate cycles, e.g. North Atlantic Oscillation; and long-term changes in climate, e.g. climate-change related influences on hydrology, vegetation and water temperature. Climate change-driven changes in precipitation can have greater effects than could be expected by the measures already in place, outweighing water quality response to measures (e.g. Gregory 2007).

BOX 1A. Examples of catchment factors with a consistently positive or negative relationship with pollutants along the pollutant transfer continuum across studies.

- Area of forests or wetlands and nutrients and sediment in-stream (negative).
- Rainfall and FIO from wastewater discharges due to dilution (negative).
- Rainfall and the mobilisation of sediment-bound nutrients and FIO (positive).
- Area of forests and mobilisation of sediment and sediment-bound pollutants in runoff (negative).
- Farmland runoff and nutrients delivered in-stream (positive).
- Rainfall and pollutants from diffuse sources in surface and subsurface runoff (positive).
- Natural grassland and wetlands and diffuse pollutants in runoff (negative).

Source: APPENDIX III.1

BOX 1B. Examples of factors with an inconsistent relationship with pollutants along the pollutant transfer continuum across studies.

- Erosion and pollutants in-stream. Erosion rate for each soil type varies in space and time as it is determined by complex relationships between a wide range of factors, including slope length and gradient, intensity frequency of storm events and the type of measures implemented.
- Soil sorption capacity and pollutants in-stream. Biogeochemical transformations between particulate (adsorbed) and solute (de-sorbed) forms of pollutants vary and are influenced by complex interactions between rainfall, runoff and deposition.
- Catchment slope and different forms of pollutants. This is a complex relationship. A simple explanation is that while sediments and sediment-bound pollutants are more easily mobilised in steeper slopes (positive relationship), these pollutants may settle out of surface runoff in areas with shallower slopes and contribute to soil legacies and thereafter high concentrations in-stream.

Source: APPENDIX III.2

4.0 Obj. 2: Studies reviewed to explore patterns in lags in water quality response to measures

This section reviews the evidence on lags in water quality response to measures in temperate regions based on 17 catchment-scale studies[9]. These studies helped to explore patterns in lags in water quality response to measures. Metadata on specific aspects of these studies (i.e. monitoring data, uptake of measures and efficiency of measures) are provided in APPENDIX IV. Table 2 summarises the results of studies that reported a lack of water quality response to measures. Table 3 presents observed lags by pollutant in relation to monitoring/ analysis design, catchment size, land use, precipitation/ discharge, soil/bedrock properties, types of measures implemented and the interpretation of results, as reported by the authors of the studies. Figure 5 presents the relationship between catchment size, monitoring design and observed lags. The results are described in the form of a "Questions and Answers" section to help understanding (Sections 4.1-4.7). Section 4.8 provides evidence on policy challenges based on the findings of the overall literature review on lags.

4.1 What are the factors leading to a lack of water quality response to measures?

10 out the 17 studies reviewed reported a lack of response to measures (lack of significant improvements) for at least one of the pollutants studied (Table 2). Box 2 outlines reasons for a lack of response to measures given in the studies reviewed in Table 2. The post-implementation monitoring period in studies reporting a lack of response to measures ranged from 1 to 26 years. It must be also noted that the evidence in Table 2 largely depends on the availability of monitoring data in each catchment and should not be interpreted as an indication of lags in water quality response to measures (observed lags are discussed in Section 4.2 and shown in Table 3). The examples in Table 2 show that it may take many years before any response to measures is detected (e.g. Dupas et al., 2019; Pearce and Yates 2017; Schilling et al., 2011; Wilcock et al., 2013; Zhang et al., 2016).

Lack of response to Source control measures. Reduction in N inputs and optimisation of P application did not lead to groundwater and stream quality response due to legacies from past inputs and biogeochemical legacies in farm soils (Table 2: Hansen et al., 2019; Zhang et al., 2016; see also reviews by Jarvie et al., 2013; Vero et al., 2018). Dupas et al. (2019) attributed the lack of a phosphorus response to measures in Britany's rivers to the insufficient extent of measures and gaps in long-term data series but also highlighted the need to plan for more diffuse pollution control measures and a longer monitoring time to detect a phosphorus response. Lack of FIO response to livestock exclusion measures was reported in England (Davey et al., 2020) and in small catchments in New Zealand (Wilcock et al., 2013). Lack of FIO response was attributed to insufficient extent and distribution of livestock restrictions for access to streams across the catchment.

Lack of response to in-field Transport control and riparian Trapping. No sediment response to measures could be detected in a study on the effectiveness of prairie reconstruction to reduce sediment mainly due to lack of addressing sediment inputs from streambank erosion (Schilling et al., 2011). No sediment or nutrient response was found in studies implementing a combination of livestock exclusion, *Transport (erosion) control* and *Trapping measures* (Davey et al., 2020[10]; Table 2: Pearce and Yates 2017), the reasons provided being: (i) poor targeting of hydrologic paths; (ii) the measures were not yet fully effective; (iii) varying performance of the measures over the range of site-specific levels of precipitation and temperature; and (iii) need for more water quality and catchment data (i.e. longer-term and higher resolution) to capture a significant response to measures. Steinman et al. (2018) observed that lack of response to wetland restoration could be due to the restoration being still very recent: at the start of the post-implementation period in their study, the restored sites were not fully functional. Pearce and Yates (2015) observed that the function of measures can be negated

9 Melland et al. (2018) also reviewed 25 studies from mesoscale catchments (<100km²) and identified relationships between pollutant response and factors such as monitoring, catchment size, hydrologic pathway (as a proxy of travel time) and type of uptake (mandatory or voluntary) of the measures implemented. The review by Melland et al., (2018) reported lags in the context of catchment size and hydrologic pathways but provided no evidence on trajectories in water quality response to measures and did not include FIO. There were studies from a range of climatic conditions (i.e. from Northern Europe and Australia to tropical South America) in the studies reviewed but no evidence on how streamflow or precipitation affected lags. Also, some of the studies referred to spatial comparisons without temporal information on practice change or there were no measures over the monitored timeframe. For these reasons, it was decided to carry out a new review of the literature focusing on temperate regions and lags in relation to catchment factors.

10 The study by Davey et al. (2020) reports results for the median catchment size (62.5 km²); the range of catchments was 11-2276 km². For this reason, this study is not included in Table 2 and 3.

by environmental conditions. For example, under warm, dry conditions in the summer the land is less likely to generate significant surface runoff and measures designed to capture nutrients and sediments in runoff have less influence on surface water quality. Lack of appropriate ongoing maintenance of the measure (e.g. removing sediment from traps) means they have a finite working lifespan (e.g. Stutter et al., 2019). Constructed wetlands collecting tile drainage in the US Midwest were virtually ineffective in reducing P (Kovacic et al., 2000 cited in Osmond et al., 2019; this study is not reviewed in Table 2). This was because the wetlands received mainly SRP that was initially sequestered by wetland plants but then released when vegetation died.

Trajectories: Increases in FIO and nitrate were also observed post-implementation (e.g. Wilcock et al., 2103). Studies that reported a lack of water quality response to measures observed a lack of declining trends or inconsistent fluctuations in water quality in the post implementation period response to measures (e.g. Bergfur et al., 2012; Zhang et al., 2016).

Catchment size. It is unclear how a lack of response may be related to catchment size (Table 2). For relatively small catchments (1-25km^2) the post-implementation period without detecting a response ranged from 1 year for TP and sediment (Meal 2001) to 15 years for nutrients (Pearce and Yates, 2017).

Box 2. Interpretation of the lack of response to measures.

Lack of water quality response to measures was attributed to combinations of the following reasons:

1. Uncertainties about the effectiveness of measures and the level of implementation required for a water quality response.
2. Lack of integrated implementation of combinations of measures at catchment scale/ poor targeting.
3. Low efficiencies of the measures in the context of background catchment variability and pressures such as climate change, which translate to small, undetectable, improvements.
4. A longer time required for the measures to become fully effective.
5. Variable function and performance of the measures in response to environmental conditions.
6. Lack of appropriate long-term water quality and catchment data to account for catchment factors.
7. Poor understanding of the start time of the post-implementation period at the catchment scale, which affects statistical analyses and study designs.
8. Monitoring design, which may be introducing a statistical lag, or is unable to detect the magnitude of improvement that has occurred or can occur under site-dependent circumstances.

Source: Meals 2001; Simon and Makarewitz 2009; Line et al., 2016; Bergfur et al., 2012; Pearce and Yates 2017; Wilcock et al., 2013; Schilling et al., 2011; Steinman et al., 2018; Zhang et al., 2016; Dupas et al., 2019; see also Table 2.

Table 2. Examples of studies that reported a lack of water quality response to measures. *References: 1. Meals 2001; 2. Steinman et al., 2018; 3. Simon and Makarewitz 2009; 4. Line et al., 2016; 5. Bergfur et al., 2012; 6. Pearce and Yates 2017; 7. Wilcock et al., 2013; 8. Schilling et al., 2011; 9. Zhang eta l., 2016; 10. Dupas et al., 2019.*

Pollutant	Years post-implementation without detecting a response	Total monitoring time (years)	Monitoring design	Catchment size (km^2)	Measures (examples)	Way of uptake of measures	Ref.
TP, sediment	1	5	BACI	14	Livestock restrictions, Streambank stabilisation, Riparian restoration	Voluntary uptake, fully implemented by end of baseline period, joint maintenance by Ministry-landowners	1
TP, SRP, TP, nitrate, ammonia	2	3.5	BACI	464	Wetland restoration (overall restored area: 0.45km^2)	Planned / implemented by end of baseline period	2
Total coliforms	4.5	5	BACI	1-10	Improved farm infrastructure, Manure management	Uptake within a year from start of sampling	3
Nitrate	3.5		BACI	1	Livestock exclusion from riparian areas with alternative watering	Voluntary uptake by the end of baseline period	4
SRP, nitrate, ammonium	5	10	BACI	1.2-6	Livestock fencing, Riparian broadleaf trees planted	Voluntary uptake within few months by the end of baseline period	5
TP, SRP, TN, nitrate, ammonium	3-15	3-15	• No baseline data • One-off sampling for 2 weeks • Multiple regression between (i) "Indicators of Abundance and Location of measures within the catchment" and stream metabolism and nutrients, and (ii) stream metabolism and nutrients	5-12	Manure storage, Livestock restrictions, Erosion control structures, 30-m wide riparian buffers	Voluntary uptake in 13 headwater catchments	6
E.coli, nitrate	10 (Increase)	10	Trend	24.8 (Bog Burn)	Source control (slurry storage, away wintering, livestock exclusion), Trapping (riparian buffers strips, low slurry irrigation)	Voluntary, Gradual (Phased) for all measures except irrigation related measures	7
Nitrate	15 (Increase	15	Trend	63.2 (Waikahi)	Source control (livestock exclusion), Transport control (bunds), Trapping (riparian buffers strips, spray irrigation)	Voluntary Gradual-Phased	7
Sediment	10	10	• Control vs Impact (daily and weekly sampling) • Analysis: Multiple linear regression (using seasonality, discharge, and sediment from Control site as covariates)	47-52	Voluntary uptake	Voluntary uptake at all planned sites with a few months	8
TP, SRP, TN, nitrate	26	26	• Trend • Pollutant concentrations from 25- 40 days in combination with daily streamflow data monthly and 8 stormflow samples	1217-20194	Manure or fertiliser management for TP, Control of atmospheric deposition for TN	Voluntary uptake	9
TP	20	20-30 in total but with data gaps	Trend	183.7-300	Manure or fertiliser management, Gradual implementation of EU POMs	Uptake of measures under EU Nitrate Directive	10

Table 3A. Observed lags (in this table: first time a significant improvement was detected) per pollutant, and in relation to catchment factors.

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km^2)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
FIO	Voluntary, fully implemented by end of baseline period, joint maintenance by Ministry-landowners	1	5	BACI	14	150-400m	Annual P: 1049mm	Unstratified glacial drift	Source control, Erosion control, Trapping	Extensive and integrated implementation	No	1
FIO (E.coli)	Voluntary within a year from start of sampling	4	5	BACI and Trend, Pristine area as Control	0.4	Steep	Average Daily Q: 1571 m^3, 4-year AVRG of non-event Q=77% of Total Annual Q	Well-drained limestone	Source control (roof water seaparation, livestock fencing, off-site water pumps, no winter manure spreading), Tile drainage	Extensive and integrated implementation	Maxima gradually declined over the years: annual average values fluctuated seasonally but independently from livestock presence	2
FIO (E.coli)	Gradual-Phased	6	6	Trend, Pristine area as Control, no baseline data	38		Annual P:598-1421mm (study period, Q: 0.01 (Dry season)-70 (Wet season) m^3, Intermittent flow	Unclear	Source control (stream fencing)	Extensive and gradually increasing, targeted implementation	Interannual variability related to 24-h antecedent precipitation	3*
FIO	Gradual-Phased	15	15	Trend, no baseline data	63.2	Flat	Annual P: 520mm	Draining silt-loams	Source control (livestock exclusion), Transport control (bunds), Trapping (riparian buffers strips, spray irrigation)	Effectiveness of Irrigation management and livestock exclusion, need for faster adoption	Increase within 5 years then reduction	4

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
FIO	Gradual-Phased	10	10	Trend, no baseline data	20.9	Flat	Annual P: 1250mm	Volcanic silt loams	Source control (livestock exclusion), Trapping (riparian buffers strips, slurry irrigation,)	Need for longer-term data, Need for faster adoption of measuress	Marginal reduction	4

References: 1. Meals et al., 2001; 2. Simon and Makarewitz, 2009; 3. Lewis et al., 2019; 4. Wilcock et al., 2013; 5. Makarewitz et al., 2009; 6. Line et al., 2016.

3*: This study refers to Mediterranean climate

P: Precipitation; Q: Discharge

Table 3B. Observed lags (in this table: first time a significant improvement was detected) per pollutant, and in relation to catchment factors.

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
TP	Voluntary, fully implemented by end of baseline period, joint maintenance by Ministry-landowners	1	5	BACI	14	From 150m to 400m	Annual P: 1049mm	Unstratified glacial drift	Source control, Erosion control, Trapping	Long-term monitoring and effectiveness of irrigation management and fencing	No	1
TP	Voluntary within a year from start of sampling	3	4	BACI and Trend, Pristine area as Control	0.4	Steep	Average Daily Q: 1571 m³, (study period)	Well-drained limestone	Source control (manure lagoons, soil testing, fertilization rates, crop/livestcok removal); Transpot control (sediment control basins, gully plugs riparian buffer strips, terraces, catch crops and crop roations); Trapping (riparian buffer strips)	Extensive and integrated implementation	Maxima gradually declined over the years but average values fluctuated seasonally but independently from livestock presence	5
TP	Voluntary within a year from start of sampling	2	3	BACI and Trend, Pristine area as Control	6	Steep	Average Daily Q: 7739m³ (study period)	Well-drained limestone	Source control (removal of 37% of crops, removal of cows)	Effective source reduction of TP sources	Anuual average fluctuated but significant decline after the second year	5
TP	Voluntary by the end of the pre-BMP period	3.7	3.7	BACI, pasture non-treated with measures as Control	0.6	6% (pasture), 15% (woodland)	Anuual P: 907(pre)-988(post)	Loamy	Source control (fencing, alternative watering, manure/N fertilisation application, soil P testing), Trapping (riparian vegetation -3m from the top of streambank)	Good statistical design and effectiveness of measures with high efficiency	No	6

SRP	Voluntary within a year from start of sampling	1	2	BACI and Trend, Pristine area as Control	6	Steep	Average Daily Q: 7739m³ (study period)	Well-drained limestone	Source control (removal of 37% of crops, removal of cows)	Effective source reduction of SRP sources	Anuual average fluctuated but significant decline after the second year of monitoring	5
SRP	Gradual-Phased	7	7	Trend, no baseline data	6	Flat	Annual P: 4800mm	Free draining, stony	Source control (riparian fencing, bridges or culverts across streams), Trapping (improved farmyard effluent irrigation)	Improved effluent irrigation	Climate-related remperature-dependent fluctuations	4

References: 1. Meals et al., 2001; 2. Simon and Makarewitz, 2009; 3. Lewis et al., 2019; 4. Wilcock et al., 2013; 5. Makarewitz et al., 2009; 6. Line et al., 2016.

TP: Total Phospshorus

SRP: Soluble Reactive Phosphorus

P: Precipitation,; Q: Discharge

BMP: Best Management Practices (See Appendix I.1)

Table 3C. Observed lags (in this table: first time a significant improvement was detected) per pollutant, and in relation to catchment factors.

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km^2)	Land use % grazing grassland-dairy farming	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Ref.
Sediment	Gradual-Phased	15	15	Trend, no baseline data	63.2	90%	Flat	Annual P: 520mm	silt-loams	Source control (livestock exclusion), Transport control (bunds), Trapping (riparian buffers strips, spray irrigation)	Effectiveness of Irrigation management and livestock exclusion, need for faster adoption	4
Sediment	Gradual-Phased	10	10	Trend, no baseline data	20.9	99%	Flat	Annual P: 1250mm	volcanic silt loams	Source control (livestock exclusion), Trapping (riparian buffers strips, slurry irrigation)	Effectiveness of irrigation management	4
Sediment	Gradual-Phased	7	7	Trend, no baseline data	6	100%	Flat	Annual P: 4800mm	stony	Source control (riparian fencing, bridges or culverts across streams), Trapping (improved farmyard effluent irrigation)	Improved effluent irrigation	4
Sediment	Gradual-Phased	16	16	Trend, no baseline data	15.8	83%	Unclear	Annual P: 1160mm	volcanic silt loams	Source control (livestock exclusion, and grazong management, including stand-off pads), Trapping (riparian buffers strips, slurry irrigation)	Effectiveness of Irrigation management and livestock exclusion	4

References: 1. Meals et al., 2001; 2. Simon and Makarewitz, 2009; 3. Lewis et al., 2019; 4. Wilcock et al., 2013; 5. Makarewitz et al., 2009; 6. Line et al., 2016.

P: Precipitation

Q: Discharge

Table 3D. Observed lags (in this table: first time a significant improvement was detected) per pollutant, and in relation to catchment factors.

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km^2)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
Nitrate	Voluntary within a year from start of sampling	2	3	BACI and Trend, Pristine area as Control	0.4	Steep	Average Daily Q: 1571 m^3, 4-year AVRG of non-event Q=77% of Total Annual Q	Well-drained limestone	Source control (manure lagoons, soil testing, fertilization rates, crop/livestcok removal); Transpot control (sediment control basins, gully plugs riparian buffer strips, terraces, catch crops and crop roations); Trapping (riparian buffer strips)	Extensive and integrated implementation	Maxima gradually declined over the years but average values fluctuated seasonally but independently from livestock presence	5
Nitrate	Voluntary within a year from start of sampling	1	2	BACI and Trend, Pristine area as Control	18.8	steep	Average Daily Q: 2189 m^3	Well-drained limestone	Source control (livestock fencing), Transport control (crop rotation to promote N-uptake), Rotational grazing and "gully plugs" accounted for less than 9.5% of the entire catchment, tile drainage	Extensive and integrated implementation of crop rotation	After decline within a year average values fluctuated -no significant trend	5
Nitrate	Gradual-Phased	10	10	Trend, no baseline data	20.9	Flat	Annual P: 1250mm	Volcanic silt loams	Source control (livestock exclusion), Trapping (riparian buffers strips, slurry irrigation)	Need for longer-term data, Need for faster adoption of measuress	Marginal reduction	4

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
Nitrate	Gradual-Phased	7	7	Trend, no baseline data	6	Flat	Annual P: 4800mm	Free draining stonie	Source control (riparian fencing, bridges or culverts across streams), Trapping (improved farmyard effluent irrigation)	Improved effluent irrigation	Change studies as 4-year mean	4
Nitrate	Mainly referring to uptake of the Nitrate Directive in 1991	10	30	Trend	260.7	7%	Annual Runoff:683mm	Bedrock (Granite/ Schist): 34.1 / 61.1	Reduce N surplas and in-stream N concentrations	Not mentioned	Influences of seasonal variability and wet-dry cycles from North Atlantic Oscillation	7
Nitrate	Mainly referring to uptake of the Nitrate Directive in 1991	10	30	Trend	183.7	5.30%	445.6mm	Bedrock (Granite/ Schist): 67.5 / 30.2	Reduce N surplas and in-stream N concentrations	Not mentioned	Influences of seasonal variability and wet-dry cycles from North Atlantic Oscillation	7
Nitrate	Mainly referring to uptake of the Nitrate Directive in 1991	10	30	Trend	299.8	8.50%	531.5mm	Bedrock (Granite/ Schist): 10.4 / 84.4	Reduce N surplas and in-stream N concentrations	Not mentioned	Influences of seasonal variability and wet-dry cycles from North Atlantic Oscillation	7
Nitrate	Not mentioned	20 7 (Fall) - 21 (Spring)	1973-2014	Trend	64.14	Unclear	A high ratio of surface runoff to baseflow	Diamicton (clay, mud, sand, boulders) / low permeability	N-fertiliser and manure management (not explicitly mentioned if the measures include riparian buffers and constructed wetlands)	It was a study of quantifying lags not studying effectiveness of measures	Strong influences of trajectories in N inputs from all sources, seasonality of hydrological paths	8

Pollutant	Way of uptake of the measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
Nitrate	Not clearly mentioned	12 4 (Fall) - 19 (Spring)	1973-2014	Trend	113.57	Unclear	A high ratio of surface runoff to baseflow	Diamicton (clay, mud, sand, boulders) / low permeability	N-fertiliser and manure management (not explicitly mentioned if the measures include riparian buffers and constructed wetlands)	It was a study of quantifying lags not studying effectiveness of measures	Influence from Tile drainage and point sources	8
Nitrate	Not clearly mentioned	25 16 (Spring) - 32 (Summer)	1979-2014	Trend	230.23	1.5-4.5	Low surface water to baseflow ratio	Gravel, permeable with high levels of groundwater recharge	N-fertiliser and manure management (not explicitly mentioned if the measures include riparian buffers and constructed wetlands)	It was a study of quantifying lags not studying effectiveness of measures	Influence of groundwater flow paths	8

References: 4. Wilcock et al., 2013; 5. Makarewitz et al., 2009; 7. Dupas et al., 2019; 8. Van Meter and Basu 2017; 9. Hansen et al., 2019.

N: Nitrogen

P: Precipitation

Q: Discharge

Table 3E. Observed lags (in this table: first time a significant improvement was detected) per pollutant, and in relation to catchment factors.

Waterbody type	Pollutant	Way of uptake of measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
Groundwater	Nitrate	Gradual	8	1989-2016	Trend	8	Flat	Annual P: 600-1000mm	Sandy (Tile drainage:5-10%) Bedrock:limestone and chalk	Source control (Statutory norms for manure N utilization, Max N allowance for crops equalling economic optimum, Max N allowance for crops ≈ 10% below economic optimum, N leaching reduction, reducing livestock density, handling of manure, crop plans, groundwater protection zones, organic farming, protection of wetlands), Transport control (catch crops, crop rotation), Trapping (10m buffer zones)	Statutory norms for manure N utilization, Maximum N allowance for crops equalling economic optimum, Maximum N allowance for crops ≈ 10% below economic optimum, higher N use efficiency measures in sandy catchments, sandy aquifers had higher N concentrations	Fluctuating and variable between sites, upward trend afyter 2009 due to less efficient measures	9

Waterbody type	Pollutant	Way of uptake of measures	Lag (first response -years post-implementation)	Total monitoring time (years) up to detection of a response	Monitoring design	Catchment size (km²)	Slopes	Climate (P or Q)	Soils	Intended effect of measures implemented	Reason for detecting a response	Interannual variability / Trajectory described	Ref.
Groundwater	Nitrate	Gradual	21	1989-2016	Trend	8	Flat	Annual P: 600-1000mm	Loamy (Tile drainage: 8-70%), Bedrock: Limestone and Chalk	Source control (N leaching reduction, reducing livestock density, handling of manure, crop plans, groundwater protection zones, organic farming, protection of wetlands), Transport control (catch crops, crop rotation), Trapping (10m buffer zones)	Lower N use efficiency in loamy catchments, aquifers had lower N concentrations	Fluctuating and variable between sites, upward trend afyter 2009 due to less efficient measures	9

References: 4. Wilcock et al., 2013; 5. Makarewitz et al., 2009; 7. Dupas et al., 2019; 8. Van Meter and Basu 2017; 9. Hansen et al., 2019.

4.2 What are the observed lags for each pollutant?

The ranges of lags vary between pollutants (Table 4).

Table 4. Ranges of observed lags per pollutant based on the examples shown in Table 3.

Pollutant	Range of observed lag (years)
In stream FIO	1- 15
In stream TP	1-3.7
In stream SRP	1-7
In stream nitrate	1-25
In stream Sediment	7-16
Groundwater nitrate	8-21

It is important to remember that these are examples of lags and <u>should not be interpreted as fixed lags to inform policy</u>, because they are catchment-and pollutant-specific. In addition, as shown in Table 2, many studies have reported a lack of response to measures for a pollutant with longer-term monitoring than that used for detecting a response with a lag.

- FIO appears to be the most sensitive pollutant to *Source control* (i.e. livestock exclusion) as it shows a response in relatively short timescales (Table 3A). Fast response occurs on the condition that there is extensive implementation of measures across a catchment (Lewis et al., 2019) or accounting of dominant hydrologic paths (Wilcock et al., 2013).
- TP and SRP appear to respond to measures (combinations of all types of measures) within less than five years following the start of the post implementation period (Table 3B). As a result of this perceived "success", the studies reviewed in Table 3B do not expand on phosphorus legacies. However, this "success" must be interpreted in the context of the studies reviewed in Table 2, whereby a wide range (i.e. 2-26 years) can be observed of post-implementation period without detecting a TP or SRP response to measures (See Section 4.1).
- Sediment lags ranged from 7 to 16 years (Table 3C). Compared to other pollutants and for the same types of measures and range of sizes of catchments, this finding suggests that a longer time may be required for a sediment response to measures ("resistance to measures") than for other pollutants. However, one study reported compliance of sediment with water quality standards within 5 years of implementation (Makarewitz et al., 2009). This illustrates that available evidence on lags is catchment-specific and generalisations are impractical.
- Nitrate lags ranged from 1 to 25 years (Table 3D and E). It is unclear whether nitrate response is faster in smaller catchments, but soil properties and tile drainage were key determinants of the faster response. For example, shorter lags can be observed for sandy catchments (Table 3E) and where tile drainage can help remove past (historical) and biogeochemical legacies (Van Meter and Basu 2017). The authors of studies reviewed in Table 3D and E highlighted the positive role of Source control measures in reducing nitrate and the importance of past legacy effects and hydrologic legacy effects as causes of longer lags, in line with our review (APPENDIX II.2-5).

The range of lags in the studies reported in Table 3 should be compared with the findings of the recent review by Melland et al. (2018), who found that the time needed for detecting the first significant water quality response to the measures ranges between 1 and more than 10 years after the measures are implemented. **Our review shows that a longer time may be needed and that it cannot be concluded with certainty what the range of lags can be.**

4.3 What is the relationship between lags and catchment factors?

4.3.1 Lags and type (intended effect) of measures

Shorter lags (<5 years) for all pollutants were observed when a combination of different types (intended effect) of measures were implemented (Table 3: Meals et al., 2001; Simon and Makarewitz 2009; Makarewitz et al., 2009; Line et al., 2016).

Source control

The implementation of *Source control* measures is associated with reduced amounts of nutrients and FIO available for mobilisation and delivery at a farm-scale (Rittenburg et al., 2015). For example, extensive implementation across a catchment of manure management (including slurry storage)and measures preventing livestock access to watercourses reduced nutrients and FIO downstream of grazed farmland (Table 3: Lewis et al., 2019; Meals et al., 2001; Makarewitz et al., 2009; Van Meter and Basu 2017; Willcock et al., 2013).

Livestock exclusion measures from the riparian zone through fencing were found to be effective in reducing TP, TN and ammonia (Table 3: Line et al., 2016). Similarly, reduction in fertilizer application (optimizing to the crop requirements) and timing (application during plant growth periods) can reduce the amount of excess nutrients such as SRP, TP, TN or nitrate available for leaching or surface runoff (Table 3: Dupas et al., 2018; Hansen et al., 2019; Makarewitz et al., 2009; Simon and Makarewitz 2009; Van Meter and Basu 2017).

Box 3. Combinations of measures implemented in studies that reported a water quality response to measures. References from Table 2.

Source control measures • Reduction in N and P fertilisers • Optimisation of N • Manure storage • Livestock restrictions (fencing) • Reduction of stocking rates • Farmyard runoff control • Removal from crop production • Irrigation management In-field *Transport control* measures • Catch crops Riparian *Trapping* measures • Constructed wetland • >8 m Riparian buffers • Riparian restoration with woody vegetation	*Source control* measures • Livestock exclusion (from riparian zone (fencing) with off-site watering, bridges, crossings • Improved farm infrastructure Riparian *Trapping* measures • Streambank stabilisation (incl. revetments) • Riparian restoration (2-8m up to 30m) with woody vegetation	*Source control* measures • Livestock exclusion with off-site watering, bridges, crossings In-field *Transport control* measures • In-field erosion control Riparian *Trapping* measures • Streambank stabilisation • Riparian restoration

The method and timing of slurry application are also important for soil FIO loadings. For example, Hodgson et al. (2016, study not included in Table 3) observed significant reductions in soil FIO following broadcast application of slurry to grassland soil surface because of die-off upon exposure to UV radiation and significant reductions in FIO available for surface runoff following shallow injection. It can be assumed that FIO die-off on site will prevent FIO delivery to receiving waters.

In-field Transport Control and riparian Trapping

In-field *Transport Control* and riparian *Trapping* can immobilise nutrients through crop uptake, precipitation or adsorption on soil matrix, and sediment erosion control (Davey et al., 2020; Table 3: Makarewitz et al., 2009; Simon and Makarewitz 2009; Hansen et al., 2019; Meals et al., 2001; Meals and Hopkins 2002; Wilcock et al., 2013). These measures have the potential to reduce biogeochemical legacy effects and delay transport to receiving water to achieve pollutant immobilisation or removal, e.g. through denitrification (Chen et al., 2019) or FIO die-off (Kay et al., 2012).

Observed lags refer to a combination of different types of measure in each study (Table 3, Box 3). In this respect, it is impossible to infer lags per type of measures or a particular measure. It must be noted that these combinations of measures were also implemented in studies that reported a lack of response to measures (see Table 2). However, our review agrees with the conclusion by Melland et al. (2018) that lags in water quality response to measures are broadly shorter in response to the combined implementation of *Source control*, *Transport control* and *Trapping* when these address pollutant sources, pathways and delivery at impacted sites.

4.3.2 Lags and the threshold of implementation

It must be noted that the right level of implementation across a catchment for a water quality response to occur is not fixed or known (Kroll et al., 2019; Meals et al., 2010; Melland et al., 2018; Stets et al., 2020; Steinman et al., 2018).

Extent, i.e. level of implementation required for a water quality response.

Based on our review, studies with long-term data on gradual uptake of measures show that a water quality response is more likely to be detected with extensive and spatially integrated implementation of measures across a catchment (Davey et al., 2020; Lewis et al., 2019; Wilcock et al., 2013). Studies that compared many different catchments (Makarewitz et al., 2009; Wilcock et al., 2013; Dupas et al., 2019; Van Meter and Basu 2017) observed a range of outcomes for similar measures and pollutants but different ways of uptake of the measures (e.g. gradual versus voluntary, with voluntary implementation being gradual and not occurring at the same time) and level of uptake.

Distribution: Integrated implementation.

Evidence from small catchments (<50 km^2) showed that lags were shorter where multiple spatially integrated measures were implemented (Line et al., 2016; Makarewitz et al., 2009; Meals et al., 2001).

Start time of the post-implementation period

Different start times for the implementation of *Source control* measures such as fertiliser application control and livestock measures as a result of different rates of uptake by farmers may further complicate explanations of differential pollutant improvement (Table 3: Line et al., 2016; Wilcock et al., 2013). Voluntary implementation of measures may contribute to asynchronous (not occurring at the same time) implementation. For example, Line et al. (2016), who applied a BEFORE-AFTER/CONTROL-IMPACT (BACI) design to assess water quality response to measures, highlighted problems related to farmers' resistance to adopting livestock restriction measures (i.e. fencing) at the same time and at the extent initially planned; for this reason, they only used data from the main stem of the river and not from tributaries.

4.3.3 Lags and targeting

Targeting hydrologic processes

All studies reviewed in Table 3 acknowledged the role of: (i) targeting critical source areas; (ii) accounting for hydrologic legacy effects when targeting the measures, especially in-stream/channel sediment storage; and (iii) implementing a combination of *Source control*, *Transport control*, and *Trapping* measures at both the farm-plot and catchment scales. Wilcock et al. (2013) emphasised the role of targeting hydrologic paths. Van Meter and Basu (2017) reported that tile drainage was the key determinant of relatively shorter lags in areas with past legacy effects. Many studies reported in Table 3 mentioned targeting as a factor contributing to a faster response but did not explicitly report the evidence or the criteria their targeting was based on.

4.3.4 Lags and effectiveness of measures at the catchment scale

Studies reviewed in Table 3 attributed relatively fast responses (i.e. within five years) to aspects of effectiveness[11] of measures at catchment scale, i.e. extensive and spatially integrated implementation, targeting to match pressures to biogeochemical and hydrologic processes at farm scale and application of a combination of *Source control*, *Transport control* and *Trapping* across the landscape.

4.3.5 Lags and efficiency[11] of measures

The magnitude of improvements in response to measures determines the efficiency of measures. The magnitude due to measures must be larger than that related to inputs of pollutants due to precipitation, slope or point source discharges, or seasonal/climatic variability. Wilcock et al. (2013) observed that the 3-7 years climatic variation (due to the South Oscillation Index (SOI) (which is associated with changes in precipitation regime and temperature) masked response to measures and lowered their efficiency. Dupas et al. (2019) observed that in Brittany's rivers the phosphorus improvements due to reductions in point source discharges were larger than those due to the measures. They also observed that unveiling the interannual variation due to the North Atlantic Oscillation required more than five years of monitoring (Dupas et al., 2019). It must be noted that only long-term monitoring could unveil such interannual and decadal effects on the efficiency of measures.

APPENDIX IV presents efficiencies of measures in the studies reviewed (where possible). Efficiencies ranged widely, as expected, but were above 20% at the time of detecting the first significant improvement and in some cases reached 90% (e.g. for FIO after 19 years of monitoring, Lewis et al., 2019). This result may be related to the monitoring data, as there is a minimum detectable change based on the frequency and number of samples available and depending on the efficiency of measures (Meals et al., 2010; Akoumianaki et al., 2016). By contrast the efficiencies of the measures implemented in the studies that reported lack of response were generally low (APPENDIX IV: Davey et al., 2020; Pearce and Yates 2017).

11 APPENDIX I.1.

4.3.6 Lags and catchment size

Table 3 and Figure 5 show that there is wide range of lags for a given range of catchment sizes (1-300 km^2). The relationship between lags and catchment size is not consistent for each pollutant. For example, nitrate lags (assessed mostly based on Trend analysis on long-term data) do not increase with catchment size (Figure 5 and Table 3D), which is expected as nitrate lags depend on past legacies and bedrock geology underpinning hydrologic legacies (e.g. Dupas et al., 2019; Van Meter and Basu 2017); see also Table 1 in Section 3.1. On the other hand, TP lags (assessed using a BACI design) decreased with catchment size (Figure 5); see also Table 3B. FIO lags broadly increased with catchment size (coefficient of determination R^2=0.4, based on data from five catchments), but the catchments reviewed had different precipitation regimes and soils and showed different improvement trajectories (Table 3A). Sediment lags showed a weak relationship with catchment size (coefficient of determination R^2=0.32, based on data from four catchments) (Figure 3C). These results must also be assessed in the context of the studies reviewed in Table 2, whereby a lack of water quality response to measures was reported for catchments with similar characteristics and size to those reviewed in Table 3.

A previous review concluded that lags broadly increase with catchment size, but this relationship was relatively weak (R^2<0.5 based on 14 mesoscale catchments, i.e. size<65km^2) (Melland et al., 2018). Our review shows that the relationship between lags and catchment size is not consistent and is partly in agreement with the findings by Melland et al. (2018), i.e. only for FIO and sediment.

The effect of monitoring design is further discussed in Section 4.6.

4.3.7 Lags and land use, slope, soil properties, precipitation and hydrologic paths

Only one study reviewed in Table 3 presents analyses that link lags to catchment factors (Van Meter and Basu 2017). They found that nitrate lags are negatively associated with both tile drainage and catchment slope, with tile drainage being a dominant control of nitrate delivery in autumn and watershed slope being a significant control of nitrate delivery during the spring snowmelt period (Van Meter and Basu 2017). Table 3 presents results from three sub-catchments, where lags ranged between 12 and 25 years. Lags for in-stream nitrate response to measures in all sub-catchments of that large river catchment ranged from 12 to 34 years.

Based on Table 3, there is a wide range of characteristics of catchment factors but also ambiguity in the range of lags per catchment factor. For example, lags in flat catchments ranged between 7 and 15 years for all

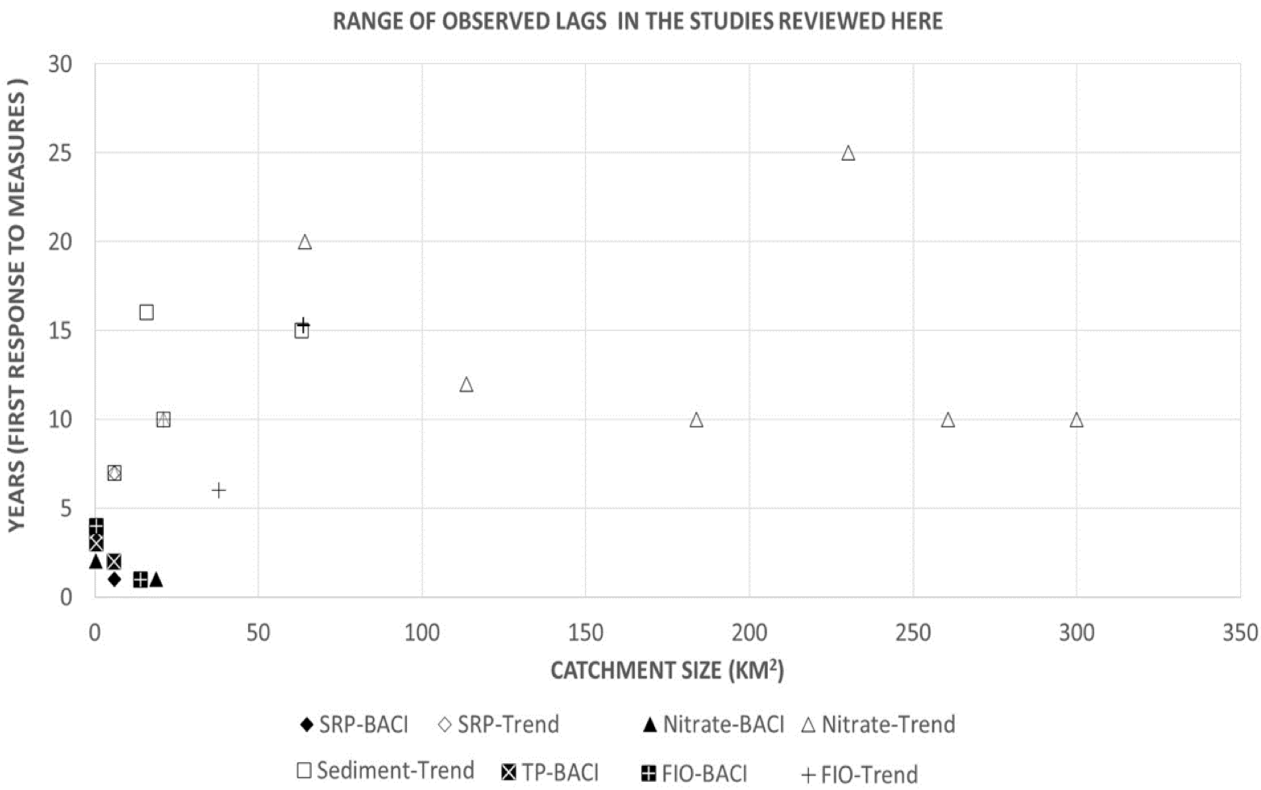


Figure 5. Relationship between time (years) elapsed between the start of implementation of measures and the first detectable water quality response to measures in river waterbodies in temperate regions based on BACI and Trend monitoring designs. No compliance with water quality standards was observed in any of these studies.

pollutants. Lags in steeper catchments ranged between 1 and 25 years but range of sizes of steeper catchment was also very wide (0.4-299.8 km^2). Lags in well-drained catchments ranged between 1 and 25 years.

Groundwater nitrate lags were longer in loamy than in sandy catchments presumably due to longer travel times in the soil matrix in the loamy catchments (Table 3E: Hansen et al., 2019). This finding is in line with the conclusion by Melland et al. (2018) that lags broadly increase with the length or travel time of the hydrologic pathway from sources to receiving waters.

Precipitation was not reported in all studies and information on discharge was referred in a variety of ways, e.g. daily, annual, or average during the course of the study. As a result, comparisons and generalisations are difficult.

4.4 What is the evidence on water quality trajectories in response to measures

Of the 17 studies reviewed to assess pattern in lags, only two mention the term **trajectory and refer to nutrients** (Dupas et al., 2019; Van Meter and Basu 2017); see also Table 3 for in-stream nitrate. These two studies provide explicit data on trajectories of in-stream pollutants in relation to changes in their inputs. These changes refer to the implementation of measures or policies on reducing point sources, atmospheric pollution. Other factors that were shown to affect trajectories were seasonality and interannual (climatic) variability in precipitation. Wilcock et al. (2013) reported long-term data (7-15 years) and presented the evolution of 2-year averages during the course of-15 years (but not in all of the catchments studied). There was limited evidence for linear reduction in pollutants (only TP in one of the catchments of the study) (Wilcock et al., 2013). Hansen et al. (2019) expressed concerns about voluntary uptake of measures as their study demonstrated that a shift of policy from mandatory to voluntary uptake resulted in deterioration in groundwater quality after a period of sustained improvements in response to measures.

Based on the studies reviewed in Table 3, the following remarks can be made on trajectories:

1. The study of trajectories requires long term data in pollutant inputs from all sources, catchment and water quality data.
2. Response to seasonal variation may vary between sub-catchments of the same river catchment.
3. Climatic (interannual) variation such as that determined by climatic periodicity has a similar effect upon sub-catchments in the same region.
4. Trajectories of in-stream pollutant concentrations following implementation of measures are non-linear.
5. Trajectories of in-stream pollutant concentrations are parallel to trajectories of inputs only in the period of increasing inputs and only for inputs from point sources. In-stream nitrate and SRP increase in parallel with increasing inputs from both diffuse and point sources and decrease immediately in response to abatement of point sources (which elicit a lag-less response, as described by Van Meter and Basu 2017). However, long lags in response to abatement of agricultural diffuse pollution through the measures have been observed for nitrate and SRP (Dupas et al., 2019; Van Meter and Basu 2017).

4.5 Do lags depend on the variables we can measure?

Water quality response to measures occurs as a result of the interplay among all catchment factors and their integrated effect on the pollutant transfer continuum (Heathwaite 2010; Li et al., 2013; Lintern et al., 2018; Stets et al., 2020). The systematic study of the integrated effect of all catchment factors on water quality is a key challenge in catchment management (Kroll et al., 2020; Lintern et al., 2018; Meals et al., 2010; Stets et al., 2020). As Lintern et al. (2018) noted: "still required is a better understanding of the interactions and relationships between catchment features".

However, collecting data for all catchment factors and processes influencing water quality is logistically impossible. For example, estimates of lags in groundwater nitrate response to measures require site-specific knowledge of parameters such as: past inputs of nitrogen, crop nitrate, pore water and groundwater nitrate concentrations across hydrogeological gradients, subsurface runoff, soil properties, groundwater recharge, depth of the water table, type and size of the aquifer, denitrification rate, and dispersion and dilution in the aquifer (e.g. Jiang et al., 2017, not included in Table 3 because it referred to farm-scale data). All these parameters are difficult to measure at the appropriate sampling resolution on a site/season- specific basis. A shift of focus can also be observed in the studies reviewed in Table 2 and Table 3 from trying to detect a step change under well designed research/experimental circumstances and without accounting for lags (Meals et al., 2001) towards using a combination of long-term monitoring of inputs and high frequency water quality and modelling approaches to help understand the interplay between catchment factors and the effect of climatic or weather variability (Makarewitz et al., 2009; Pearce and Yates 2017; Wilcock et al., 2013;).

Van Meter and Basu (2017) quantified lags in response to measures in a river catchment in Canada based on multi-decadal data series of nitrogen input and meteorological/streamflow (since 1920) and 25-50-year worth of data of stream water quality and catchment factors (e.g. bedrock geology, soils and tile drainage). Such long-term data series are rare and can rarely be collected.

4.6 What is effect of monitoring (duration/frequency) on lags in response to measures?

It is important to note that statistical time lags (APPENDIX I.1) do not cause the lack of response to measures or the lags but they delay detection of response, thus prolonging the time needed to detect a response to measures. Sampling frequency (number of samples per year) accounting for background variability (e.g. baseflow vs stormflow and wet weather vs dry spells), baseline data (i.e. before the implementation of measures and in control catchments) and long-term data post-implementation are key determinants of our ability to detect the "minimum detectable change" and understand how long it will take to document change[12] (Meals et al., 2010).

The ability to detect a water quality response to measures also depends on the efficiency of the measures: the smaller the improvement expected as a result of the measures (e.g. based on model projections or knowledge of the extent of implementation), the longer the time needed to detect an improvement in water quality against the backdrop of catchment variation. The design, BACI or Trend, used to plan for monitoring and statistically analyse results, also depends on the availability of baseline and control data and the uptake of measures (gradual or completed within a short period of time. Both types of design can be used if there are baseline and control data.

4.6.1 Is lack of a response to measures a result of insufficient monitoring?

The studies reviewed in Table 2 used data from relatively high frequency monitoring, flow weighted concentrations and sampling across a range of flows (APPENDIX IV). For example, Schilling et al. (2011) sampled sediment daily and weekly for 10 years; Zhang et al. (2016) sampled nutrients weekly to fortnightly and took stormflow samples for 26 years. In this context, the design and frequency of monitoring was not considered to be the reason for not detecting a response to measures. Studies that used the BACI design and did not detect a water quality response for certain pollutants, attributed this to uncertainties regarding the start time of the post-implementation time due to asynchronous (not occurring at the same time) uptake across the catchment (Table 2: Davey et al., 2020; Line et al., 2016; Makarewitz et al., 2009; Steinman et al., 2018). Studies that used a trend design (Wilcock et al., 2013) or statistical modelling to detect relationships between the changes brought about by the measures and water quality (Pearce and Yates 2017) highlighted issues with the function of the measures (i.e. the measures were not fully effective at the time of monitoring or because of environmental conditions and to the extent of implementation required for a response.

4.6.2 Are longer lags in response to measures due to monitoring or the design?

Lags in temperate regions (Table 3, Figure 5) were in the range of 1-20 years for river waterbodies and 8-21 years for a groundwater nitrate response. Studies based on a BACI design used 2-4 years of baseline data. The Trend design and analysis were reported in studies with short-term or no baseline data, and were used in cases of gradual implementation of the measures across medium-sized catchments (size: 20-280km^2) and when there was voluntary uptake of measures (e.g. Lewis et al., 2019; Wilcock et al., 2013). Studies using the Trend design were based on long-term (usually over ten years) monitoring and comparisons between sub-catchments.

The results of our review are similar to those by Melland et al. (2018), who reported that the total monitoring time to detect an improvement in water quality (including baseline monitoring before the implementation of measures and post-implementation monitoring) ranges from 4 years (usually 1 or 2 years pre-implementation and 2 or 3 years post-implementation) to more than 20 years if "subsurface/groundwater pathways" are involved.

The consensus among the studies reported in Table 3 is that regardless of design /analysis (BACI or Trend), long-term monitoring of water quality and catchment data especially on nutrient inputs, precipitation, streamflow and drainage are key to understanding water quality response to measures and trajectories and quantifying lags.

All studies highlighted the need for long-term monitoring when assessing the effectiveness of measures and water quality response. Davey et al. (2020) based on their experience from the water quality response to measures

12 Statistical time lags (APPENDIX II.1) in water quality response to measures were discussed extensively in two earlier CREW reports to SEPA (Akoumianaki et al., 2016a;b) and are not further discussed in this report as they do not cause lags in water quality response to measures.

implemented under the Catchment Sensitive Farming (CSF) initiative in England stressed the importance of long term, consistent monitoring programmes for understanding the contribution of the measures to water quality trends. An earlier CREW report recommended that assessing the effectiveness of measures and minimising statistical time lags requires longer than four (4) years of post-implementation monitoring with at least weekly frequency (or finer) of flow-weighted concentrations (Akoumianaki et al., 2016). Long-term studies suggested that at least five (5) years of high frequency monitoring are required to understand the effect of climatic variability due to the North Atlantic Oscillation on water quality response to measures (Dupas et al., 2019).

4.6.3 Can monitoring help to explore policy relevant questions

Monitoring can help to explore a crucial, policy-relevant question (Jarvie et al., 2013; Kroll et al., 2019; Rittenburg et al., 2015; Stets et al., 2020):

- Are pollutant reductions simply not large enough compared to background catchment effects or climate change on water quality to result in downward trends, or are legacy effects from past inputs and lack of matching pressures to local biogeochemical and hydrologic processes introducing additional, significant lags in water quality response to measures?

Box 4 summarises the possible outcomes of monitoring

Box 4. Monitoring outcomes and questions about water quality response to measures

Monitoring can help to explore the following possible monitoring outcomes:

- Did the measure work well by this point in time and we failed to observe a response[13]?
- Did the measures not work yet, and we may or may not have the ability to determine outcomes/ responses should they occur in future[15]?
- Did the measures not work by this point in time and we correctly observed a lack of response[15]?
- What are the catchment factors related to the observed water quality data?
- What is the trajectory of water quality at catchment scale in the context of catchment background variability?
- Is there any evidence on what measures could be retrofitted (expanded, re-installed, supplemented with additional measures) to deliver water quality improvements?

13 Marc Stutter (James Hutton Institute) pers. com. August 2020.

4.7 Can we predict or quantify lags based on catchment typologies?

Despite the practical and policy interest in understanding lags in response to measures, we lack appropriate techniques that can account for the diversity of landscape and management drivers that may impact the time scales over which change may occur (Van Meter and Basu 2017). Lags can be better understood using a more elaborate study design than regulatory programmes (e.g. by including discharge data, a high frequency, event sampling, a Before vs After the measures / Control (pristine or without measures) vs Impact (where the measures are implemented); or long term-monitoring to gather time-series data.

There are studies that calculated long-term (over 100 years) N and P balance (input versus output) trajectories in nutrient loadings and nutrient concentrations but not in relation to measures or catchment typologies (David et al., 2010; Goyette et al., 2016; Howden et al., 2010). More recently, Hashemi et al (2020) provided catchment typologies for catchment-pollutant combinations based on catchment-discharge relationships but their study does not account for lags in response to measures.

Current understanding is that the exact duration of lags for surface water waterbodies can rarely be predicted but may exceed the length of research programmes (usually <3-6 years) and policy cycles in catchment management (usually <5-12 years) (Chen et al., 2019; Gregory et al., 2007; Jarvie et al., 2013; Rittenburg et al., 2015; Meals et al., 2010; Melland et al., 2018; Kroll et al., 2019; Stets et al., 2020; Steinman et al., 2018; Vero et al., 2018. Fenton et al., 2011). The exact duration of lags for nitrate in groundwater can be predicted but only for cases where the complexity of the factors influencing the nitrate transfer continuum is well understood spatially and temporally (APPENDIX II).

Only one out of the 17 studies we reviewed to explore patterns in lags, applied specific statistical methods to quantify lags between reductions in inputs or delivery of pollutants and their concentrations in receiving waters (Van Meter and Basu 2017). Van Meter and Basu (2017) quantified lags through long-term monitoring (25-51 years) and the use of the cross-correlation methodology, whereby they compared reductions in-stream nitrate concentrations with reductions in nitrogen inputs in Ontario, Canada. They quantified the effect and significance of each catchment factor (TN inputs, wetland area, % organic matter, cropland area, slope, population density, extent of tile drainage) on the observed lags based on interannual and seasonal data.

Two studies used indirect methods to estimate lags in response to measures (Davey et al., 2020; Dupaet al., 2019). For example, Davey et al. (2020) used process-based model projections on pollutant loads, Generalised Additive Mixed Models (GAMMs) and sensitivity analysis to quantify the strength of the water quality response to the measures implemented under the Sensitive Catchment Farming (CSF) in England using an assumed time lag between 1 and 5 years. The results showed a "strong" SRP response to CSF within three years but the absence of any detectable water quality response to CSF advice when using shorter or longer lags demonstrated "how a failure to correctly account for lags in the system can lead to the effectiveness of pollution reduction schemes being under-estimated or over-looked altogether" (Davey et al., 2020). In addition, Dupas et al. (2019) estimated breakpoints in long-term data series (50 years) to identify changes in the trajectory of water response to measures and other drivers of change in water quality in the context of discharge variability.

It is also worth noting approaches that account for the interplay of catchment factors influencing water quality. For example, in Scotland, an earlier CREW report developed a methodology based on a Weight-of-evidence approach to help take into account and quantify the interplay between all catchment factors influencing water quality at catchments where DP GBR and SRDP measures are implemented (Akoumianaki et al., 2016a). The method involved quantification of catchment factors such as DP GBR uptake across a catchment, spend for SRDP measures with the potential to benefit water quality, nitrogen and phosphate fertiliser annual inputs, % of high erosion risk area (depending on crop type and land cover); total and grazing livestock density; and deviation of annual rainfall from the 30-year average[14]. the catchment factors proposed in the methodology developed by CREW have been found to be among the key determinants of water quality variability in a catchment (Lintern et al., 2018). The CREW approach can help quantify or understand lags, provided that long term water quality data become available.

Finally, it is useful to draw parallels with the "land type framework" (Figure 4 in Section 3.2) and the suggestion by Rittenburg et al. (2015) for the development of process-based decision-support tools that use readily available catchment and water quality data to advance the land type approach. For example, Brooks et al. (2015) developed a simplified process-based tool building on the land type framework. They included site-specific spatial and temporal catchment (e.g. slope, management scenarios, soil sorption capacity and transformation rates of pollutants) and water quality data. The tool enabled Brooks et al. (2015) to identify the causes of a lack of effectiveness and a longer than 10 years lag in water quality response, despite extensive implementation of measures.

14 Despite the availability of adequate data on catchment factors, the method could not be used for further evaluating their influence on catchment-scale water quality response and lags because of lack of adequate water quality monitoring data.

To sum up:

The review of 17 studies showed there is paucity of empirical evidence and lack of understanding about precisely how long it takes for a water quality response to measures to occur, whether it be the first detectable improvement or the trajectory to the first response or to the end-point of compliance with water quality standards.

All studies in Table 2 and the studies in Table 3 that monitored for longer than 5 years call for: (i) adjusting expectations for the timescales of water quality response; (ii) collecting long-term time series data (longer than 5-10 years) of water quality indicators in combination with catchment factors related to land use, soil properties, runoff, streamflow and precipitation to enable a response to be detected against the backdrop of catchment variation; and (iii) planning for a monitoring programme of at least 10 years post-implementation of the measures for surface water pollutants and much longer (potentially longer than 20 years) for groundwater nitrate) is required to detect a measurable water quality response to measures irrespective of form, type of pollutant and type of measures, and monitoring design.

4.8 Further evidence on policy challenges

Previous reviews on water quality response to measures have also highlighted the paucity of evidence on lags despite their importance in diffuse pollution control but also provide additional evidence in relation to the factors influencing lags (Kroll et al., 2019; Lintern et al., 2018; Meals et al., 2010; Melland et al., 2018). Here, we summarise this additional evidence in the context of policy challenges.

Evidence on measures is site-dependent. Kroll et al. (2019), after reviewing 158 studies reporting timescales of effectiveness and efficiencies of measures, concluded that the localised case-study nature of most studies makes it unreliable to transfer findings to a different region or even a nearby catchment. Cherry et al. (2008) concluded that effectiveness of measures in one farm-plot or catchment should not be extrapolated to predict water quality response in catchments with similar measures in place but different soil type/ land use combinations or spatial scales (e.g. field to farm to catchment to regional scale). Effectiveness of measures at both farm-plot and catchment scales depends on the interplay of catchment-specific factors, e.g. soil sorption capacity, slope, crop type, livestock numbers, climate, hydrologic paths, presence of other source of pollutants and past pollutant inputs (Kroll et al., 2019; Meals et al., 2010; Rickson, 2014). These factors vary in space and time in a catchment-specific way (Agouridis et al., 2005; Chen et al., 2018; Kroll et al., 2019; Lintern et al., 2018; Liu et al., 2017; Osmond et al., 2019; Rittenburg et al., 2015; Rickson 2014; Stutter et al., 2019).

The measures are spatially dispersed. Evidence from small[15] catchments (<50km^2) has shown that there are cumulative benefits from spatially integrated implementation of combined types of measures (Kroll et al., 2019; Meals et al., 2010; Melland et al., 2019; Rittenburg et al., 2015). However, Kroll et al. (2019) observed that the lack of water quality response to measures at the catchment scale is related to the lack of integrated implementation. They observed that the implementation of measures to achieve water quality improvements at catchment scale "is often opportunistic, involving widely dispersed farms throughout large geographic regions" (Kroll et al., 2019).

Catchment complexity. Complexities of pollution sources, and pollutant mobilisation and delivery through river catchments mean that monitored outcomes will take years to decades to confirm successful impacts arising from targeted on-farm implementation of the measures (Collins et al., 2018; Meals et al., 2010; Melland et al., 2018).

There is a call for long-term, high frequency monitoring. A new approach to observing water quality response and accounting for lags is gaining ground: evaluation through long-term regulatory[16] monitoring post-implementation (Lewis et al., 2019; Hansen et al., 2019; Vero et al., 2018; Davey et al., 2020; Wilcock et al., 2013; Van Meter and Basu 2017; Dupas et al., 2019). It is premised on observations that (1) pre-implementation monitoring is not always feasible; (ii) the uptake and implementation of measures is usually gradual and over many years; (iii) legacy effects cause lags ranging from years to many decades; and (iv) climatic/weather phenomena influencing precipitation regime and temperature in an area can have a decadal periodicity. However, there is also a call for finer resolution monitoring (higher frequency and more sampling sites) to address events that occur rarely but have a large impact on water quality, such as storm events (see review by Meals et al., 2010 on magnitude of change) and travel time to groundwater (e.g. Jiang et al., 2017). Securing continuous funding for long-term and high-frequency water quality monitoring remains a challenge (Melland et al., 2018). But there are suggestions for using alternative water quality measurements to allow early estimation of the direction of travel of the measures before the signal reaches receiving waters. For example, Vero et al. (2018) suggested the use of soil pore-water sampling to detect reductions in nitrate in response to *Source control*. Meals et al. (2010) also suggested that in-stream measurement of nutrients can be accompanied by soil nutrient measurements at all farms to demonstrate an immediate response to fertilisation management measures.

Further action may involve retrofitting. In general, there is a consensus that planning for long-term monitoring is the only way to quantify lags (see Section 4.7). It is also the only way to understand whether a lack of water quality response to measures has resulted from poor targeting or low efficiency and inform further action. This action may involve retrofitting the correct measures to site-specific losses (Melland et al., 2018). For example, Tomer et al. (2014) conducted retrospective studies (i.e. studies that examined the suitability of targeting after the measures were implemented) and found that streambank rather than field erosion was the cause of high in-stream sediment following the implementation of in-field erosion (vegetated buffers) measures. Retrofitting involved riparian re-vegetation at local to basin scale to improve stream water quality (Tomer et al., 2014). In addition, adjusting the vegetation and width of riparian buffer strips can help delay the delivery of pollutants though preferential pathways and improve the effectiveness of the buffer strip (Rittenburg et al., 2015; Stutter et al., 2019).

Targeting is narrow-scope and based on assumptions about effectiveness and not local evidence. Chen et al. (2018) found that the lack of appropriate targeting and planning results in greater uptake of measures that cannot address biogeochemical and hydrologic legacy effects. In Europe, and especially in England, narrowly addressing WFD non-compliances has been implicated in lack of targeting and integrated implementation of the measures to deliver temporally consistent water quality responses with wider ecological benefits (Giakoumis and Voulvoulis, 2019). Targeting is largely based on general considerations such as the proximity of grazing pastures to streams, grazing intensity in relation to vegetative cover (i.e., grasses), and livestock access to streams. These considerations **assume without evidence that targeted and well maintained and managed measures will always deliver a water quality response in short timescales (Agouridis 2005; Liu et al., 2017).** Agouridis (2005) argued that site-dependant natural stream processes (e.g. balance between delivered via in-field erosion and removed via streambank erosion) must be accounted when targeting livestock pressures and erosion in-field to ensure that the measures will be effective. Schoumans et al. (2015) observed that there is uncertainty about the true effectiveness of measures, such as riparian buffers:

15 Some studies mention these catchment as mesoscale catchments (Melland et al., 2018) or mini-catchments (Hashemi et al., 2020).

16 Sampling frequencies using regulatory monitoring vary by country and availability of autosampler devices.

most evidence comes from short-term (single rainfall event) plot studies.

Links between performance and function of the measures and lags in response is poorly understood. Chen et al. (2018) observed that Transport control and Trapping measures, which can address biogeochemical legacies, require considerable planning and management. A common issue in the case of *Trapping* (e.g. buffer strips) is pollution swapping (Stutter et al., 2019). Weaver and Summers (2014) identified that for sandy catchments dominated by subsurface nutrient flows, riparian fencing and vegetation were likely to decrease sediment loss, but increase the proportion of bioavailable phosphorus, entering waterbodies. The function and performance of the measures can vary over time irrespective of maintenance due to factors such as the variation of vegetation, natural degradation of structures, and accumulation of pollutants (Liu et al., 2017; Rittenburg et al., 2015; Stutter et al., 2019). Limited empirical data have been collected to describe the performance of the measures against the backdrop of environmental variation. However, growing evidence shows that under certain conditions some in-field Transport control and riparian Trapping measures (e.g. winter stubble, grass strips and riparian buffers) may become sources for dissolved (leached) nutrients (Chen et al., 2019; Dodd and Sharpley, 2016; Meals et al., 2010; Stutter et al., 2019); FIO (Knapper et al., 2012); and sediment (Rickson, 2014).

The link between lags and efficiencies of measures is poorly understood. Efficiencies of measures at catchment scale vary widely (Box 5) in the context of background catchment variability and emerging pressures such as climate change, which translate to small, undetectable, improvements (Kroll et al., 2019). The wide range of efficiencies means that it is difficult to know the minimum detectable change for a specific catchment without monitoring (see Section 4.3.5).

Inconsistency of study approaches. Lags have not been quantified or studied systematically. Results from different studies are often inconsistent based on the monitoring methods employed, especially in terms of forms of sediment and nutrient compounds studied, units of measurement used, spatial and temporal scales and monitoring design (Kroll et al., 2019; Meals et al., 2010; Melland et al., 2018).

Lags are poorly addressed in policy frameworks. Lags have not been quantified; only few long-term, studies have applied proper methodologies to quantify lags[17]. As a result, lags in water quality response to measures are rarely considered in policy frameworks requiring compliance with water quality standards within designated/fixed timeframes as in the WFD (Gregory et al., 2007; Davey et al., 2020; Melland et al., 2018; Kroll et al., 2019; Vero et al., 2018). In 2010, for example, the European Environmental Bureau stated that 'time lags' were simply a 'generic excuse' generated to avoid implementation of more stringent policy measures (Scheure and Naus 2016 cited in Vero et al., 2016). Van Meter and Basu (2017) pointed out that long lags in achieving improvements in water quality can lead to disillusionment among stakeholders regarding the potential for achieving meaningful changes in water quality and, unnecessarily, to calls for investment in more drastic mitigation strategies. **There are concerns among scientists that long lags or lack of understanding or quantification of lags in water quality response to measures increases the risk of reducing funding/ support for long-term monitoring of the effectiveness of measures or promoting voluntary instead of mandatory uptake of the measures** (Hamilton 2012; Jarvie et al., 2013; Meals et al., 2010; Steinman et al., 2018; Van Meter and Basu 2017; Vero et al., 2017). Scientists also point out that new strategies for the effective communication of both the theory and realities of lags in water quality response to measures must be investigated, (Meals et al., 2010; Vero et al., 2017). Osmond et al. (2019) emphasised the need for adaptive management also known as "learning by doing".

17 Van Meter and Basu 2017; Dupas et al. 2019; Davey et al., 2020.

Box 5. Range of efficiencies of measures to reduce pollutants at catchment scale.

A review of 158 studies showed that different combinations of agricultural measures have resulted in (Kroll et al., 2019):

- 3–85% reductions in TN, with reductions of up to 92% for NH4+, 82% for nitrate and 78% for TKN.
- 0–79% reductions in TP, with reductions of up to 91% for SRP.
- 0–90% reductions in total suspended solids.

There are discrepancies between scientific evidence and policy. The literature also highlights the discrepancies between timeframes in current regulations (approximately 6 years) and the decadal timescales associated with the travel times through groundwater (as in the case of nitrate), legacy phosphorus removal, and sediment storage (Jarvie et al., 2013; Meals et al., 2010; Osmond et al., 2019; Rickson 2014; Vero et al., 2018). There is a call for accounting for lags in the design of water quality policies (Kroll et al., 2019; Meals et al., 2010; Vero et al., 2018). More specifically, certain targets and deadlines prescribed by current policies in WFD may need review (Vero et al., 2018).

5.0 Obj. 3: Can lags based on catchment typologies be identified in the literature?

Based on the evidence presented in Sections 3 and 4 no catchment typologies for lags in water quality response could be identified because of:

- Complexity: Multiple interacting catchment factors involved in diffuse pollution control.
- Paucity of long-term water quality and catchment dataset, which are required to quantify lags.
- Localized, case-study nature of most studies on effectiveness of measures and related response.
- Inconsistent reporting of catchment factors in studies of water quality response to measures.
- Lack of knowledge about the function of measures across a range of conditions and environments.
- Inconsistencies in available evidence per type of pollutant.
- Difficulty in quantifying the complex processes determining catchment typologies.

6.0 Recommendations

6.1 Literature-based recommendations

- Account for dominant legacy effects and hydrologic paths when targeting the measures to address pressures.
- Promote spatially integrated implementation of a combination of different types of measures.
- Include *Source control* measures, as their intended effect (i.e. reduce inputs at source) is independent of legacy effects and hydrologic paths.
- Collect long-term monitoring data from catchments where the measures are implemented and from control catchments (pristine, or without measures); control data are key to separating effect of measures from effects of other factors on water quality.
- Apply a Before-After/Control-Impact or trend monitoring design using long-term data depending on availability of pre-implementation data or on mode of uptake of measures (e.g. gradual or not).
- Account for catchment-scale influences on water quality of factors such as rainfall, land use, application of fertiliser, livestock numbers, streamflow, discharges from point-sources, and data on soil sorption capacity and rates of biogeochemical processes.
- Model water quality responses to catchment processes to derive typologies, understand sensitivity to measures over time and guide further action.

6.2 Practical implications for Scotland

- Keep monitoring water quality to help understand lags and inform further action, bearing in mind that monitoring does not cause lags in water quality response to measures.
- Adjust expectations for water quality response and recovery, i.e. there is no evidence supporting fixed timeframes for waterbody improvement.
- Plan for lags in water quality response. This may involve:
  - ✓ Planning for longer-term monitoring of water quality and catchment data, more analyses, and more flexible water quality objectives (ultimately, the universality of lags calls for adaptive management approaches);
  - ✓ Prioritising measures that deliver immediate results by accounting for hydrologic paths e.g. fencing livestock out of streams has been shown to give faster, and in some cases, immediate water quality improvement, compared to waiting for riparian buffers to deliver their intended effect. However, as Meals et al. (2010) highlighted, "Quick-fix" practices with minimum lag time should not automatically replace practices implemented in locations that can ultimately yield permanent reductions in pollutants in the long-term;
  - ✓ Targeting sources nearest to receiving waters;
  - ✓ Taking account of sediment storage and travel processes. For example, in areas where sediment and sediment-bound pollutants from cropland erosion are primary concern, implementing measures that target the largest sediment sources closest to the receiving water may provide a more rapid water quality benefit than erosion controls in the upper reaches of a catchment;
  - ✓ Demonstrating results to the public at areas that are likely to deliver immediate water quality responses, which may be easier at small scales. This may involve focusing on nested sub-catchments within larger river catchments before demonstrating benefits at the larger river catchment scale, e.g. the protected area scale in Scotland. It may also involve focusing on areas with small travel times (i.e. where delivery to streams does not involve groundwater), and where nutrient legacies from the past are unlikely.

- Account for dominant hydrologic paths at farm scale during catchment characterisation surveys and when targeting; this means collecting evidence on soil properties, soil sorption capacity, legacy nutrients, bedrock geology and precipitation regime along with observations on land use and pressures.
- Match the measures to the pollutant(s), pollutant source(s), and hydrologic transport pathways.
- Promote spatially integrated implementation of a combination of measures at farm- and catchment-scale.
- Avoid inputs at source (Source control).
- Consider demonstrating immediate response to Source control, such as fertilisation management measures, by soil nutrient measurements at all farms.
- Consider retro-fitting the correct measure(s) to site-specific losses.
- Develop modelling approaches (e.g. a decision support tool) examining the effect of a suite of catchment factors on water quality. The aim of modelling/tool should be to explore patterns between catchment characteristics and water quality response and understand the causes of observed lags to guide further action. In Scotland, relevant catchment-scale GIS-based data are readily available for desk-top surveys or processing and may include:
  - ✓ Soil leaching potential, Soil runoff risk, Soil compaction, available from Scotland's soils web page;
  - ✓ Land Use/Land Cover data - JHI has free access to 2007 map (UKCEH no date);
  - ✓ Digital elevation model;
  - ✓ Data from uptake of measures per catchment;
  - ✓ Meteorological data;
  - ✓ Streamflow data;
  - ✓ Fertiliser application data, (available from British Survey of Fertiliser Practice);
  - ✓ Livestock numbers, locations, and discharges from point sources;
  - ✓ WFD water quality monitoring data (available form SEPA).
- Investigate strategies for the effective communication of scientific evidence on lags and adaptive management approaches in the context of cost-effectiveness of the measures.

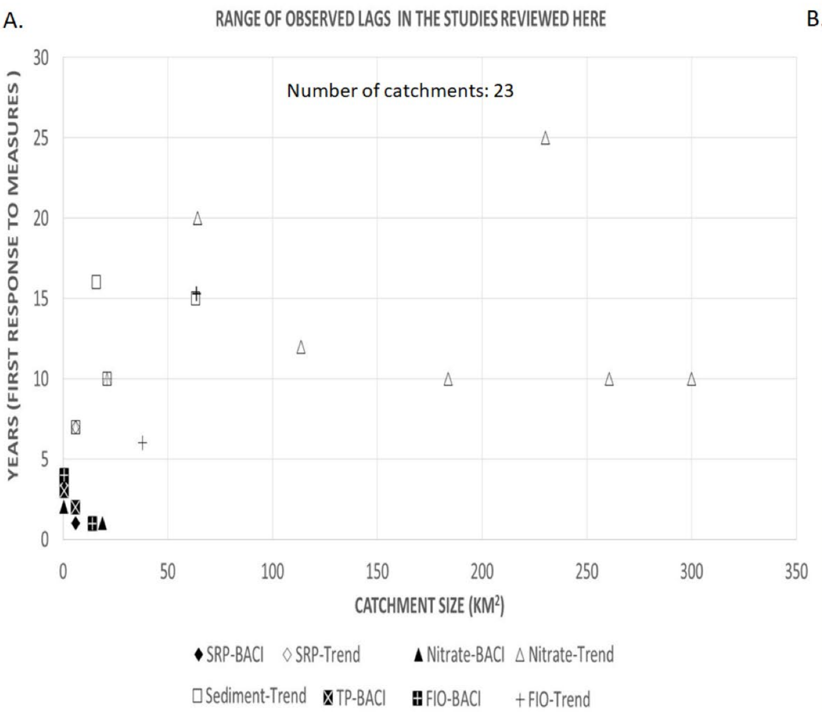
7.0 Conclusion

This project reviewed over 70 studies on riverine and groundwater water quality, mainly focusing on small catchments (<50 km^2) and temperate climates. The overwhelming consensus of the scientific community is that the study of effectiveness of measures at catchment scale and related policies must account for long-term, site-specific lags in water quality response to measures. The key findings are outlined below.

1. There is paucity of empirical evidence and lack of understanding about precisely how long it takes a water quality response to measures to occur, whether it be the first detectable improvement or the trajectory to the first response or to the end-point of compliance with water quality standards. There is no evidence that fixed timeframes for a water quality response to measures can be set. Predicted lags in water quality response based on a catchment typology were not found in the literature. Long-term water quality and catchment data are key to quantifying lags.
2. Lack of water quality response to measures was attributed to combinations of the reasons below:
   - Uncertainties about the effectiveness of measures and the level of implementation required for a water quality response;
   - Low efficiencies of the measures in the context of background catchment variability and pressures such as climate change, which translate to small, undetectable improvements;
   - Lack of effectiveness due to non-optimal implementation of the measures;
   - Longer time required for the measures to become fully effective;
   - Variable function and performance of the measures in response to environmental conditions;
   - Lack of appropriate long-term water quality and catchment data to account for catchment factors;
   - Poor understanding of the start time of the post-implementation period at the catchment scale, which affects statistical analyses and study designs;
   - Monitoring design, which may be introducing a statistical lag, or is unable to detect the magnitude of improvement that has occurred or can occur under site-dependent circumstances.

3. Studies that observed a water quality response to measures, found that lags broadly increase with:
    - Catchment size (Fig. ES1) but differ for the same pollutant in catchments with similar size/measures;
    - Legacy effects from past pollutant inputs stored in the soils and in-stream;
    - Travel time from sources to receptors, e.g. when groundwater hydrologic pathways predominate;
    - Residence time in-field, in-stream and in the aquifer, which is generally enhanced by storage;
    - Presence of multiple, non-agricultural sources (e.g. wastewater discharges) of that pollutant.
4. Lags reported in temperate regions were in the range of 1-25 years for river waterbodies (Fig. ES1. A) and potentially longer than 20 years for a groundwater nitrate response. Studies based on a Before-After/Control-Impact (BACI) design used 2-4 years of baseline data. The Trend design involved long-term (over ten years) monitoring and comparisons between catchments and was used in cases of gradual implementation of the measures across medium-sized catchments (size: 20-300km²). Based on long-term data, river and groundwater water quality trajectories of response to measures are subject to site-dependent seasonal, interannual and decadal, climate-related, variation.
5. Studies that reported a water quality response within five years post-implementation of the measures (Fig. ES1) attributed the relatively fast response to optimal implementation, i.e. extensive and spatially integrated implementation, targeting to match pressures to biogeochemical and hydrologic processes at farm scale and application of a combination of *Source control*, *Transport control* and *Trapping* across the landscape.
6. No catchment typologies for lags in water quality response were identified because of:
    - Complexity: Multiple interacting catchment factors involved in diffuse pollution control;
    - Paucity of long-term water quality and catchment datasets, which are required to quantify lags;
    - Localised, case-study nature of most studies on effectiveness of measures and related response;
    - Inconsistent reporting of catchment factors in studies of water quality response to measures;
    - Lack of knowledge about the function of measures across a range of conditions and environments;
    - Inconsistencies in available evidence per type of pollutant;
    - Difficulty in quantifying the complex processes determining catchment typologies.

A.



B.

Nutrients	FIO
Source control measures • Reduction in N and P fertilisers • Optimisation of N • Manure storage • Livestock restrictions (fencing) • Reduction of stocking rates • Farmyard runoff control • Removal from crop production • Irrigation management In-field *Transport control* measures • Catch crops Riparian *Trapping* measures • Constructed wetland • >8 m Riparian buffers • Riparian restoration with woody vegetation	*Source control* measures • Livestock exclusion with off-site watering, bridges, crossings • Improved farm infrastructure Riparian *Trapping* measures • Streambank stabilisation (incl. revetments) • Riparian restoration (2-8m up to 30m) with woody vegetation **Sediment** *Source control* measures • Livestock exclusion In-field *Transport control* measures • In-field erosion control Riparian *Trapping* measures • Streambank stabilisation • Riparian restoration

Figure ES1. A. Relationship between time (years) elapsed between the start of implementation of measures and the first detectable water quality response to measures in river waterbodies in temperate regions based on BACI and Trend monitoring designs. No compliance with water quality standards was observed in any of these studies. B. Combinations of measures implemented per catchment per pollutant in the studies presented in A.

Key practical implications-recommendations for Scotland based on these findings are listed below.

1. Keep monitoring water quality to help understand lags and inform further action.
2. Adjust expectations for water quality response and recovery, i.e. there is no evidence supporting fixed timeframes for waterbody improvement.
3. Plan for lags in water quality response. This may involve:
    - Planning for longer-term monitoring and flexible objectives as in "learning by doing ";
    - Prioritising measures that deliver immediate results by accounting for hydrologic paths;
    - Targeting sources nearest to receiving waters for faster improvements;
    - Demonstrating results to the public in areas delivering immediate water quality responses.
4. Account for dominant hydrologic paths at farm scale during catchment characterisation surveys and when targeting; this means collecting evidence on soil properties, soil sorption capacity, legacy nutrients, geology, streamflow and precipitation along with evidence on land use and pressures.
5. Match the measures to the pollutant(s), pollutant source(s), and hydrologic transport pathways.
6. Promote spatially integrated implementation of a combination of types of measures.
7. Avoid inputs at source (*Source control*).
8. Consider soil pore-water nutrient measurements to demonstrate effectiveness of *Source control*.
9. Consider retro-fitting the correct measure(s) to site-specific losses when assessment of the measures in place shows that the predominant sources of pollutants have not been addressed.
10. Develop modelling approaches (e.g. a decision support tool) examining the effect of a suite of catchment factors on water quality using readily available desk-based GIS data.
11. Investigate strategies for the effective communication of scientific evidence on lags and adaptive management approaches in the context of cost-effectiveness of the measures.

It is believed that undertaking a systematic review of the literature on water quality response and lags in response to the measures implemented helped to better understand the factors causing lags and the limitations in identifying catchments typologies for lags. It also helped to gather information on what can be done to improve understanding of lags in water quality response. Finally, it provided practical recommendations for Scotland towards accounting for lags in river basin management planning, reducing the lags, and communicating the scientific evidence on lags with practitioners and policy makers.

References

Web sites

Scottish Government (no date). Agri-environment Climate Scheme. Retrieved from: https://www.ruralpayments.org/publicsite/futures/topics/all-schemes/agri-environment-climate-scheme/.

Scotland's Soils (no date). Risk maps. Retrieved from: https://soils.environment.gov.scot/maps/risk-maps/.

UK CEH (no date). Retrieved from: https://www.ceh.ac.uk/services/land-cover-map-2007.

Legislation

Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Union L 327, 22.12.2000, p. 1–73. Available from: http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32000L0060 [Accessed: December 2016].

Water Environment (Controlled Activities) (Scotland) Regulations (CAR No 209) 2011. Available: http://www.legislation.gov.uk/ssi/2011/209/contents/made [Accessed: January 2019].

Grey literature

AKOUMIANAKI, I., PAGALIN, E., Coull, M., AVERY, L. & POTTS J. 2020. Review of monitoring techniques and sampling strategies to identify the most significant sources of Faecal Indicator Organisms (FIO) within a catchment. CRW2018_14.

AKOUMIANAKI, I., POTTS, J. & J, M. 2016b. *Monitoring guidance to assess the effectiveness of the Rural Diffuse Pollution Plan. CD2014/14.* [Online]. Available: https://www.researchgate.net/publication/308504339_Monitoring_guidance_to_assess_the_effectiveness_of_the_Rural_Diffuse_Pollution_Plan [Accessed: May 2019].

AKOUMIANAKI, I., POTTS, J., BAGGIO, A., GIMONA, A., SPEZIA, L., SAMPLE, J., VINTEN, A. & MACDONALD, J. 2016a. *Developing a Method to Monitor the Rural Diffuse Pollution Plan: Providing a Framework for Interpreting Catchment Data. CRW2014/13.* [Online]. Available: https://www.researchgate.net/publication/308505096_Developing_a_Method_to_Monitor_the_Rural_Diffuse_Pollution_Plan_Providing_a_Framework_for_Interpreting_Catchment_Data [Accessed: May 2019].

DEAKIN, J., FLYNN, R., ARCHBOLD, M., DALY, D., O'BRIEN, R., ORR, A. & MISSTEAR, B., 2016. Understanding pathways transferring nutrients to streams: review of a major Irish study and its implications for determining water quality management strategies. In: *Biology and Environment: Proceedings of the Royal Irish Academy* (Vol. 116, No. 3, pp. 233-243). Royal Irish Academy.

DOCHARTAIGH, B., MACDONALD, A., FITZSIMONS, V. & WARD, R. 2015. *Scotland's aquifers and groundwater bodies*. Available: http://nora.nerc.ac.uk/id/eprint/511413/1/OR15028.pdf [Accessed: February 2018].

ENVIRONMENT AGENCY 2019. Catchment Sensitive Farming Evaluation Report – Water Quality Phases 1 to 4 (2006-2018) (NE731). Available online: http://publications.naturalengland.org.uk/publication/4538826523672576. [Accessed: June 2020].

SEPA 2020. Significant water management issues. Available: https://consultation.sepa.org.uk/rbmp/swmi_scotland/supporting_documents/191219_SWMI_final.pdf. [Accessed: June 2020].

WANG, L., STUART, M., BLOOMFIELD, J., GOODDY, D. & BUTCHER, A. 2012. The nitrate time bomb: a national view of nitrate storage and lag time in the unsaturated zone of the UK. [Lecture] In: *Sino-European Symposium on Environment and Health (SESEH 2012), Galway, Ireland, 20-25 Aug 2012* (Unpublished).

WESTGATE, M. 2013. The evidence on adaptive management. Connecting conservation policy makers, researchers and practitioners. *Decision Point Online*, 74. Available: http://decision-point.com.au/article/the-evidence-on-adaptive-management/. [Accessed: July 2020].

Peer-reviewed literature

ALLAIRE, S. E., SYLVAIN, C., LANGE, S. F., THÉRIAULT, G. & LAFRANCE, P. 2015. Potential efficiency of riparian vegetated buffer strips in intercepting soluble compounds in the presence of subsurface preferential flows. *PLoS One,* 10(7): e0131840.

BIDDULPH, M., COLLINS, A., FOSTER, I. D. & HOLMES, N. 2017. The scale problem in tackling diffuse water pollution from agriculture: Insights from the Avon Demonstration Test Catchment programme in England. *River Research and Applications,* 33(10): 1527-1538.

BRIERLEY, G., FRYIRS, K. & JAIN, V. 2006. Landscape connectivity: the geographic basis of geomorphic applications. *Area,* 38(2): 165-174.

BROOKS, E., SAIA, S., BOLL, J., WETZEL, L., EASTON, Z. & STEENHUIS, T. 2015. Assessing BMP Effectiveness and Guiding BMP Planning Using Process-Based Modeling. *JAWRA Journal of the American Water Resources Association,* 51(2): 343-358.

BÜNEMANN, E. K. 2015. Assessment of gross and net mineralization rates of soil organic phosphorus–A review. *Soil Biology and Biochemistry,* 89: 82-98.

Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. *Global Biogeochemical Cycles*, 30 GB005384.

CHEN, D., SHEN, H., HU, M., WANG, J., ZHANG, Y. & DAHLGREN, R. A. 2018. Legacy nutrient dynamics at the watershed scale: principles, modeling, and implications. *Advances in agronomy.* Elsevier, 237-313.

COLLINS, A. & ZHANG, Y. 2016. Exceedance of modern 'background' fine-grained sediment delivery to rivers due to current agricultural land use and uptake of water pollution mitigation options across England and Wales. *Environmental Science & Policy,* 61: 61-73.

COLLINS, A., PRICE, J. N., ZHANG, Y., GOODAY, R., NADEN, P. & SKIRVIN, D. 2018. Assessing the potential impacts of a revised set of on-farm nutrient and sediment 'basic' control measures for reducing agricultural diffuse pollution across England. *Science of the total environment,* 621: 1499-1511.

CONROY, E., TURNER, J., RYMSZEWICZ, A., O'SULLIVAN, J., BRUEN, M., LAWLER, D., LALLY, H. & KELLY-QUINN, M. 2016. The impact of cattle access on ecological water quality in streams: Examples from agricultural catchments within Ireland. *Science of the Total Environment,* 547: 17-29.

DAVEY, A. J., BAILEY, L., BEWES, V., MUBAIWA, A., HALL, J., BURGESS, C., DUNBAR, M. J., SMITH, P. D. & RAMBOHUL, J. 2020. Water quality benefits from an advice-led approach to reducing water pollution from agriculture in England. *Agriculture, Ecosystems & Environment,* 296: 106925.

DAVID, M.B., DRINKWATER, L.E. & MCISAAC, G. F. 2010. Sources of nitrate yields in the Mississippi river basin. *Journal of Environmental Quality*, 39, 1657–67.

DODD, R. J. & SHARPLEY, A. N. 2016. Conservation practice effectiveness and adoption: unintended consequences and implications for sustainable phosphorus management. *Nutrient Cycling in Agroecosystems,* 104(3): 373-392.

DUPAS, R., MINAUDO, C., GRUAU, G., RUIZ, L. AND GASCUEL- ODOUX, C., 2019. Multidecadal trajectory of riverine nitrogen and phosphorus dynamics in rural catchments. *Water Resources Research*, *54*(8), pp.5327-5340.

FENTON, O., SCHULTE, R. P. O., JORDAN, P., LALOR, S. T. J. & RICHARDS, K. G. 2011. Time lag: a methodology for the estimation of vertical and horizontal travel and flushing timescales to nitrate threshold concentrations in Irish aquifers. *Environmental Science & Policy,* 14(4): 419-431.

GAGLIARDI, J. V. & KARNS, J. S. 2000. Leaching of Escherichia coli O157: H7 in diverse soils under various agricultural management practices. *Applied and environmental microbiology,* 66(3): 877-883.

GIAKOUMIS, T. & VOULVOULIS, N. 2019. Water Framework Directive programmes of measures: Lessons from the 1st planning cycle of a catchment in England. *Science of The Total Environment,* 668: 903-916.

GOYETTE, J. O., BENNETT, E. M., HOWARTH, R. W. & MARANGER, R. 2016. Changes in anthropogenic nitrogen and phosphorus inputs to the St. Lawrence sub-basin over 110 years and impacts on riverine export. *Global Biogeochemical Cycles*, *30*(7): 1000-1014.

GRANGER, S., BOL, R., ANTHONY, S., OWENS, P., WHITE, S. & HAYGARTH, P. 2010. Towards a holistic classification of diffuse agricultural water pollution from intensively managed grasslands on heavy soils. *Advances in Agronomy,* 105:83-115.

GREGORY, S., ALLEN, A., BAKER, M., BOYER, K., DILLAHA, T. & ELLIOTT, J. 2007. Realistic expectations of timing between conservation and restoration actions and ecological responses. In Schnepf, M. & C. Cox (Eds): *Managing Agricultural Landscapes for Environmental Quality. Soil and Water Conservation Society*, Iowa, 111-142.

GUO, D., LINTERN, A., WEBB, J., RYU, D., LIU, S., BENDE- MICHL, U., LEAHY, P., WILSON, P. & WESTERN, A. 2019. Key factors affecting temporal variability in stream water quality. *Water Resources Research,* 55(1): 112-129.

HAMILTON, S. K. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. *Freshwater Biology,* 57: 43-57.

HAMILTON, S. K. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. *Freshwater Biology,* 57: 43-57.

HANSEN, B., THORLING, L., KIM, H. & BLICHER-MATHIESEN, G. 2019. Long-term nitrate response in shallow groundwater to agricultural N regulations in Denmark. *Journal of environmental management,* 240: 66-74.

HASHEMI, F., POHLE, I., PULLENS, J.W., TORNBJERG, H., KYLLMAR, K., MARTTILA, H., LEPISTÖ, A., KLØVE, B., FUTTER, M. AND KRONVANG, B. 2020. Conceptual Mini-Catchment Typologies for Testing Dominant Controls of Nutrient Dynamics in Three Nordic Countries. *Water*, *12*(6), p.1776.

HEATHWAITE, A. 2010. Multiple stressors on water availability at global to catchment scales: understanding human impact on nutrient cycles to protect water quality and water availability in the long term. *Freshwater Biology,* 55: 241-257.

HODGSON, C. J., OLIVER, D. M., FISH, R. D., BULMER, N. M., HEATHWAITE, A. L., WINTER, M. & CHADWICK, D. R. 2016. Seasonal persistence of faecal indicator organisms in soil following dairy slurry application to land by surface broadcasting and shallow injection. *Journal of environmental management,* 183: 325-332.

HOLMAN, I. P., HOWDEN, N. J., BELLAMY, P., WILLBY, N., WHELAN, M. J. & RIVAS-CASADO, M. 2010. An assessment of the risk to surface water ecosystems of groundwater P in the UK and Ireland. *Science of the Total Environment,* 408(8): 1847-1857.

HOLMAN, I. P., WHELAN, M. J., HOWDEN, N. J., BELLAMY, P. H., WILLBY, N., RIVAS- CASADO, M. & MCCONVEY, P. 2008. Phosphorus in groundwater—an overlooked contributor to eutrophication? *Hydrologic Processes: An International Journal,* 22(26): 5121-5127.

HONG, E.-M., PARK, Y., MUIRHEAD, R., JEONG, J. & PACHEPSKY, Y. A. 2018. Development and evaluation of the bacterial fate and transport module for the Agricultural Policy/Environmental eXtender (APEX) model. *Science of the Total Environment,* 615: 47-58.

HOWDEN N J K, BURT T P, WORRALL F, WHELAN M J & BIEROZA M 2010 Nitrate concentrations and fluxes in the River Thames over 140 years (1868–2008): are increases irreversible? *Hydrological Processes*, 24 2657–62.

JARVIE, H. P., SHARPLEY, A. N., SCOTT, J. T., HAGGARD, B. E., BOWES, M. J. & MASSEY, L. B. 2012. Within-river phosphorus retention: accounting for a missing piece in the watershed phosphorus puzzle. *Environmental science & technology,* 46(24): 13284-13292.

JARVIE, H. P., SHARPLEY, A. N., SPEARS, B., BUDA, A. R., MAY, L. & KLEINMAN, P. J. 2013. Water quality remediation faces unprecedented challenges from "legacy phosphorus". *Environmental science & technology*, 47(16): 8997–8998.

KAUSHAL, S. S., MAYER, P. M., VIDON, P. G., SMITH, R. M., PENNINO, M. J., NEWCOMER, T. A., DUAN, S. W., WELTY, C. & BELT, K. T. 2014. Land use and climate variability amplify carbon, nutrient, and contaminant pulses: a review with management implications. *Journal of the American Water Resources Association,* 50(3): 585-614.

KAY, D., CROWTHER, J., STAPLETON, C. M., WYER, M. D., FEWTRELL, L., EDWARDS, A., FRANCIS, C. A., MCDONALD, A. T., WATKINS, J. & WILKINSON, J. 2008. Faecal indicator organism concentrations in sewage and treated effluents. *Water Research,* 42(1): 442-454.

KAY, D., CROWTHER, J., KAY, C., MCDONALD, A., CHRISTOBEL FERGUSON, C., STAPLETON, C. & WYER, M. 2012. Effectiveness of best management practices for attenuating the transport of livestock-derived pathogens within catchments World Health Organization (WHO). In Dufour, A., Bartram, J., Bos, R., Gannon, V., (Eds): *Animal Waste, Water Quality and Human Health*; Available: http://www.who.int/water_sanitation_health/publications/2012/ ch6.pdf. [Accessed: June 2020].

KNAPPETT, P. S., MCKAY, L. D., LAYTON, A., WILLIAMS, D. E., ALAM, M. J., HUQ, M. R., MEY, J., FEIGHERY, J. E., CULLIGAN, P. J. & MAILLOUX, B. J. 2012. Implications of fecal bacteria input from latrine-polluted ponds for wells in sandy aquifers. *Environmental science & technology,* 46(3): 1361-1370.

KROLL, S. A. & OAKLAND, H. C. 2019. A Review of Studies Documenting the Effects of Agricultural Best Management Practices on Physiochemical and Biological Measures of Stream Ecosystem Integrity. *Natural Areas Journal,* 39(1): 58-77.

LEWIS, D. J., VOELLER, D., SAITONE, T. L. & TATE, K. W. 2019. Management Scale Assessment of Practices to Mitigate Cattle Microbial Water Quality Impairments of Coastal Waters. *Sustainability,* 11(19).

LI, S., XIA, X., TAN, X. & ZHANG, Q. 2013. Effects of catchment and riparian landscape setting on water chemistry and seasonal evolution of water quality in the upper Han River basin, China. *PLoS One,* 8(1): e53163.

LINE, D. E., OSMOND, D. L. & CHILDRES, W. 2016. Effectiveness of livestock exclusion in a pasture of central North Carolina. *Journal of environmental quality,* 45(6): 1926-1932.

LINTERN, A., WEBB, J. A., RYU, D., LIU, S., BENDE-MICHL, U., WATERS, D., LEAHY, P., WILSON, P. & WESTERN, A. W. 2018. Key factors influencing differences in stream water quality across space. *Wiley Interdisciplinary Reviews-Water,* 5(1).

LIU, Y. Z., ENGEL, B. A., FLANAGAN, D. C., GITAU, M. W., MCMILLAN, S. K. & CHAUBEY, I. 2017. A review on effectiveness of best management practices in improving hydrology and water quality: Needs and opportunities. *Science of the Total Environment,* 601: 580-593.

MCDOWELL, R., DEPREE, C. & STENGER, R. 2020. Likely controls on dissolved reactive phosphorus concentrations in baseflow of an agricultural stream*. Journal of Soils and Sediments*, 20: 3254–3265.

MEALS, D. & HOPKINS, R. 2002. Phosphorus reductions following riparian restoration in two agricultural watersheds in Vermont, USA. *Water Science and Technology,* 45(9): 51-60.

MEALS, D.W., 2001. Water quality response to riparian restoration in an agricultural watershed in Vermont, USA. *Water Science and Technology,* 43(5): 175-182.

MELLAND, A. R., FENTON, O. & JORDAN, P. 2018. Effects of agricultural land management changes on surface water quality: A review of meso-scale catchment research. *Environmental Science & Policy,* 84: 19-25.

MELLANDER, P.-E., JORDAN, P., SHORE, M., MCDONALD, N., WALL, D., SHORTLE, G. & DALY, K. 2016. Identifying contrasting influences and surface water signals for specific groundwater phosphorus vulnerability. *Science of the Total Environment,* 541: 292-302.

OEURNG, C., SAUVAGE, S., SÁNCHEZ-PÉREZ, J.M., 2010a. Dynamics of suspended sediment transport and yield in an intensive agricultural catchment, Southwest France. *Earth Surface Processes and Landforms*, 35, 1289–1301.

OEURNG, C., SAUVAGE, S., SÁNCHEZ-PÉREZ, J.M. 2010b. Temporal variability of nitrate transport through hydrologic response during flood events within a large agricultural catchment in south-west France. *Science of the Total Environment*, 409: 140–149.

ONDERKA, M., WREDE, S., RODNÝ, M., PFISTER, L., HOFFMANN, L. & KREIN, A. 2012. Hydrogeologic and landscape controls of dissolved inorganic nitrogen (DIN) and dissolved silica (DSi) fluxes in heterogeneous catchments. *Journal of Hydrology,* 450: 36-47.

OSMOND, D., SHOBER, A., SHARPLEY, A., DUNCAN, E. & HOAG, D. 2019. Increasing the effectiveness and adoption of agricultural phosphorus management strategies to minimize water quality impairment. *Journal of environmental quality,* 48(5): 1204-1217.

PEARCE, N. J. & YATES, A. G. 2017. Intra-annual variation of the association between agricultural best management practices and stream nutrient concentrations. *Science of the Total Environment,* 586: 1124-1134.

PEARCE, N. J. & YATES, A. G. 2017. Intra-annual variation of the association between agricultural best management practices and stream nutrient concentrations. *Science of the Total Environment,* 586: 1124-1134.

REISCHER, G. H., KOLLANUR, D., VIERHEILIG, J., WEHRSPAUN, C., MACH, R. L., SOMMER, R., STADLER, H. & FARNLEITNER, A. H. 2011. Hypothesis-driven approach for the identification of fecal pollution sources in water resources. *Environmental science & technology,* 45(9): 4038-4045.

RICKSON, R. J. 2014. Can control of soil erosion mitigate water pollution by sediments? *Science of the Total Environment,* 468: 1187-1197.

RITTENBURG, R. A., SQUIRES, A. L., BOLL, J., BROOKS, E. S., EASTON, Z. M. & STEENHUIS, T. S. 2015. Agricultural bmp effectiveness and dominant hydrologic flow paths: concepts and a review. *Journal of the American Water Resources Association,* 51(2): 305-329.

SCHILLING, K. E., ISENHART, T. M., PALMER, J. A., WOLTER, C. F. & SPOONER, J. 2011. Impacts of land-cover change on suspended sediment transport in two agricultural watersheds. *Journal of the American Water Resources Association,* 47(4): 672-686.

SCHOUMANS, O. F., BOURAOUI, F., KABBE, C., OENEMA, O. & VAN DIJK, K. C. 2015. Phosphorus management in Europe in a changing world. *Ambio,* 44(2): 180-192.

SHARPLEY, A., JARVIE, H. P., BUDA, A., MAY, L., SPEARS, B. & KLEINMAN, P. 2013. Phosphorus legacy: overcoming the effects of past management practices to mitigate future water quality impairment. *Journal of environmental quality,* 42(5): 1308-1326.

SHERRIFF, S. C., ROWAN, J. S., FENTON, O., JORDAN, P., MELLAND, A. R., MELLANDER, P.-E. & HUALLACHAIN, D. O. 2016. Storm event suspended sediment-discharge hysteresis and controls in agricultural watersheds: implications for watershed scale sediment management. *Environmental Science & Technology,* 50(4): 1769-1778.

SIMON R.D. & MAKAREWITZ, J.C. 2009. Impacts of manure management practices on stream microbial loading into Conesus Lake, NY. *Journal of Great Lakes Research,* 35: 66-75.

STEINMAN, A. D., HASSETT, M. & OUDSEMA, M. 2018. Effectiveness of best management practices to reduce phosphorus loading to a highly eutrophic lake. *International journal of environmental research and public health,* 15(10): 2111.

STETS, E. G., SPRAGUE, L. A., OELSNER, G. P., JOHNSON, H. M., MURPHY, J. C., RYBERG, K., VECCHIA, A. V., ZUELLIG, R. E., FALCONE, J. A. & RISKIN, M. L. 2020. Landscape drivers of dynamic change in water quality of US rivers. *Environmental Science & Technology,* 54(7): 4336-4343.

STUTTER, M., KRONVANG, B., HUALLACHAIN, D. O. & ROZEMEIJER, J. 2019. Current Insights into the Effectiveness of Riparian Management, Attainment of Multiple Benefits, and Potential Technical Enhancements. *Journal of Environmental Quality,* 48(2): 236-247.

TOMER, M., SADLER, E., LIZOTTE, R., BRYANT, R., POTTER, T., MOORE, M., VEITH, T., BAFFAUT, C., LOCKE, M. & WALBRIDGE, M. 2014. A decade of conservation effects assessment research by the USDA Agricultural Research Service: Progress overview and future outlook. *Journal of Soil and Water Conservation,* 69(5): 365-373.

TRIMBLE, S. W. 2010. Streams, valleys, and floodplains in the sediment cascade. *Sediment Cascades: An Integrated Approach*: 307-343.

VAN METER, K. J. & BASU, N. 2017. Time lags in watershed-scale nutrient transport: an exploration of dominant controls. *Environmental Research Letters,* 12(8): 084017.

VERO, S. E., BASU, N. B., VAN METER, K., RICHARDS, K. G., MELLANDER, P.-E., HEALY, M. G. & FENTON, O. 2018. The environmental status and implications of the nitrate time lag in Europe and North America. *Hydrogeology journal,* 26(1): 7-22.

WANG, L., STUART, M.E., BLOOMFIELD, J.P., BUTCHER, A.S., GOODDY, D.C., MCKENZIE, A.A., LEWIS, M.A. & WILLIAMS, A.T., 2012. Prediction of the arrival of peak nitrate concentrations at the water table at the regional scale in Great Britain. *Hydrological Processes*, *26*(2), pp.226-239.

WANG, L., BUTCHER, A., STUART, M., GOODDY, D. & BLOOMFIELD, J. 2013. The nitrate time bomb: a numerical way to investigate nitrate storage and lag time in the unsaturated zone. *Environmental Geochemistry and Health,* 35(5): 667-681.

WANG, Y., BIAN, J., LAO, W., ZHAO, Y., HOU, Z. & SUN, X., 2019. Assessing the impacts of best management practices on nonpoint source pollution considering cost-effectiveness in the source area of the Liao River, China. *Water*, *11*(6), p.1241.

WEAVER, D. & SUMMERS, R. 2014. Fit-for-purpose phosphorus management: do riparian buffers qualify in catchments with sandy soils? *Environmental Monitoring and Assessment,* 186(5): 2867-2884.

WILCOCK, R.J., MONAGHAN, R.M., QUINN, J.M., SRINIVASAN, M.S., HOULBROOKE, D.J., DUNCAN, M.J., WRIGHT-STOW, A.E. AND SCARSBROOK, M.R., 2013. Trends in water quality of five dairy farming streams in response to adoption of best practice and benefits of long-term monitoring at the catchment scale. *Marine and Freshwater Research*, *64*(5), pp.401-412

ZHANG, Q., BALL, W. P. & MOYER, D. L. 2016. Decadal-scale export of nitrogen, phosphorus, and sediment from the Susquehanna River basin, USA: Analysis and synthesis of temporal and spatial patterns. *Science of the Total Environment,* 563: 1016-1029.

ZHU, Q., SCHMIDT, J. P. & BRYANT, R. B. 2012. Hot moments and hot spots of nutrient losses from a mixed land use watershed. *Journal of Hydrology,* 414: 393-404.

APPENDIX

APPENDIX I.1 Key concepts and their terminology

Note: This Section is not intended to provide regulatory or legal definitions of terms. Instead, it provides a literature-based definition of terms used in this project.

Measures (as of Programmes of Measures - POMs)

The Water Framework Directive (WFD; 2000/60/EC, OJEC, 2000) set out the requirements for PoMs where an individual waterbody has been classified as below or at risk of not reaching "good ecological status" by 2015 and for preventing further deterioration of the status of freshwater environments. PoMs to tackle agricultural diffuse pollution are implemented as agricultural Best Management Practices (BMP).

Categories of on-farm measures in terms of intended effect

Diffuse pollution control measures installed on land to deliver water quality benefits[18] can be divided into three categories:

(i) Measures avoiding inputs of pollutants at source, i.e. on the farm.

These measures are referred to as: Avoiding (Osmond et al., 2019); BMPs that reduce application of a pollutant at the source (Rittenburg et al., 2015); source management (Rittenburg et al., 2015; Agouridis et al., 2005); nutrient source controls (Chen et al., 2019); practices addressing sources of pollutants (Melland et al., 2019); measures limiting sources of pollutants (Lintern et al., 2019); and source control (Wang, et al., 2019).

Examples include: Manure/slurry storage; nutrient management (i.e., planning the fertiliser amount, rate, timing, and method of application); preventing livestock access to watercourses (including fencing and provision for alternative watering and bridges-crossings for livestock); control and reduce farmyard runoff; and reduction of stocking rates to reduce nutrient and FIO inputs.

This report refers to these measures as: "Source control".

(ii) Measures controlling and delaying mobilisation of pollutants by capturing, retaining or interrupting surface and subsurface transport pathways of pollutants in-field.

These measures are also referred to as: Controlling (Osmond et al., 2019); mobilisation control (Lintern et al., 2019); practices addressing pathways of pollutants (Melland et al., 2019); in-field structural and vegetative BMPs (Rittenburg et al., 2015); cultural BMPs (Agourides 2005); practices addressed at pathways of pollutants (Melland et al., 2019); and BMPs that increase the soil's ability to infiltrate and store water, reducing overland flow (Rittenburg et al., 2015).

Examples include: retention of winter stubble, cover crops, in-field grass strips; hedges; reduced or no-till; and reduction of stocking rates to reduce livestock poaching.

This report refers to these measures as: "in-field Transport control" or simply Transport control.

(iii) Measures intercepting and trapping/retaining pollutants in riparian areas before they reach waterbodies.

These measures are also referred to as: Trapping (Osmond et al., 2019); delivery control (Lintern et al., 2019); and riparian structural and vegetative BMPs (Agourides 2005; Rittenburg et al., 2015; Stutter et al., 2019); nutrient retention and sediment trapping (Stutter et al., 2019); and practices addressed at delivery or impact of pollutants (Melland et al., 2019).

Examples include: vegetated (grass or trees) riparian buffer strips adjacent to grassland and arable fields, constructed wetlands and sediment trap bunds.

This report refers to these measures as: "Riparian Trapping" or simply Trapping.

Agricultural Best Management Practice (BMP)

A technique, process, activity, or structure intended to remove, reduce, delay, or prevent agricultural pollutants from reaching receiving waters (e.g. Strecker et al 2001). This term is commonly reported in international research and regulatory literature to reflect both mandatory and voluntary implementation of these practices.

This report uses the term "measures" when referring to studies using the term BMP, unless otherwise stated.

Effectiveness of measures

The degree to which a measure or a combination of measures implemented achieves its intended effect at (Strecker et al 2001; Kay et al., 2010; Liu et al., 2017):

(i) the farm/field scale, e.g. prevention or reduction of a pollutant's losses from a field to a waterbody; and

(iii) catchment scale, e.g. compliance of receiving waterbodies, such as streams, groundwater, lochs, estuaries, or coastal waterbodies, with water quality as well as ecological (where relevant) standards.

18

Efficiency of measures

How well a measure or combination of measures removes pollutants, usually expressed as changes of pollutants in compared to original conditions, due to the implementation of the measures. It may refer to:

(i) Efficiency at the field scale, e.g. percentage reduction of loadings of pollutants in the soil, percentage reduction of losses from farm plots where the measures are implemented to receiving waterbodies.

(ii) Efficiency at catchment/scale, e.g. percentage change of a pollutant's concentrations at the catchment outlet.

Form of pollutant

This includes specific forms of pollutants that determine their behaviour in the soil/groundwater, stream-bed, in-stream, and brackish, saline, or turbid waters, e.g. particulate versus dissolved, organic versus inorganic, and bioavailable or not.

Water quality response to measures

- Meals et al. (2010): the first measurable (i.e. statistically significant) improvement in water quality in the waterbody downstream of the catchment where the measures are implemented at the level projected to reduce diffuse pollution at a catchment scale.
- Melland et al. (2018): the number of years from when a threshold or maximum rate of implementation of a practice is reported or inferred to have been achieved, to when a (significant) effect on water quality was deduced to have occurred.
- This report: any measurable (i.e. statistically significant) improvement in water quality in the waterbody downstream of the catchment where the measures are implemented at or above the level projected to reduce diffuse pollution at a catchment scale, i.e. at or above the threshold

Lag in the water quality response to measures

(also reported as time lag, lag time, environmental or water quality response time[19], delayed response, memory effect, residence effect, and legacy effect)

The time elapsed between implementation (adoption, uptake, installation) of measures at the level projected to reduce diffuse pollution at a catchment scale and a water quality response to measures in the waterbody downstream of the catchment where the measures are implemented (Fenton 2011; Meals et al 2010;).

19 SEPA's project request

Statistical lag time in water quality response

The time between a water quality response to measures and its detection due to insufficient monitoring design (Meals et al., 2010).

Legacy effects

(mainly referring to nitrogen and phosphorus) pollutants that remain stored in terrestrial and aquatic landscapes due to (Chen et al., 2019):

- Excessive anthropogenic inputs before the implementation of the measures, including fertiliser use, fossil fuel combustion and food and feed import.
- Biogeochemical legacy effects, which arise from the time elapsed from inputs to complete removal from a given landscape via gaseous emission and plant uptake, and export from the catchment through water flow and are associated with biogeochemical nutrient cycling within or among soil/sediment, biota, and water.
- Hydrologic legacy effects, which arise from the hydrologic travel times required for nutrient delivery from the sources to the receiving waters along the hydrologic pathways at the catchment scale.

Hydrologic connectivity

The probability that a certain point in the landscape is capable of transmitting material to another point. For a point to be considered hydrologically connected, it must be generating runoff and transmitting the flow vertically downwards and downslope to the stream/river channel and thereof downstream (Brierley et al., 2006).

Adaptive management

Its basic premise is that as management proceeds, information is collected that improves knowledge of the system being managed. This knowledge is then used to improve future management practice, in an iterative process sometimes described as **'learning by doing'.** Consequently, adaptive management should be a good way to manage systems that are poorly understood. Source: Westgate 2013.

APPENDIX I.2
Literature review methodology

Computerised searches for peer-reviewed and grey literature were performed using web-based search engines such as Google Scholar (GS n.d.); Web of Science (WoS n.d.); and Science Direct (SD 2020). The reason for using three different search engines was to take advantage of the different benefits arising from the use of each one of them. GS enabled the detection of published peer-reviewed and grey literature (e.g. reports from government organisations and regulatory agencies) on the basis of full document searches including results drawn from references; however, the results (in terms of numbers and content) were not 100% reproducible. WoS enabled a detection of peer-reviewed articles tagged for their high scientific impact and close relevance of their title and keywords with the search terms. In addition to the advantages referring to the GS and WoS search engine, SD allowed for the targeted search of whole document, i.e. de-emphasising results from references as in GS.

The following words-phrases were used as search terms or keywords: "best management practices[20]" OR BMP OR land management OR WFD agricultural measures OR agri-environment measures OR measures. Other terms used for further refinement of the results included: Diffuse pollution OR non-point source OR nonpoint source; Catchment OR watershed OR landscape; Response OR lag time OR time lag OR "water quality response time"; Legacy nutrient.

20 See APPENDIX I.1.

Given time constraints, the review focused on peer-reviewed (both review and research) articles and book chapters from SD and WoS related to UK and international evidence on lag times in water quality response to measures/BMPs published post-2010 (including). Grey literature was also used, reports and technical notes referring to any relevant evidence from Scotland, regardless of year of publication. Evidence on response to SEPA's measures was also extracted by searching the web sites of environment protection agencies such as the Irish EPA, the US EPA, and the Environment Agency. Finally, citations from selected articles and reports were also checked to develop an understanding on how evidence and knowledge gaps are interpreted by the researchers themselves and to identify evidence that was not captured by the search engine.

The search string and associated output is presented in Table I.1.

Case studies for Objective 2. Most studies refer to water quality in catchment up to 300 km^2 in temperate regions. Data from larger river catchments or subtropical catchments were included, but only where the measures implemented, or the monitoring design were similar to SEPA's measures. In addition to catchment size, measures and climate, selection criteria for these studies included clear information on: catchment characteristics, such as waterbody type, land use/land cover, hydrology, types and uptake of measures implemented and, where available, soil properties; monitoring design and efficiency of measures. The research articles selected included studies from Ireland, England, Denmark, France, USA, Canada, and New Zealand. Studies reporting lags in water quality response to measures were summarise in a Table and a graph and were discussed at the sense checking workshop.

Table I.1. Search string.

Number	Keywords	Google scholar	WoS	Science Direct
9	"best management practices" OR BMP OR land management OR WFD agricultural measures OR agri-environment measures OR measures	3,930,000	4,656,775	
9.1	Refine: Diffuse pollution OR non-point source OR nonpoint source	24,400	3,018	
9.2	Catchment OR watershed OR landscape		1894	
9.3	Response OR lag time OR effectiveness	6610	334	92 [1 (cited in 80 papers in Google Scholar and 42 in Web of science and Scopus)]
9.4	2010-2020	3360	203	77 [See above: Cited: 23]
	Since 2020	92		

APPENDIX II
Hydrologic paths and forms of pollutants along the pollutant transfer continuum

1. Past legacy effects and multiple non-agricultural sources (Processes at Source[21] in Figure 2). Pollutants may enter catchments through external inputs to the soils (e.g. application of fertilisers, intensively grazed grassland, septic tank soakaways,) or directly into streams through wastewater and stormwater discharges in both urban and rural areas. Atmospheric deposition, wildlife and naturally occurring sources, e.g. bedrock, can constitute sources of pollutants both in soils and in-stream. Excessive anthropogenic nutrient inputs through agriculture and wastewater discharges are major causes of the build-up of legacy nutrients in catchments in soil and streambed, respectively.

2. Biogeochemical legacy effects (Processes referring to Mobilisation[22] in Figure 2). Pollutants in agricultural land may be mobilised through: agricultural activity such as tillage, poaching by livestock and erosion from overgrazing; biogeochemical transformations, e.g. sorption-desorption cycles, mineralization, crop uptake of nutrients, leaching (i.e. loss of non-adsorbed pollutants from the soil to groundwater or receiving streams through subsurface hydrologic pathways); FIO die-off; denitrification; hillslope and gully erosion, and freeze–thaw cycles. In-stream mobilization of pollutants can also occur by streambank erosion, sediment deposition/ resuspension cycles and organic matter decay or nutrient cycling in the water column or the sediment. Warming also increases microbial activity, desorption of phosphorus from sediments, and decomposition and mineralization of organic matter (Kaushal et al., 2014). For pollutants with high sorption potential by most soils and sediments, as in the case of SRP, ammonium and FIO, sinks or storages are ubiquitous. As a result, there is: (i) continuous accumulation of these pollutants in various terrestrial and aquatic landscapes, and (ii) continuous re-mobilisation in runoff, streamflow, or through disturbance of the soil or resuspension from streambed.

21 The amount of inputs is measured at the sites that are considered sources of pollutants along the source-mobilisation-delivery continuum in a catchment, e.g. in the soil in farmland or remote areas and in effluent discharges.

22 Measuring mobilisation of pollutants depends on the form of pollutant mobilised (i.e. nutrient, FIO or sediment) and the biogeochemical processes investigated. Measurement can be in-situ, as in erosion studies (e.g. measurement of suspended sediment in-stream or in runoff), or also involve in-vitro experiments (e.g. FIO die off rates, nutrient crop uptake).

3. Hydrologic legacy effects (Processes referring to Delivery[23] in Figure 2). Hydrologic flow pathways can be surface or subsurface, depending on the sites of delivery of mobilised pollutants within a catchment and hydrologic connectivity. In general, particulate forms of pollutants are largely transported by overland (surface) flow whilst dissolved forms are transported by both surface and subsurface (lateral and vertical) runoff. Dissolved pollutant infiltration along the soil matrix is influenced by mobilisation processes, mainly adsorption potential. Artificial (tile) drainage can cause a preferential, lateral flow path, transporting weakly and non-absorbed pollutants in dissolved form and strongly adsorbed pollutants associated with fine soil colloids to waterbodies (Rittenburg et al., 2015). Sediment and particulate nutrients, such as phosphorus (Sharpley et al., 2013), can accumulate rapidly between rainfall and flood events in the soil, floodplain and streambed and at locations where slope or stream channel geometry serves to lower runoff velocity and streamflow. However, measuring and quantifying the time component of these processes requires long term water quality and catchment data and as well as complex hydrogeological studies (Chen at al., 2018), which are rarely available.

II.1 Surface hydrologic paths

Key facts on delivery of pollutants to receiving waters via overland flow, artificial drainage, and streamflow:

- Delivery of pollutants from sources to receiving waters in overland runoff and in-stream has a relatively short hydrologic travel time ranging from days to months (Chen et al., 2019; Jarvie et al., 2013).
- Particulate and dissolved pollutants entrained in overland flow could be partially deposited onto soil if runoff is slow or filtered out of flow due to the presence of vegetation (hydraulic reduction) (Chen et al., 2019; Rittenburg et al., 2015).
- Dissolved pollutants such as nitrate can be delivered to streams faster in loamy, relatively impermeable soils than in sandy soils due to increased infiltration in sandy soils (Hansen et al., 2019).
- Pollutants deposited in streambed during low-regime streamflow could be stored in the sediment but storage times can be relatively short term, i.e. until the next high-flow event remobilizes them (Jarvie et al., 2012), with residence times of <1 year in-stream in many river systems (Chen et al., 2019; Jarvie et al., 2013). For P, residence time in-stream is estimated to

23 Delivery is measured through measurement of the concentration of pollutants in surface and subsurface runoff, in soil matrix, groundwater and streamflow (stormflow and baseflow). Dyes and isotopes can effectively help trace subsurface pollutant transport pathways.

be less than a year irrespective of flashiness, however sorption to streambed/streambank sediments enhances the influence of sorption-desorption and deposition-resuspension effects prolonging residence time (Jarvie et al., 2013).

- Residence times for FIO in-stream depend on many other factors such as temperature, solar radiation, salinity and die-off rates (extensively reviewed in an earlier CREW report by Akoumianaki et al., 2020).
- Shallower slopes favour net deposition of particulate pollutants and sediment in runoff and streamflow and prolong residence times of pollutants in lowland soils and streambed sediments (Chen et al., 2019).
- Fine sediment and total sediment storage in the river channel are generally higher in main-stem reaches than in tributaries in poorly-drained (flashy hydrographs) catchments (Sheriff et al., 2016).
- Flooding events can increase residence time (by years or decades) of nutrients and sediment in adjacent inundated floodplain or riparian zone (Hamilton, 2012; Sharpley et al., 2013).
- Heavy or extreme rainfall events could remobilise pollutants in runoff and streamflow, thus facilitating their export from the catchments (Jarvie et al., 2012), but as these events are rare, this effect is described as "fast in–slow out" (Trimble, 2010).

II.2 Subsurface hydrologic paths

i. Key facts on transport of pollutants in the zone between above-ground sources and the water table (i.e. the unsaturated zone) via throughflow, vertical or lateral preferential flow and soil matrix flow:

- Delivery of pollutants from sources to receiving waters (i.e. adjacent streams other types of surface waterbodies and the groundwater-water table) has a relatively long hydrologic travel time ranging from months to years or decades (Chen et al., 2019; Jarvie et al., 2013).
- Nitrate and dissolved phosphorus can move through preferential pathways at a similar rate as water and relatively fast as there is not sufficient time for adsorption on the soil particles (Rittenburg et al., 2015).
- Dissolved pollutant movement through the soil matrix can be slow and gradual and may favour removal of certain pollutants such as N through denitrification (Vero et al, 2018), or immobilisation such as phosphorus precipitation into clay minerals or phosphorus adsorption onto the soil matrix, and FIO sorption onto soil particles (Rittenburg et al., 2015; Kay et al., 2012).
- Nitrate and phosphorus forms transport through the unsaturated zone is dependent upon soil hydraulic properties[24] as well as on effective rainfall or recharge, depth of the zone above the water table, and the clay or mineral content of the soil (Sharpley et al., 2013; Vero et al., 2018). Other important factors include the cropping pattern, levels of fertiliser application and the type of fertiliser applied (Vero et al., 2018).
- Due to the variety of soil properties in any single area, both rapid preferential flow and slow matrix flows are frequently observed within the same area (Sharpley et al., 2013; Vero et al., 2018).
- Steeper slopes are associated with a thicker unsaturated zone (low water tables) and therefore longer travel times for nitrate to the water table (Vero et al., 2018).
- As a result of leaching, soil nitrate concentrations have been found to increase with soil depth years after the implementation of N *Source control* measures in areas with thick unsaturated zones and historic excessive agricultural N inputs. This effect is known as the "nitrate time bomb" (Wang et al., 2013).
- For P, the travel time from farm soil to stream may range from 5 to 30 years (Jarvie et al., 2013). For N, the travel time can vary from farm soil to groundwater is in the range of decades (Wang et al., 2011;2012;2013).
- In certain regions of Scotland underlain by Old Red Sandstone and Carboniferous Sandstone (e.g. East Scotland, Northern Highland, Central Belt and Scottish Borders), N travel times in the unsaturated zone have been estimated to exceed 20 and in some cases 50 years, from 2009 (Wang et al., 2011; 2012).

ii. Key facts on transport of pollutants in groundwater:

- Travel time within groundwater begins once pollutants break through the water table and become available for transport within the aquifer (deep groundwater) (Vero et al., 2018).
- Pollutants that can be transported to groundwater are dissolved nutrients such as nitrate (Vero et al., 2018) and SRP (Holman et al., 2008; 2010; Jarvie et al., 2013; McDowell et al., 2020), and potentially leached FIO into shallow groundwater (Knapett et al., 2012).
- Travel time of pollutants in the groundwater depends on the geological characteristics of the aquifer (e.g.

24 i.e. a soil's ability to permit water movement through its pores.

flow type, flow length, water residence time, age, permeability) and may not be uniform within the same groundwater waterbody (Vero et al., 2018). In general, longer water residence times are associated with sandstone aquifer types and shorter water residence times are associated with karstic aquifers (O'Dochartaigh et al., 2015).

- Pollutant residence time can range from several months to years in shallow groundwater and from several years to decades in deep groundwater in many regions (Chen et al., 2019; Jarvie et al., 2013; Vero et al., 2018), including the UK (Wang et al., 2011).
- Prolonged pollutant residence time in groundwater can facilitate removal of pollutants (Meals et al., 2010), e.g. N removal through denitrification (Vero et al., 2018), or P sorption onto the aquifer matrix (Holman et al., 2010). However, residence times can be up to 50 years or longer (Vero et al., 2018; Sharpley et al., 2013; Jarvie et al., 2013).
- In Scotland, certain types of aquifer such as Old Red Sandstone (Fife, Strathmore and Moray) and Carboniferous sandstone (mainly Scotland's Central Belt, and Southern Scotland) are characterised by groundwater residence times often in excess of 60 years (O'Dochartaigh 2011 cited in O'Dochartaigh et al., 2015).

iii. Key facts on delivery of pollutants from groundwater to surface waters via baseflow:

- Pollutants in groundwater enter streams at sites of streambed/streambank–stream interface (a.k.a. hyporheic exchange), or transitional and coastal waters through groundwater upwelling from the sediments.
- Pollutants that can be delivered from groundwater to streams include mainly nitrate (Vero et al., 2018) and solutes (not further discussed here). However, growing evidence shows that considerable amounts of SRP in streamflow can originate from groundwater (McDowell et al., 2020; Holman et al., 2008;). Some studies have shown a link between shallow groundwater SRP concentrations in headwater catchments and baseflow SRP concentrations (Mellander et al., 2016).
- In climatic conditions such as those in Scotland, delivery of pollutants via baseflow is most important in summer or during low-flow periods (Holman et al., 2010). In Scotland, groundwater is estimated to sustain more than a third of the annual flow in all river waterbodies, even in small upland streams, rising to over 60% in some rivers in drier East Scotland (Gustard *et al.*, 1987 cited in O'Dochartaigh *et al.*, 2015).
- The transport of pollutants via baseflow depends on redox conditions in streambed and sediment and sediment properties (e.g. grain size); e.g. oxygen gradients leading to hypoxia in the sediment profile may favour denitrification (Vero et al., 2018), or sediment properties may enhance SRP sorption before release into the streambed (McDowell et al., 2020).

II.3 Forms of pollutants in sources

Sources (soil, stream banks, sediment)

P forms:

(1) Inorganic P, mainly SRP added as fertiliser;
(2) Organic P in litterfall, crop residue, livestock/wildlife faeces, and sewage discharges.

Fully bio-reactive inorganic nitrogen forms (in particulate or dissolved forms):

(1) Oxidized nitrogen from atmospheric deposition (as nitrogen oxides-NO_x) and fertiliser application (as nitrite and nitrate);
(2) Reduced nitrogen from atmospheric deposition (mainly as fine particulate ammonium salts) and fertiliser application (as ammonium in dissolved or particulate forms).

Partially bio-reactive (upon microbially-mediated transformation) nitrogen forms:

(1) Dissolved and particulate organic N (as protein and urea) in litterfall, crop residue, livestock/wildlife faeces, and sewage discharges.

Sediment:

(1) industrial (including mining), and domestic wastewater in urban and rural areas;
(2) Construction activities in urban and rural areas.

FIO*:

(1) domestic sewage discharges from the public network and private septic tanks;
(2) livestock;
(3) manure/ biosolid spreading;
(4) wildlife.

* An earlier CREW report to SEPA delivered an extensive review of catchment FIO sources.

References: Lintern et al., 2018; Rittenburg et al., 2018; Bunemann, 2015; EEA: https://ec.europa.eu/environment/integration/research/newsalert/pdf/IR6_en.pdf.

II.4 Forms of pollutants associated with different mobilisation processes

Mobilisation (soil, stream banks, sediment, runoff, in-stream)

P transformations:

(1) Dissolved P (Orthophosphate) solubilised during rock-weathering;
(2) Crop P (SRP) uptake;
(3) Mineralised P from organic P through microbial decay;
(4) Particulate P, i.e. P adsorbed on clay minerals in soil/sediment particles (thereafter undergoing similar mobilisation processes to soil and sediment);
(5) Precipitated P (immobilised) as non-bioavailable phosphate minerals with aluminium, iron, manganese, or calcium;
(6) Microbial P, immobilised into microbial biomass;
(7) Re-mineralised P through microbial transformation of microbially-bound P to dissolved P;
(8) Leached P into soil solution as dissolved P (SRP) when sorption potential is low.

N transformations:

(1) Microbially-mediated atmospheric N-fixation to produce ammonium;
(2) Microbial uptake of ammonium (immobilisation);
(3) Crop uptake of ammonium and nitrate;
(4) Leaching of excess N in soil into infiltrating water as dissolved nitrate;
(5) Nitrification: microbial transformation of ammonium to crop-available nitrites and nitrates under aerobic conditions;
(6) Ammonification: microbial transformation of nitrate to crop-available ammonium under aerobic conditions;
(7) Denitrification: microbial transformation of water-soluble nitrate into dinitrogen N_2 (atmosphere) under anaerobic conditions;
(8) Ammonia volatilisation: release of water-soluble ammonium into the atmosphere as ammonia;
(9) Adsorption of ammonium on clay particles (thereafter undergoing similar mobilisation processes to soil and sediment).

Sediment mobilisation processes

(1) Hillslope and gully erosion by precipitation (weathering) producing both organic and inorganic particulate material;
(2) streambank erosion by streamflow or due to livestock poaching;
(3) Sediment resuspension and deposition in-stream.

FIO*:

(1) Die-off or growth depending on responses of different types of bacteria to oxygen and nutrient levels, and exposure to solar radiation and salinity;
(2) Leaching is possible, especially in freely drained sediments (Gagliardi and Karns 2000);
(2) Adsorbed FIO undergo similar mobilisation processes to soil and sediment.

* An earlier CREW report to SEPA delivered an extensive review of catchment FIO sources.

References: Lintern et al., 2018; Rittenburg et al., 2015; Bunemann, 2015.

II.5 Forms of pollutants in different delivery pathways

P forms:

(1) Particulate P (adsorbed) is entrained in surface runoff;
(2) Particulate P (adsorbed) adheres on soil matrix;
(3) Dissolved P (non-adsorbed or weakly adsorbed) in surface runoff, preferential flow (vertical or lateral) upon leaching into soil solution and in streamflow;
(4) P adsorbed to very fine colloidal soil particles can travel significant distances, laterally or vertically, and reach groundwater waterbodies;
(5) Particulate P can be taken out of the delivery flow pathway (e.g. runoff and streamflow) by filtration from vegetation, sedimentation and infiltration during slow surface runoff, or sedimentation and deposition during slow streamflow followed by resuspension during high-flow regime;
(6) Under hypoxic (or even anoxic) conditions, P is released from the sediment resulting in the possibility that upwelling groundwater could contribute significantly to baseflow concentrations of P (McDowell et al., 2020);

Nitrogen forms:

(1) Nitrogen in the unsaturated zone, i.e. above the water table, can be transported in subsurface flow pathways (vertical or lateral) through fast preferential flow as nitrite/nitrate (dissolved), or through slow soil matrix flow paths as organic nitrogen sorbed to the soil matrix, where there is opportunity for biogeochemical transformation (see APPENDIX II.4);
(2) Dissolved nitrate in baseflow and streamflow;
(3) Ammonium (usually adsorbed in soil particles) can be found in overland flow, preferential flow, soil matrix flow and streamflow;
(4) Ammonia can be found in streamflow;
(5) Organic N can be found overland flow, in preferential flow, soil matrix flow and stream flow.

Sediment:

(1) Suspended sediments are generally transported by overland flow and streamflow into receiving surface waters;
(2) Finer sediments can sometimes be transported by subsurface flows;
(3) Sediment can be taken out of the delivery flow pathway by filtration from vegetation, sedimentation and infiltration during slow surface runoff or sedimentation and deposition during slow streamflow;

FIO forms:

(1) Freely drained soils, *E. coli* O157:H7 can travel below the top layers of soil for more than 2 months after manure initial application .and can reach the water table of shallow groundwater.

References: Lintern et al., 2018; Oeurng et al., 2010a, b; Perks et al., 2016; Rittenburg et al., 2015; Gagliardi and Karns, 2000.

APPENDIX III Relationships between catchment factors and pollutants

This Section discusses the relationships between each key catchment factor separately and the levels of pollutants along the source-mobilisation-delivery continuum. Table III.1 summarises these relationships.

III.1 Consistent relationships

Positive correlations between land use and inputs of pollutants in soils and streams.

- Intensively grazed grassland and in-stream livestock defecation are sources of both particulate and dissolved nutrients (P and N) and FIO in the soil and in-stream, respectively (Kay et al., 2012; Lintern et al., 2018);
- Application of manure and fertilisers on crop and grazing lands is a source of particulate and dissolved forms of nutrients (Zhu et al., 2012) and FIO (Kay et al., 2012; Hodgson et al., 2016);
- Septic tank soakaways and wastewater discharges are major inputs of nutrients and FIO in soils and in-stream (Heathwaite 2010; Kay et al., 2012).

Negative correlations between non-intensive land uses and rainfall and inputs of pollutants.

- Rainfall is negatively associated with pollutants from point sources due to the dilution effect of rainfall on stormwater and sewage effluent (Kay et al., 2008a);
- Extensive cover (>45%) by forest, wetlands or undeveloped areas are associated with reduced nutrient inputs compared to other land uses and types of cover (Stets et al., 2020; Kay et al., 2008b). The is clearly illustrated in a USA-wide study showing that river nutrient and sediment concentrations were lower in catchments dominated by wetlands and natural grassland than in agricultural catchments, despite the 20-year or longer implementation of BMPs in the agricultural catchments (Stets et al., 2020).

Positive correlations between agricultural land use and rainfall and mobilisation of pollutants.

- Extensive livestock grazing enhances mobilisation of sediment and FIO. More specifically, livestock can increase the mobilisation of sediment and sediment-bound nutrients and FIO enhancing susceptibility of soil to erosion from overgrazing and livestock poaching (Agouridis et al., 2005; Conroy et al., 2016).
- Rainfall can increase mobilisation of sediment-bound nutrients and FIO in surface and subsurface runoff and enhance infiltration in permeable soils (Chen et al., 2019; Kay et al., 2012).

Negative correlations between forest land cover and mobilisation of pollutants.

- Extensive forest land cover can immobilise nutrients, sediment and sediment-bound pollutants through root uptake, precipitation or adsorption on soil matrix, and erosion control (Lintern et al., 2018).

Positive correlations between farmland runoff and rainfall and delivery of pollutants to receiving waters.

- Extent of agricultural land cover is a key determinant of in-stream and groundwater pollutant concentrations (Heathwaite 2010; Lintern et al., 2018; Vero et al., 2018).
- Rainfall can enhance delivery of both particulate and dissolved pollutants via surface and subsurface hydrologic pathways (See also Section 3.1.2 and APPENDIX II). Livestock or application of fertiliser or manure do not affect, in themselves, the delivery of mobilised nutrients and FIO from farmland to watercourses or in-stream. However, the presence of impervious surfaces within a catchment, as

in farmyards, compacted farmland , built-up or deforested areas, facilitates the delivery of pollutants via overland flow pathways and by increasing hydrologic connectivity (see APPENDIX I.1) and streamflow (Lintern et al., 2018). Rainfall is positively correlated with in-stream pollutants from diffuse, mainly agricultural, pollution sources.

Negative correlations between land cover and delivery of pollutants to receiving waters.

- River nutrient and sediment concentrations are lower in catchments dominated by wetlands and natural grassland than in agricultural catchments (Heathwaite 201; Lintern et al., 2010; Stets et al., 2020).

III.2 Inconsistent relationships

Measures

Inconsistencies have been observed in relation to water quality response to measures. For example, Makarewitz et al. (2009) and Simon and Makarewitz (2009) observed that in-stream sediment reductions (within a year) and recovery (regulatory compliance) five years after the implementation of measures such as removal from crop production and gully plugs were not followed by compliance with standards for nutrients and FIO. This is extensively discussed in Section 3.2 (see also APPENDIX IV).

Erosion

Predominance of soils susceptible to erosion increases soil loss, and sediment in runoff and artificial drainage and potentially in streamflow (Rickson, 2014). However, it must be borne in mind that that erosion rate for each soil type varies in space and time as it is determined by complex relationships between (see review by Rickson 2014):

- Soil properties;
- Runoff intensity, which determined sediment transport and deposition processes;
- Frequency of storm events and duration of periods between storm events;
- Hydrologic paths from sediment source sites to watercourses;
- Slope length and gradient; and land use;
- Soil conservation measures; and
- Length, morphology and density of the river network.

For example, annual erosion rates in the UK in silty clay loamy soils ranged from 0.33 to 7.44 t/ha in arable land and from 2.82 to 4.92t/ha in pasture land (Walling et al., 2002; 2003, 2006; Tetzlaff et al., 2013 cited in Rickson, 2014). An additional problem is the discrepancies between erosion rates and sediment concentrations in-stream. These may reflect lack of: (i) targeting critical source sediment areas with erosion control measures (Biddulph et al., 2017); (ii) accounting for deposition of eroded sediment within the catchment before reaching the stream network (Parsons et al., 2004 cited in Rickson 2014); (iii) accounting for streambank erosion which is independent of farm-plot erosion control measures (Schilling et al., 2011); and lack of accounting for weather (event vs non-event) variability (Sheriff et al., 2016). It is also useful to recognise the importance of hydrologic connectivity: fields with low erosion risk may represent a higher environmental risk if the connectivity with the receiving waters is uninterrupted (see review by Rickson 2014).

Finally, Sheriff et al. (2016) suggested the following catchment controls on sediment erosion and transport, which also highlight the importance of site-specific catchment factors:

- Catchment size and shape
- Drainage ratio
- Soil type and location of soils susceptible to erosion in relation to stream network
- Slope and the location of steep slopes in relation to stream network
- Vegetation (area covered, type, temporal fluctuations)
- Stream discharge, with positive correlation between event streamflow and resuspension and streambank/ streambank erosion
- Rainfall duration, with positive correlation between rainfall and in-stream suspended sediment indicative of sediment loss in land runoff.
- Rainfall intensity, with positive correlation between rainfall and in-stream suspended sediment indicative of topsoil loss in areas with high hydrologic connectivity and low groundcover.
- Antecedent (prior to rain) catchment wetness, with positive correlation with in-stream suspended sediment indicative of hydrologic connectivity due to sustained wetness.

Soil sorption capacity

Soil sorption capacity can be correlated positively with sediment-bound nutrient and FIO mobilisation, as nutrients adsorbed on clay and silt particles can be delivered to receiving waters primarily via overland flow (see Section 3.1.2 and APPENDIX II.1-3). For example, Lintern et al. (2018) reviewed evidence showing that in-stream concentrations of ammonium, TN and TP were positively correlated with the % of silt and clay in catchment soils. However, biogeochemical transformations between particulate (adsorbed) and solute (de-sorbed) forms of pollutants vary in space and time and can be determined by complex interactions between rainfall, runoff generation and deposition (Rickson 2014).

Catchment slope

Slope correlates positively with the mobilisation of sediments and sediment-bound pollutants, i.e. higher concentration in overland flow (e.g. Onderka et al., 2012). This is because overland flow has higher velocities on steeper slopes, and therefore has greater erosive and transport power. However, steeper slopes have been found to correlate negatively with dissolved in-stream pollutants such as nitrate and total dissolved solids as a result of interacting catchment physical characteristics such as soil properties (soil texture and soil drainage), morphological variables (drainage density and elongation) and vegetation cover (Li et al., 2013). On the other hand, slower overland flow at shallower slopes allows for particulates to settle out of the flow and be deposited in the catchment before delivery to watercourses, potentially contributing to biogeochemical legacies mobilised through storm events (Lintern et al., 2018; Chen et al., 2019). The lack of consistent relationship between slope and the mobilisation and delivery of pollutants may also be related to the fact that shallow slopes are often used as farmland which can deliver higher amounts of sediments and nutrients compared to steeper natural areas, especially if they are covered by woodland (Lintern et al., 2018).

Table III.1. Summary of how each key catchment factor is related to inputs, mobilisation, and delivery of pollutants at landscape/catchment scales. Studies reviewed examined cause-effect relationships and correlations between a factor and pollutants at locations of inputs, mobilisation and delivery.

+ represent a positive relationship (factor leading to or associated with increase of pollutant); - represent a negative relationship (factor leading to or associated with a decrease of pollutant); 0: No relationship; +/-, or +/0 or -/0 represent evidence that is inconsistent across studies.

Correlations between catchment factors and pollutants	Pollutant inputs (i.e. external sources in soils or in-stream)				Pollutant mobilisation (i.e. soil, streambed, in-stream)				Pollutant delivery (i.e. from sources via surface and subsurface flow pathways)			
Catchment Factors	P	N	Sed	FIO	P	N	Sed	FIO	P	N	Sed	FIO
Livestock	+	+	0	+	+	+	+	+	0	0	0	0
Fertiliser application **Manure application**	+ +	+ +	0 +	- +	0	0	0	0	0	0	0	0
Urban wastewater/Septic tank discharges	+	+	+	+	+	+		+	+	+		+
Forest land cover	-	-	0	0	-	-	-	0	-	-	-	0
Implementation of combined measures	-	-	0	-	+/-	+/-	+/-	+/-	-/0	-/0	+/-	-/0
Geology: erosion	0	0	0	0	+/0	+/0	+/0	+/0	0	0	0	0
Climate: rain duration / intensity	0	0	0	0	+	+	+	+/-	+	+	+	+/-
Catchment size	0	0	0	0	0	0	0	0	+/-	+/-	+/-	0
Elevation/slope	0	0	0	0	+	+	+	0	+/-	+/-	+/-	0
Baseflow contribution	0	0	0	0	0	0	0	0	+/-	+/-	-	+/0
References	1	1	2	3	4	4	5	3	6	6	7	8

References: 1. Zhu et al., 2012; Reviews by Heathwaite (2010) and Lintern et al., 2018; 2. Conroy et al., 2016; Review by Lintern et al., 2018; 3. Agouridis et al., 2005; Hodgson et al. (2016); Kroll et al., 2019; Conroy et al., 2016; Kay et al., 2008a,b; 2012; 4. Agouridis et al., 2005; Reviews by Chen et al., 2018; Lintern et al., 2018; Rittenburg et al., 2015; Akoumianaki et al., 2020 5. Agouridis et al., 2005; Conroy et al., 2016; Reviews by Lintern et al., 2018 and Rickson, 2014. 6. Allaire et al., 2015; Davey et al., 2020; Onderka et al., 2012; Stets et al., 2020; Biddulph et al., 2017, Davey et al., 2020; Reviews by Stutter et al., 2019; Schoumans et al., 2015; and Lintern et al., Akoumianaki et al., 2020 2018; 7. Onderka et al., 2012; Li et al., 2013; Review by Lintern et al., 2018; 8. Agouridis et al., 2005; Kay et al., 2012; Lewis et al., 2019; Hong et al., 2018 (and literature cited therein on groundwater-stream bed bacterial exchange); Akoumianaki et al., 2020

APPENDIX IV. Metadata from selected studies on lags in water quality response.

Table IV.1 Metadata from selected studies on lags in water quality response. Q: Discharge/streamflow; Sed/TSS: Sediment; TON: Total Organic Nitrogen; TDP: Total Dissolved Phosphorus; BACI: Before-After/Control-Impact; GW: Groundwater.

	Study / Country	Monitoring design	Measures	Efficiency (%)
1 & 2	England, UK Davey et al 2020 EA 2019	• Years: 4 Phases of CF implementation (2006-2019) • Parameters: FIO, TP, SRP, Sed, Total Oxidised N • Design: BACI in eight river catchments representative of 69 catchments targeted with CSF. Pre-CSF: 2006-09 Post-CSF: 2010-2018 • Monitoring: o Monthly: 2000-06 o Weekly:2007-today o Control (non-CSF/modelled) =61 sites o Impact (CSF)=49 sites • Additional data: weather, cropping patterns, livestock densities • Analysis: Generalised Additive Mixed Models (GAMMs)	• CSF: Greatest uptake o Fencing o Farm infrastructure o Reduce livestock o Feeders o Artificial wetlands o Nutrient management • Uptake: Voluntary • 67.4% of farms implemented 50% of the measures advised	FIO=4-35* TP=4-21 SRP=3-20 Sed=? Total oxidised N=?
3	Vermont, USA (Meals 2001)	• Parameters: Q, FIO, Nutrients, Sediment • Design: BACI o Pre-BMP: 4yrs o Post-BMP:1yr o Control: no-BMP o Impact: BMP • Monitoring: o Q: Continuous o Nutrients: Composite weekly o FIO: grab twice weekly	• Livestock exclusion/watering/ bridges/culverts/crossings • Streambank stabilisation (incl. revetments) • Riparian restoration (2-8m) with woody vegetation • Uptake: Voluntary (7 farms -considered extensive)	TP: 25 FIO:46-52 TN=? Sediment=?
4	Meals and Hopkins 2002 Vermont	• Parameters: TP • Design: BACI o Pre-BMP:4 years o Post-BMP: 2 years o Control: no BMP o Impact: X2 BMP • Monitoring o P: weekly	See Meals 2001	TP:21%

	Study / Country	Monitoring design	Measures	Efficiency (%)
5 & 6	Conesus Lake Catchment, NY, USA Makarewicz et al., 2009 Simon and Makarewitz 2009 Catchment size <10 km² Six catchments	• Parameters: TP, SRP, TON, TKN, Sed, FIO, Q • Design: BACI Control: pristine Pre-BMP: 9 months Post-BMP:>4yrs • Monitoring o Q=daily o TP, SRP, TON, TKN, Sed, FIO: autosampler weekly composite/flow proportional o TP, SRP, TON, TKN, Sed, FIO: Grab samples • Additional data: Soil and drain data • Analysis: ANCOVA & Trend analysis	Nutrient management Gully plugs Rotations Removal from crop production Improved farm infrastructure *Source control* and combinations of measures most successful	Variable 30-70%
7	N. Carolina USA Line et al., 2016	• Parameters: TP, TN, NH3. nitrate-, Sed • Design: BACI o Control: Non-BMP o Impact: BMP o Pre-BMP: 3.7yrs o Post-BMP:3.7yrs • Monitoring: o Automated and grab sampling o Collection every 2 weeks o for TP, TN, NH3. nitrate-, Sed o Flow-proportional samples during storm events • Analysis: ANCOVA and Least Squares means test	• Livestock exclusion with alternative watering (3 m fencing off riparian areas) on only the main stem – Landowners resisted fencing to the extent recommended • Evaluating the effectiveness of excluding cattle from less area and length than is commonly recommended was a high priority.	TKN= 34 ammonia=54 TN=33 TP=47 TSS=60 No significant change for nitrate-
8	Pearce and Yates 2017 Lake Erie basin Nith and Conestoga sub watersheds, USA	• Parameters: Stream Metabolism (Dissolved Oxygen) and nutrients (TP, ammonium, TN, TDP, SRP, Nitrate). Turbidity and TSS • 13 headwater catchments • Sampling for 2 weeks in Summer 2014 • Range of years post-BMP:3-15 years • No pre-BMP data • Use of BMP metrics (i.e. BMP abundance, BMP location) • Use of GIS mapping to locate BMPs • Nutrient monitoring: grab sampling twice • Flow velocity monitoring • Analysis: Multiple regression between (i) BMP metric and stream metabolism and nutrients, and (ii)stream metabolism and nutrients	• Manure storage • Livestock restrictions • Erosion control structures • 30-m wide riparian buffers	• TP (average >x9 standard, i.e. 0/3mg/l) • ammonium (average)>10 standard, i.e. 0.019mg/l

	Study / Country	Monitoring design	Measures	Efficiency (%)
9	Lewis et al., 2019 Olema Creek Watershed (Tomales Bay, USA)	• Parameters: FIO, Precipitation, Streamflow • Analysis: Trend analysis of a 19-year data set of FIO • Design: No pre-BMP data, One Control site (only wildlife influences) • Monitoring: o at multiple sampling stations at confluences and downstream the measures: o Faecal coliforms (MPN): quarterly (dry season) to 12 times a year both low flow and storm flow	• 40 "stream corridor grazing BMPs" in 28km of stream corridor (1) Livestock fencing, (2) hardened stream crossings, and (3) off stream drinking water systems for cattle • Gradual implementation of measures in three phases, Phase 1 involved targeting	85% reduction >90% reduction
10	Wilcock et al., 2013 New Zealand	• Parameters: Turbidity, TSS, Dissolved Oxygen, pH, nitrate, TN, Filterable Reactive Phosphorus and TP • Catchments: 5 catchments representative of regional soils, rainfall and climate, topography, and farming methods • Design: Stakeholder workshop helped to develop a shared conceptual understanding of the links between water quality, pressures, and flow paths and the most appropriate BMPs • Monitoring: every two weeks for two years at 3 sites in each catchment / thereafter monthly at catchment outlet. • Monitoring time: 7-16 years • Analysis: Trend analysis (Seasonal Kendall test on parameters with LOWESS smoothing and flow-adjustment where needed.	• Stakeholder workshop helped to develop a shared conceptual understanding of the links between water quality, pressures, and flow paths and the most appropriate BMPs • on-farm management actions, e.g. livestock management, farm dairy effluent (FDE) treatment and disposal with greater use of irrigation for treated effluent, and use of nitrification inhibitors) • methods of intercepting runoff from land before entry to natural waters, e.g. use of natural and constructed wetlands, riparian management). • 'Dairying and Clean Streams Accord': fencing waterways, manage effluent effectively and have nutrient management systems that minimised environmental damage	

	Study / Country	Monitoring design	Measures	Efficiency (%)
11	Hansen et al., 2019 Denmark	• Parameter: Nitrate • Additional data: o N fertilizer input, o handling of manure, o crop plans, o yields, o catch crops • 5 catchments underlain by sandy and loamy soils • Monitoring sites (and frequency): o soil water (sandy: 52 samples/yr – loamy: 28 samples/yr) o drainage, o shallow groundwater (6 samples/yr from sandy and loamy catchments) o streams (biweekly) • Monitoring design (1989-2016): e.g. (1) Groundwater o Sandy catchments: 15-20 stations o Loamy catchments: 14-24 stations (2) Surface water o Sandy catchments: 2 stations o Loamy catchments: 3 stations • Analysis: 28-year trend analysed with linear regression (backward and forward trend analysis to detect trend reversals and time lags)	• N mitigation measures (efficiency evaluated by measurement of root zone N leaching reduction) - Max stock density - Guidelines for the handling of manure - Mandatory fertilizer and crop rotation plans - Compulsory growing of catch crops - Statutory norms for manure N utilization - Max N allowance for crops equalling economic optimum - Max N allowance for crops ≈ 10% below economic optimum - 6% obligatory catch crops - Organic farming, wetlands, extensification, and afforestation - Site-specific groundwater protection zones - More catch crops - Better manure handling - 10 m buffer zones - Max N allowance for crops ≈ 15% below economic optimum • Most effective measures for reducing Gw and Stream Nitrate: o increased utilization of N in manure o reduced N allowance for specific crops relative to the economic optimum • Need for: o targeting the measures by considering of farming characteristics and site-specific hydrogeological and geochemical conditions of the subsurface. o Ensure that voluntary approaches to uptake won't slow progress.	• Soil leaching reduction (sandy and loamy soils): 33% initially (1989-1997) but only by 2% later (2004-2016) • Gw nitrate: reduction in the 28-year period (sandy catchments): concentration dropped from 100mg/l to approximately standard levels (50mg/l) but no compliance with the standard) • GW nitrate reduction in the 28-year period (loamy catchments): Compliance reached in 15 years (due to denitrification) • Stream Nitrate (loamy catchments)>Stream nitrate (sandy catchments due to the short residence time of water in the upper groundwater aquifers or in tile drains. • Stream nitrate in sandy catchments: consistently low levels.

	Study / Country	Monitoring design	Measures	Efficiency (%)
12	Schilling et al., 2011 Iowa Walnut and Squaw Creek Watersheds	• Parameter: Sed o Monitoring years: 1996-2005 o Parameters: Stream discharge, Suspended sediment, Rainfall o Design: not truly paired, Control vs Impact (Walnut) o Sediment Monitoring: daily and weekly collection o Analysis: Multiple linear regression (using seasonality, discharge, and the Control sediment as covariates) o Additional data: o Streambank erosion survey o Modelling gross sediment erosion (RUSLE)	• Measure: Prairie reconstruction • Greatest problem for lack of effectiveness: streambank erosion, lack of hydraulically controlled (sand or gravel) source material • Proposed solutions: re-meandering, adding floodplains	Significant reduction in 10 years observed only in October (-36%) and November (-45%)
13	Zhang et al., 2016 Susquehanna River Basin and sub-basins draining to Chesapeake Bay	• Study years:1985-2011 • Parameters: TP, TN, DP, DN • Additional data: o Atmospheric deposition o Fertiliser and manure application o Point-source data • Monitoring: o Daily streamflow o Six sites o Sampled days (25- 40) o Pollutants sampled across full range of streamflows: monthly and 8 stormflow samples	• Largest declines recorded for sub-basins where there were extensive o Manure or fertiliser management for TP o Control of atmospheric deposition for TN	Percentage of declines in pollutants river loadings lower than declines in inputs at source (due to legacy effects)
14.	Britany, Dupas et al., 2018	• Parameters: In-stream Nitrate, TP and SRP. • Monitoring: monthly or bi monthly data • Duration of monitoring: 50 years • Design: Trend	• Declines in Nitrate were associated with the reductions in N inputs under the Nitrate Directive and the Urban Waste Water Management Directive. • Declines in SRP and TP were associated with control of point sources (e.g. wastewater discharges, phosphates in detergent use).	
15	Van Meter and Basu	• Parameter: Nitrate • Long term monitoring stream data (1973-2014) and catchment data (GIS data on tile drainage, slopes, land use, point sources, Dams, bedrock geology) and long term trajectories of N inputs from all sources and climate-related / seasonal variation in discharge • Daily Discharge data • Flow-weighted Nitrate concentrations • Trend analysis	• Uptake of measures: Not mentioned if gradual or not or level of implementation • Measures on N-fertiliser and manure management	

	Study / Country	Monitoring design	Measures	Efficiency (%)
16	Scotland, Tarland catchment Bergfur et al., 2012	• Parameters: ammonium, SRP, nitrate, sediment • Design: BACI • Pre-implementation: 5 years • Post implementation: 5 years • Control: streams with degraded riparian vegetation downstream of areas with mixed farming	• Site 5: Source control: livestock fencing Trapping: riparian broadleaf trees planted • Siter 13: Septic tank removal, Source control: fencing Trapping: constructed wetland, riparian tree • Site 8: Livestock fencing	
17	Steinman	• Parameters: In stream TP, TDS, SRP, Nitrate, Ammonia • Sampling monthly and during three storm events • Design: Comparison between Upstream vs Downstream of wetlands and pre (1.5yrs)- vs post (2yrs)-restoration data • Analyses: ANOVA	• Wetland restoration to slow the flow of water during storm events, thus trapping and retaining sediment and nutrients • Restoration involved (overall restored area 0.45km²): 1. Reconnecting 0.16km² of former pasture land to river by placing a pipe from the river to an excavated detention pond; 2. 0.17km² four detention basins to collect and store water during high flows • Questions because of lack of response: -Are we using ineffective BMPs? -Are we locating BMPs in the wrong areas? -Should we be more patient for the BMPs to become more effective? -Does the intensity of agricultural land use overwhelm the assimilative capacity of the BMPs? -Is there sufficient satisfaction with implementation of the management practice (output) instead of its effectiveness (outcome) that we do not push harder for better outcomes? • Explanation for lack of response: reasons: (1) Restoration is still very recent, and until the restored sites are fully functional, which should take a number of years, it is unreasonable to expect a demonstrable change; (2) the two created wetland restoration sites have relatively small footprints and volume holding capacity compared to the entire watershed; (0.45 km² : of 464 km²). Given the volume of water moving through the Macatawa River, especially during storm events, the ability to detect a signal from the noise may be very difficult at any one particular site; (3) the natural environment is variable, so it will take a number of years to detect a robust trend at any site, regardless of direction; and (4) 2017 was a dry year (43% lower than long-term average), thereby resulting in fewer opportunities for the wetlands to serve as filtering and retention basins to remove transport of pollutants.	

CENTRE OF EXPERTISE FOR WATERS

CREW Facilitation Team

Hydro Nation International Centre

James Hutton Institute

Craigiebuckler

Aberdeen AB15 8QH

Scotland UK

Tel: +44 (0)1224 395 395

Email: enquiries@crew.ac.uk

www.crew.ac.uk

CREW is a partnership between the James Hutton Institute and
all Scottish Higher Education Institutes and Research Institutes.
The Centre is funded by the Scottish Government.