Department
for Environment
Food & Rural Affairs

Evaluation of knowledge gained from the National Demonstration Test Catchment

Final report WT15115

June 2020

JWEP
Joint Water Evidence Programme

Environment Agency

UK Centre for Ecology & Hydrology

The James Hutton Institute

Joint Water Evidence Programme

This is a report of research carried out by Centre for Ecology & Hydrology (CEH), on behalf of the Department for Environment, Food and Rural Affairs.

Authors: Old, G., Hutchins, M., Miller, J., Acreman, M., Bowes, M., Redhead, J., Trill, E., Blackstock, K., and Kuhfuss, L.

© Crown copyright 2019

This information is licensed under the Open Government Licence v3.0. To view this licence, visit www.nationalarchives.gov.uk/doc/open-government-licence/

This publication is available at www.gov.uk/government/publications

Any enquiries regarding this publication should be sent to us at

Defra.helpline@Defra.gov.uk

www.gov.uk/Defra

Contents

List of Figures

List of Tables

Tables in appendices:

Supplementary materials

Supplementary material 1 Evidence extracted from documents containing physical science

Supplementary material 2 Evidence extracted from documents containing social science

Supplementary material 3 Inventory of archived datasets and their completeness

The findings from this evaluation have been compiled into a separate evidence compendium which synthesises findings in an easily understandable and useable format. This evidence compendium is published on Defra ScienceSearch: WT15115

Acknowledgements

Our thanks to Steve Addy (James Hutton Institute) who assisted with the data extraction tables. Blackstock, Kuhfuss and Addy's research capabilities are underpinned by research on natural assets within the Scottish Government's Strategic Research Programme 2016-21. Finally, we are grateful for the useful comments from the Steering Group throughout the project which improved our understanding of the project and how these findings might be most useful.

Executive Summary

Context and purpose: In many river basins in England and Wales diffuse pollution from agriculture is a major pressure that contributes to failures to meet WFD objectives. There are major policy challenges in reducing diffuse water pollution from agriculture where this is known to be an issue. Increasingly, environmental policies are seeking to use voluntary uptake of good practice by all land managers across catchments, backed up by incentives and sanctions. However, knowledge gaps on the effectiveness and targeting of specific mitigations measures currently constrains the ability to prioritise and implement many policies.

The Demonstration Test Catchment (DTC) project was commissioned in 2009 through Defra's Agriculture and Water Quality team to test the hypothesis that 'it is possible to cost-effectively reduce the impact of agricultural diffuse water pollution on ecological function while maintaining food security through the implementation of multiple on-farm mitigation measures'. The project was established as a research platform in four catchments (Eden in Cumbria, Wensum in Norfolk, Avon in Hampshire and the Tamar in Devon/Cornwall) where water quality was known to be compromised by diffuse pollution. These catchments provide good national coverage and representativeness of different physical and socio-economic factors relevant to diffuse pollution.

The primary aim of this review is to evaluate the knowledge gained, on intervention effectiveness for water quality and farmer engagement, from the DTC programme by undertaking a Rapid Evidence Assessment (REA) of the available, published evidence. Secondary aims were included to evaluate the broader physical and social science methodologies that were adopted and the approach for generalising the findings. These secondary aims allowed aspects of the DTC to be evaluated (using a traditional review approach) that would have otherwise been excluded due to the strict inclusion criteria for documents under the protocol of the REA. The REA evidence inclusion and exclusion criteria were defined for physical science using the PICO (Population, Intervention, Comparator, Outcome) and the social science using the SPICE (Setting, Perspective, Intervention, Comparison, Evaluation) frameworks before being discussed and agreed with the project steering committee. The robustness of physical evidence was assessed with respect to reporting, modelling/monitoring method and results whereas the robustness of social evidence was assessed with respect to data collection, analysis and reporting.

Given the complexity of the data sources, the draft findings were shared with the Principal Investigators (PI's) from the Avon/Tamar, Eden and Wensum to pick up any factual errors and important points of context. Two PIs responded, and their comments have informed the results presented below. Although feedback from PI's was sought using questionnaires both decided that it was most appropriate to respond at a higher level by documenting their feedback.

Supporting references are provided using standard citations. DTC material assessed as part of the rapid evidence assessment are referenced using a unique identifier, which are listed in Appendix A, to allow cross-reference to the evidence extraction tables (see Supplementary material 1 and

2). An excel file is also provided (see Supplementary material 3) to illustrate the data collected from each DTC and their completeness.

Main findings:

The focus agreed with the Steering Group and funder was on the efficacy of the agri-environment interventions on water quality; reflecting the fact that the project was sponsored by the Defra agriculture and water quality team. However, the DTC process also invested a lot of resources into understanding the nature of the problem in each catchment, including issues of flooding on ecological status, and in building communities of practice across different sectors and socio-economic configurations. PI responses emphasised that reduced funding for on-farm mitigation measures drove the re-focus of the science to understanding the catchments. The focus of this review excludes much of these wider activities. In particular, it is clear from the wealth of publications arising from the programme and other publications by DTC contractors arising from work funded through other sources that much valuable insight was gained in a number of fundamental areas that are outside the bounds of the REA as designed. In the last 10 years very substantial insights have arisen in terms of legacy pollution, ecological responses and the strong influence of storm events in controlling how these are manifested in DTC rivers. The consequence of identifying an over-riding influence of flooding has been that it masks evidence for cause and effect that might in its absence have been apparent. It should be stressed that any lack of evidence arising from the PICO and SPICE analysis is circumstantial and does not in any way imply neglect of those issues.

Main findings: physical science:

Primary Physical: How effective were DTC agri-environment interventions in DTC catchments for improving and maintaining water quality?

The DTC evidence base supports the conclusion that although DTC interventions can result in an improvement in a range of water quality parameters (reported reductions of >80% sediment, >90% for specific pesticides, and 75% for nitrate and 50% for phosphorus) their effectiveness depends on scale and intervention type. Importantly, in 5 of the 24 intervention cases an adverse impacts on at least one water quality determinant was reported. In several cases low effects were reported where interventions were applied to <20% of the upstream catchment. Thus, at the plot scale examples were found with large effects but at the larger catchment scale effects were often lower. Very few studies presented effects for individual years which meant that changes in effect through time could not be evaluated. No PICO compliant studies reported ecological effects and it is acknowledged that this may reflect the possible complex and long term response of ecology.

SP1: Have DTC monitoring methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

Robust DTC monitoring methodologies have allowed the effectiveness of a small range of interventions to be assessed at a range of scales. Although in most cases overall methodologies are robust the majority of studies lacked adequate baseline and post treatment records. Although the importance of the extent of upstream measures was noted in many instances this was difficult to quantify as they related to linear features or combinations of measures. Although interventions cases were included for plot, field and subcatchment scale only one study included all three. PI's acknowledged the lack of suitable baseline and post intervention monitoring and they described how funding constraints did not allow for a sufficiently long programme. Further, PI's explained how reduced capital funding for measures meant that fewer could be assessed and focus necessarily shifted away from landscape scale assessments.

SP2*:* Have DTC modelling methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

Modelling methodologies have assessed a small number of interventions at larger sub catchment and catchment scales (2-1700km^2) but not at plot or field scales. In summary, the body of work comprises rigorous application of peer reviewed and well-founded models. However, in some cases time periods of model validation were insufficiently long, which is likely a consequence of short monitoring periods. In most but not all cases uncertainties were not explicitly evaluated.

SP4: *What models were applied during the DTC programme?*

PICO-compliant papers covered use of FARMSCOPER, SWAT, INCA-P and CRAFT. INCA-P and SWAT have been applied extensively worldwide. Whilst they can be applied elsewhere they require estimation of parameters not readily measurable which hampers their utility for future prediction. In contrast FARMSCOPER is readily transferable without calibration but is a simplification of more detailed underpinning models and accuracy at a locally-specific level may be limited. Approaches such as CRAFT are powerful for interpretation but have detailed monitoring data requirements.

SP5: *Are models used to represent future scenarios (climatic conditions and landuse change) outside the bounds of the DTC dataset?*

None of the four PICO-compliant modelling studies included consideration of future climate or land use change as part of the specified/reported scenarios. However, PICO relevant studies have made useful contributions and are included in Section 7.

SP6: What evidence is there from the DTC programme that the effectiveness of agri-environment interventions varied between DTC catchments and was this related to differences in the design and/or management of the interventions?

The evidence does not allow a robust assessment of whether the effectiveness of interventions varies across the catchments and whether it is related to their design and/or maintenance. In total the effectiveness of 24 intervention cases was reported (10 from Wensum, 9 from Eden and 5 from the Avon). Given that the interventions in each of the catchments were diverse it is not possible to compare their effect across the catchments. To enable this assessment evidence would be needed on the effectiveness of a specific intervention type -design from several studies in each catchment. As an exception, two model based studies reported different total phosphorus loses from buffer strips in the Eden and Wensum catchments. Furthermore, a non PICO compliant study used novel magnetic tracing methods to show the relative sediment trapping efficiencies of individual buffer strips in 3 of the catchments.

SP7: What evidence is there from DTC data that confounding factors (e.g. climate, non agricultural pollution) may be important in the interpretation of the results?

Robust experimental designs were chosen in all catchments to mitigate the effects of confounding factors. Although confounding factors are often considered conceptually many of the DTC studies do not explicitly quantitatively account for them; the difficulty involved in doing this is acknowledged. However, it is important to note that the effects of several confounding factors (e.g. climatic variability) may have been reduced with longer periods of pre and post intervention monitoring. Accounting for confounding factors is a complex process when considering model applications and in each case the evaluation entails a dependency on the complexity of the model used. The project has provided both data and models that could potentially be used in combination to produce a novel quantification of the importance of confounding factors. It is acknowledged that this exercise would have been beyond the scope of the DTC project.

Main findings: social science:

Primary Social: How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?

The DTC documents illustrated that that engagement and awareness raising is necessary but not sufficient to ensure uptake of measures. However, it is difficult to find evidence of this full chain (from engagement to sustained uptake) being implemented for specific interventions on specific farms within the SPICE compliant documents. PIs suggested that the focus was more on understanding the system and building a community of practice to respond to water quality and flooding issues, rather than measuring the effectiveness of a specific agri-environment intervention.

SS1: *What evidence is there that the DTC engagement methodologies appropriately informed, consulted and actively involved farmers and other stakeholders to maximise uptake of interventions?*

The DTC programme engaged a range of appropriate stakeholders using an array of methods to inform, consult and actively involve these stakeholders. However, there was limited evidence available about the overall reach of the engagement, and little information about the active involvement strategies used when engaging farmers about the specifics of why and how they decided to implement DTC measures.

SS2: *What evidence does the DTC data provide about non-participants and why they did not engage in the process?*

There is very little evidence provided about those who did not participate in the full range of DTC engagement processes; and little explanation about why certain stakeholders did not engage with the project.

SS3: *What evidence is there that the DTC considered behavioural factors when engaging farmers in implementing interventions?*

There is evidence that behavioural factors have been taken into account when designing the project and planning engagement activities. However, the reporting of these behavioural f actors, and the evidence that these might have affected uptake of interventions, is limited.

SS4: *What evidence is there of engagement and uptake of DTC interventions varying between catchments?*

The evidence was often summarised across catchments, making distinctions between catchments more difficult to assess. Engagement does not seem to vary between catchments. There are some differences in uptake of measures; and intentions to uptake further measures, but these tend to be related to farming systems and personal attributes rather than anything specific to the geography of individual catchments, although feedback suggests that the Eden had a different socio-economic context to the other catchments.

SS5: *What evidence is presented on the cost effectiveness and benefits of DTC interventions, during the initiative and for the five-year period beyond the end of the DTC initiative?*

There were few specific cost-effectiveness analyses undertaken and only one single cost-benefit analyses done for a specific intervention (roadside wetlands). There was a reliance on the FarmScoper model for average costs until later in the project, and a tendency to focus on direct private costs, with little information on the wider costs or benefits of an intervention. There is very little information about the potential for measures to be sustained beyond the life of the DTC project, but some sources suggest that payments were not enough to entice farmers to bear the costs of mitigation.

SS6: *What evidence is provided that the confounding factors (e.g. existing non-DTC activities) were accounted for when reporting on engagement and uptake of DTC interventions?*

There was limited evidence to see whether and how confounding factors were considered for engagement, behaviour and uptake. These factors were varied and there were no clear patterns across time or catchments. There was limited evidence illustrating where the DTC project ‘added value’ to other initiatives but the DTC followed good practice by trying to integrate with existing farming and stakeholder networks as much as possible.

Broader evaluation of the DTC:

Evaluation of the monitoring methodologies: The four DTC catchments provide good national coverage and representativeness of different physical and socio-economic factors relevant to diffuse pollution. Experimental designs, selected variables and equipment deployed were appropriate and scientifically robust. However, a greater emphasis on groundwater quality may have allowed a fuller evaluation of interventions. Although data coverage was good at most sites significant gaps were present at others. Sample data allowed ground truthing of continuous nutrient data at most sites but not all. Nutrient concentration-flow relationships suggest that some sub catchments may be affected by point sources inputs of pollutants and further investigations may help understand their water quality. Novel instrumentation was deployed including various types of cameras and sediment tracing techniques. Furthermore, PI's noted that the infrastructure allowed for rapid scientific responses to policy teams and the media.

Evaluation of the modelling methodologies: Within the programme a number of studies were undertaken using widely-applied physics-based models (e.g. SWAT, INCA-P) that describe catchment diffuse pollution mechanisms and include dynamic representation of nutrient and sediment transport in rainfall events. The process-based model CRAFT has also been used. Applications of another group of models including FAMRSCOPER, PIT and PSYCHIC provide stronger emphasis on the integrated outcomes of diffuse pollution. Other models focus on risk and connectivity (e.g. SCIMAP and DBM). The wide spectrum of model applications applied provided a valuable description of hydrological and water quality response. Any reporting of extension of results to address impact of interventions on aquatic ecology however was not found.

A number of other tools were developed for site-specific analysis. They represent transferable concepts, rather than model codes, but if their detailed data requirements are met they provide powerful approaches for a wider understanding of nutrient pollution across a range of environments. For example,

(1) substantial work in the Eden has focused on empirical modelling of phosphorus transfer and the importance of storm event control, developed with capability to simulate response to future change.

(2) High temporal resolution monitoring techniques have been used in conjunction with molecular and compound-specific isotope analysis to develop a suite of Bayesian mixing models for the purposes of apportioning sources of organic matter and stream sediments. Sediment tracing has been undertaken, in particular in the Avon and Wensum.

Models of the sub-surface have also been applied (e.g. the Nitrate Time Bomb in the Eden, and conceptual models for management support in the Wensum),

The off-the-shelf process-based catchment-scale models used in the DTC research are potentially very useful for extrapolating findings to other basins, and are invaluable in quantifying climate variability, the importance of storm events and the influence of confounding factors. They have been widely used elsewhere in the UK, and given acceptable levels of performance these applications together form a pool of evidence from which to consolidate knowledge. Despite their physical basis however, representing a wide range of interventions in process models is not straightforward. Catchment model applications often lack the spatial detail to represent small-scale interventions which is instead often captured through pooling of empirical observations, field scale models and expert judgement. Furthermore, models are typically used to represent future scenarios under steady-state conditions rather than representing detailed gradual change of a catchment system in response to interventions.

Farmer Engagement: The DTC documents illustrate a good understanding of the practical aspects of engagement and illustrate: the importance of planning engagement to build up trust and social networks; the move from information provision to active involvement, and the need to embed farmer discussions in wider existing farmer and stakeholder networks. The gaps in the evidence correspond to how stakeholder analysis was undertaken, decisions made about how and where to focus effort; and the dynamics of stakeholder interactions within discussion groups, particularly regarding 'capture' by dominant interests or effects of information asymmetry.

Uptake of Measures: The DTC evidence illustrates that the researchers understood the many dimensions of farmer behaviour in the catchments and how individual decisions are influenced by others, including trusted advisors. However, there is more evidence around how farmers understood DWPA and their existing personal circumstances than evidence about decision-making to take on new mitigation measures, which corresponds to PI feedback on focussing the stakeholder engagement on understanding the system and building trusted networks. Likewise, connecting the individual farmer's identities to wider institutional drivers (both policy and the supply chain) is important, but the evidence is limited about to what extent these linkages influenced the outcomes.

Generalisation: Several methods have the potential to aid in the generalisation of findings from the DTC catchments. The DTC Catchment Matcher tool was found to be well suited. It assesses the similarity of catchments using >20 variables based on their proximity in ordination space. However, when considering the transferability of findings and the feasibility of implementing interventions it may be beneficial to include additional information such as landscape character and history. Furthermore, the usability of the tool could be enhanced with the addition of a user interface including graphing functions.

Pathways to Added Value:

The DTC linked their work to CSF and CaBa processes but insights may also be relevant to landscape partnerships and to partnerships focussed on natural flood management and conservation finance or payment for ecosystem services.

Despite being a ten year programme, more time was needed to identify, agree and implement measures and to provide evidence of ongoing uptake and impacts on water quality; and thought is needed on how to report learning when dealing with commercial sensitivities. PI's described how at early stages in the project significant effort was invested in knowledge exchange through which priority catchments were identified and permissions sought for installing infrastructure.

PI feedback suggested that once resources were invested in monitoring equipment, the investment remaining for mitigation efforts and scientific analysis were modest in relation to the complexity of the socio-ecological systems within the catchments. With increased pressure on funding throughout the project support became increasingly limited to collecting and uploading data.

Although comprehensive datasets were collected from each DTC there is great potential for added value through further analysis and engagement with current initiatives. Future, monitoring initiatives (e.g. the NERC Flood and Drought Research Infrastructure scoping project 2020-2021) should build on the lessons learnt and data collected from the DTC project.

The interaction between DWPA and flooding needs greater attention, as flood events can mask signals in monitoring data, and deflect stakeholders' attention away from the more invisible and cumulative effects of diffuse pollution.

The costs and benefits of interdisciplinary working are harder to evaluate than physical or social sciences, but the development of human capital and social networks may be important legacies of the DTC process. The DTC's clearly provided a multidisciplinary training platform that facilitated co-working of leading experts that improved the science and policy support.

Concluding recommendations can be found in section 12 covering future modelling and monitoring of interventions; engagement, uptake and setting up transdisciplinary platforms.

Table 1: Acronyms

AES	Agri-Environment Scheme	NE	Natural England
BACI	Before-after-control-impact	NFU	National Farmers Union
BQE	Biological Quality Elements	NGO	Non-Governmental Organisation
CaBa	Catchment Based Approach	PES	Payment for Ecosystem Services
CAP	Common Agricultural Policy	PI	Principal Investigator
CBA	Cost-Benefit Analysis	REA	Rapid Evidence Assessment
CEA	Cost-Effectiveness Analysis	S	Satisfactory
CSF	Catchment Sensitive Farming	SAC	Special Area of Conservation
CSFO	Catchment Sensitive Farming Officer	SP	Secondary Physical (question)
DCE	Discrete Choice Experiment	SPA	Special Protected Area
Defra	Department of Environment, Food and Rural Affairs	SPICE	Setting, Population, Intervention, Comparator, Evaluation
DTC	Demonstration Test Catchments	SRP	Soluble Reactive Phosphorus
DWPA	Diffuse Water Pollution from Agriculture	SS	Secondary Social (question)
EA	Environment Agency	SSed	Suspended sediment
E	Excellent	SSSI	Site of Special Scientific Interest
G	Good	TN	Total Nitrogen
HLS	Higher Level Scheme	TP	Total Phosphorus
ID	Identity	WFD	Water Framework Directive

The prefixes for the document identifiers were selected as follows. A refers to Academic publications; AC to documents covering all catchments; O to documents from Defra ScienceSearch, AV to documents regarding the Avon, E to documents regarding the Eden and W for documents regarding the Wensum catchments. However, these prefixes are not always reliable – for example there are some with single catchment prefixes that cover multiple catchments.

1 Introduction

1.1 Diffuse pollution from agriculture

Many surface, coastal and ground waters in England are significantly polluted with 78% of freshwater bodies failing to meet 'good' ecological status as prescribed in the European Water Framework Directive (WFD) (POSTnote, 2014). Pollutants impact water bodies from point (e.g. sewage treatment work discharges) and diffuse sources (arising from contributions of numerous small sources distributed across the catchments). Diffuse pollutants include industrial chemicals, nutrients, microbes, pesticides, herbicides and sediments that originate from a variety of sources (e.g. urban, industrial and agricultural area) with varying relative importance. Given that agriculture accounts for ~70% of the land area of England, and often higher proportions in small headwater catchments, its potential in contributing to diffuse water pollution is significant with estimates suggesting it typically accounts for a third (POSTnote, 2014). Therefore, diffuse pollution from agriculture is one of the major pressures responsible for failure to meet WFD objectives and a focus for most of England and Wales' river basin management plans.

1.2 The policy context

To meet the WFD water body thresholds for 'good ecological status' and thus avoid financial penalties it will be necessary for policies to reduce diffuse water pollution from agriculture where this is known to be a major contributor. This follows the principle of the implementation of Nitrate vulnerable Zones for meeting waterbody targets for nitrate. Given the focus on ecological health a holistic approach to catchment management is required whereby physical, chemical and biological considerations are taken into account.

There are major policy challenges in reducing diffuse water pollution from agriculture in England and Wales (see POSTnote, 2014). Policy levers traditionally consisted of advice, incentives and sanctions. In certain circumstances, catchment-scale Payments for Ecosystem Services initiatives can be effective. For example, SouthWest Water has successfully implemented their Upstream Thinking initiative (POSTnote, 2014). This demonstrates a move towards catchment-scale water management that is supported by OFWAT. However, these schemes have only worked in catchments where benefits have been quantifiable. Areas with specific designations (e.g. SSSI's) may be eligible for alternative sources of funding for catchment management. Agri-environment payments have also been used to help incentivise water quality improvements. In other areas where the implementation challenges inhibits the PES approach, reliance falls on our obligation to meet EU legislation and adopt the 'polluter pays' principle. However, it is very difficult to prove the source of pollution and prosecution is difficult.

Increasingly, environmental policies are seeking to use voluntary uptake of good practice by all land managers across catchments, backed up by incentives and sanctions. Defra, Environment

Agency and Rivers Trusts are now implementing policies (also taking into account social and economic considerations) to facilitate such integrated catchment management. For example, the Catchment Based Approach (CaBA) and the 25 year environment plan have empowered partnership working to mitigate pressures at the catchment scale. Similarly, in Wales the Well-being and Future Generations Act has driven integrated catchment management. However, knowledge gaps on the effectiveness and targeting of specific mitigations measures currently constrains the ability to prioritise and implement many policies. To close these gaps Defra continues to work in partnership with many organisations and has funded (for example) the Demonstration Test Catchment project to advance knowledge on the effectiveness of measures to mitigate DWPA and the Ecological Targeting project (Defra WQ0228) to develop methods to better target interventions.

1.3 The Demonstration Catchment Initiative

The Demonstration Test Catchment (DTC) project was commissioned in 2009 by Defra's Agriculture and water quality team to test the hypothesis that 'it is possible to cost-effectively reduce the impact of agricultural diffuse water pollution on ecological function while maintaining food security through the implementation of multiple on-farm mitigation measures'. The DTC Network was initially established to:

- Investigate approaches to delivering environmental improvements whilst maintaining farm productivity;
- Answer key research questions to support policy formulation;
- Create networks of collaboration and knowledge exchange; and
- Trial an integrated approach to catchment management.

More specifically, the summary report of the first two phases of the DTC project present policy-relevant questions which are being asked of the DTC research community. These are listed in Appendix B.

The programme was undertaken in three phases: Phase 1 (Dec 2009 to Jan 2015) focused on understanding the issues, designing interventions, understanding how to influence farmers, monitoring water quality outcomes and the role of DTC scientific evidence in supporting catchment management. PIs noted that it took several months to select and finalise the sub-catchments before monitoring could begin. Phase 2 (Jan 2015 to Mar 2018) focused on socio-economic aspects of catchment management and upscaling and extrapolating the work from Phase 1. Phase 3 (Apr 2018 to Mar 2019) focused on final reporting and further knowledge exchange. The PI feedback suggested that the approach evolved over time, becoming less focussed on interventions to mitigate DWPA (a response to a more limited implementation of on-farm measures) and more about understanding and characterising the catchment systems. The evolution of the project was overseen by a Research Advisory Group and Defra policy officers (PI feedback).

The project was established as a research platform in four catchments (Eden in Cumbria, Wensum in Norfolk, Avon in Hampshire and the Tamar in Devon/Cornwall) to build on existing infrastructure, datasets, knowledge and farming contacts all developed through previous initiatives. These catchments provide good national coverage and representativeness of different physical and socio-economic factors relevant to diffuse pollution (see section 10 on generalising findings for further detail). Furthermore, as the water quality in each of the catchments is compromised by diffuse pollution they provide case study landscapes where there is a need for mitigation measures (see below).

1.4 The DTC catchments: Ecological status and pressures

With the exception of the Eden, Dacre subcatchment, all experimental subcatchments are below the required Good-Moderate boundary based on their Biological Quality Elements (O11; Table 2). This shows that all catchments (with the exception of the Dacre) can benefit from mitigating agricultural diffuse pollution. From analysing the reasons for failure, pressures from fine sediment and nutrients have been identified as the most common issues (Table 2). The identification of these issues has informed the adoption of the most appropriate mitigation measures in each catchment. Indeed, where water bodies are classified as having moderate or lower status there is a legal requirement to put a programme of measures in place to restore their quality. This really highlights the important role that the DTC research plays in linking ecological response to physiochemical drivers at the catchment scale.

Table 2: WFD Classes for Demonstration Test Catchments (Source: 011).

Catchment	Sub-catchment	Site		WFD Class				Issues
				2010	2011	2012	2013	
Hampshire Avon	Sem	d/s Cool's Farm	Control A	M	P	M	P	Sediment, N, P, Organic
		Priors Farm	Manipulated	M	M	M	M	Sediment, N, P, Organic
		Donhead Hall	Control B	P	P	P	P	Sediment, N, P
	Ebble	u/s wetland	Control		P	P	P	N, P
		d/s wetland	Manipulated		P	B	P	N, P
	Wylye	Kingston Deverill	Control		M	P	M	N, P, Flow
		Brixton Deverill	Manipulated A	M	M	M	P	N, P, Flow
		Hill Deverill	Manipulated B	M	M	M	M	N, P, Flow
Tamar	Caudworthy Water	Caudworthy Bridge	Manipulated A		M	M	M	Sediment, Organic
		Caudworthy Ford	Manipulated B		M	M	M	Sediment, Organic
	Neet	Burracott Bridge	Control		P	M	M	Sediment, Organic
Wensum	Blackwater Drain	Eco-control	Control		B	B	B	Physical habitat, Flow
		A		P	P	B	B	Physical habitat, Flow, Sediment, N,P
		B		M	P	B	B	Physical habitat, Flow, Sediment, N,P
		C		P	P	B	P	Sediment, N, P
		D		M	P	P	P	N,P
		E		P	P	B	P	Physical habitat, Flow, Sediment, N,P
		F		B	B	B	B	Sediment, N, P
Eden	Morland	Newby Beck			M	M	M	Sediment, P
		Dedra Banks Beck			M	P	M	Sediment, P
		Sleagill Beck			M	M	M	Sediment, P
	Pow	Pow Outlet			M	B	M	Sediment, N, P
		Tributary A			M	M	M	Sediment, N, P
		Tributary B			M	M	M	Sediment, N, P
	Dacre	Thackthwaite Beck			G	G	G	Sediment, P
		Mellfell Beck				G	G	Sediment, P
		Lowthwaite Beck			M	G	G	Sediment, P

Note: Ecological status of Biological Quality Elements: H=High, G=Good, M=Moderate, P=Poor, B=Bad. Water Framework Directive is Good or above

1.5 Aims and objectives of review

The primary aim of this review is to evaluate the knowledge gained, on intervention effectiveness and farmer engagement, from the DTC programme by undertaking a Rapid Evidence Assessment (REA) of the available, published evidence. Secondary aims were included to evaluate the broader physical and social science methodologies that were adopted and the approach for generalising the findings. They were included in our proposed review methodology to meet aspects of the tender specification and to satisfy requirements of the steering committee. These secondary aims allowed aspects of the DTC to be evaluated that would have otherwise been excluded due to the strict inclusion criteria for documents under the protocol of the REA. Despite this adaptation to broaden the REA, PI feedback emphasises that the focus was much narrower than the original aims of the DTC platform, and these DTC aims evolved over time given changing policy and political priorities.

A REA is an evidence review that provides "an informed conclusion on the volume and characteristics of an evidence base, a synthesis of what that evidence indicates and a critical appraisal of that evidence" (Collins et al., 2015). A REA is seen to be appropriate for policy and practice as they are more systematic and rigorous than a literature review, but more cost-effective and less time-consuming than a full systematic review. The guidance provided by Collins et al (2015) and published by Defra and NERC on how to structure and implement evidence reviews to support policy making and evaluations is followed here. The recommended protocol has been tailored to ensure it fits with the purpose, timetable and required steps outlined in the invitation to tender for this project by Defra.

The specific objectives of this REA were to: assess the volume and characteristics of the DTC evidence base; synthesise what the DTC evidence base indicates; and to offer a critical appraisal of the DTC evidence base.

The objectives of the DTC and therefore this REA were more complex and wide-ranging than many evidence reviews. In this case we assessed the social science around engagement and uptake of proposed interventions; as well as the physical science of whether a suite of interventions has led to changes in water quality. Many reviews (e.g. Randall et al., 2015) tend to look at one aspect of these policy questions, not the full suite.

1.6 Identifying end-users and the Steering Group

The end-users of this evidence review are Defra and its associated agencies with an interest in mitigating diffuse pollution from agriculture and transferring the lessons learnt from the DTC to current thinking on catchment management in England and Wales. A broader set of stakeholders were included in dissemination stages. The project Steering Group consisted of the representatives from Defra, Forest Research, Natural England and the Environment Agency.

2 REA protocol development

The protocol development represented a crucial step in the REA process with multiple components. Therefore, consideration is given to defining the conceptual model, primary and secondary questions, and scope of the work; and their confirmation with the Steering Group.

2.1 Conceptual Model

It was important to have a conceptual model that elaborated the subsequent objectives for the REA. A schematic representation of the main issues to look for in the evidence base and how the aspects link together to generate the desired outcomes ensures the questions used in the assessment are relevant and focussed. The conceptual model for this REA is shown in Figure 1. This does not represent the overall objectives or focus on the DTC programme, but was agreed with the REA steering group to guide our narrower focus on mitigation of DWPA through on-farm measures.

The DTC interventions and adopted methodologies used for assessment are given in the orange ellipse. Orange arrows illustrate the descriptors of the DTC project that affect the catchment land use and water quality; both represented within the green square. The factors that determine the nature of the descriptors are given in the three light blue boxes that overlie the orange arrows. External confounding factors are given in the grey rectangular boxes. Primary and secondary questions are mapped onto this conceptual model. SP stands for secondary physical science question and SS stands for secondary social science question. These questions are presented in section 2.2. We have adapted the original social science conceptual model of policy interventions and influences on behaviour which influenced the development of the DTC Phase 2 report (see Figure 3) to make the model more consistent with the physical science part of the REA and ensure we have a coherent approach to the evidence review.

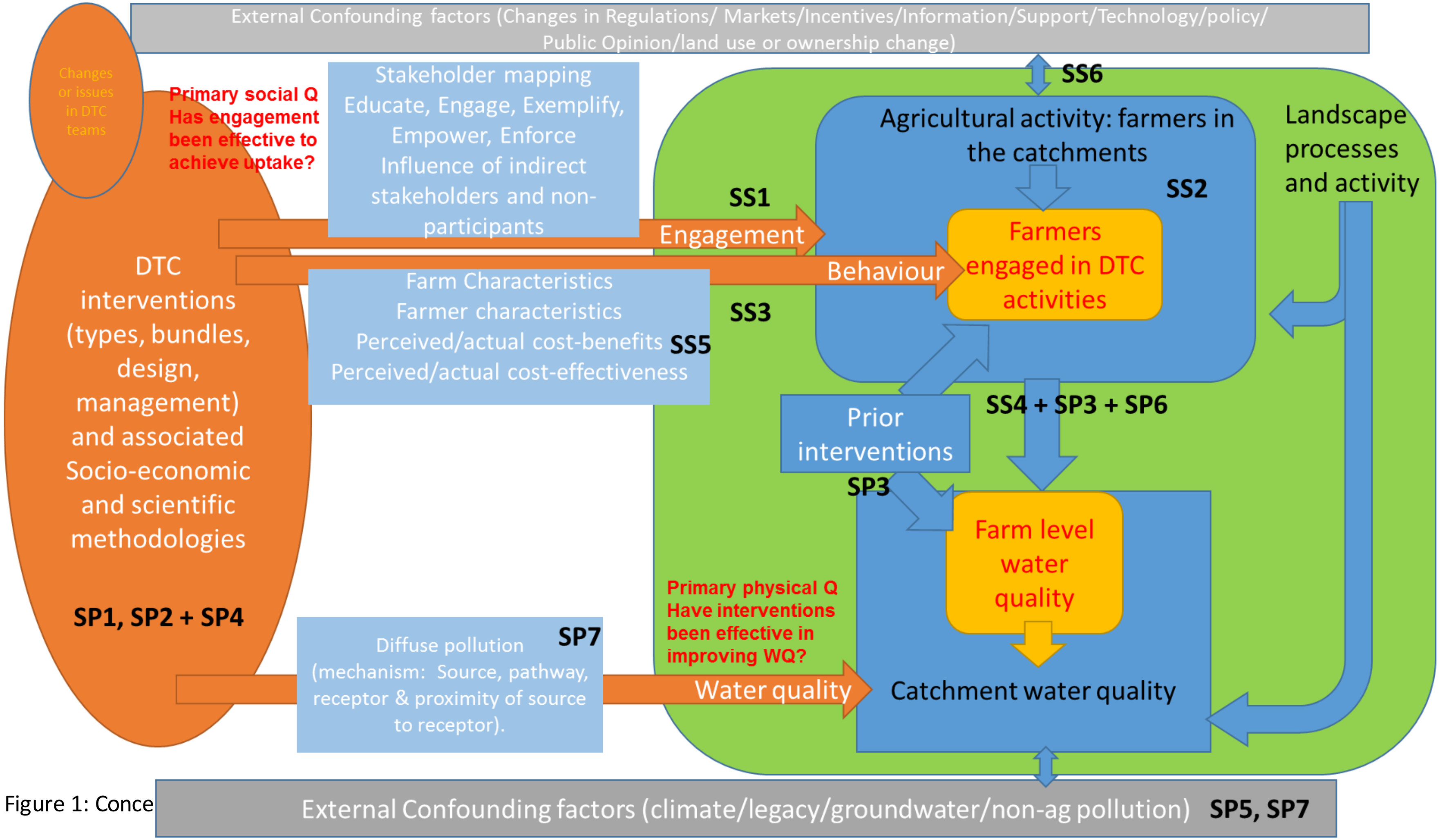

Figure 1: Conce

2.2 Primary and Secondary Questions

The primary questions represented an essential part of the REA as they guided what evidence was extracted and how it was assessed, as well as providing the narrative focus for the synthesis. Two primary questions were chosen, reflecting the physical and social science aspects of the REA. We opted to make the primary questions 'impact' questions, as we wanted to answer what impact or effect an intervention has had as recommended by Collins et al. (2015).

The Physical science question complied with the PICO (Population, Intervention, Control, Outcome) framework as set out in Table 3.

Table 3: PICO elements defined for the DTC Rapid Evidence Assessment

PICO Element	**Included**	**Excluded**
Population	DTC catchments (at various scales)	Studies where results are not explicitly DTC (e.g. national studies). Exclude methods developed in DTCs but applied elsewhere.
Intervention	Agri-environment measures. Measures that reduce water pollution from agriculture.	Studies where agri-environment measures are not included
Comparator	before and after intervention assessment, sub-catchments with and without interventions (control)	Studies that do not include water quality or biology. Studies that do not make a comparison to identify an effect.
Outcome*	Change in water quality (chemical or sediment concentration/load/yield or biology)	Studies where an explicit change in water quality or biology is not included. Exclude studies of sediments (e.g. source apportionment) or pollutants from land where no information on resulting water quality impact is given.

*Although some evidence sources were not PICO compliant they were labelled as 'PICO relevant'. Such sources did not quantify impacts on specific water quality parameters but enabled strong water quality inferences. For example, a reduction in soil erosion may have been measured and a strong water quality inference could be that stream sediment loads had reduced. In some cases this was a difficult judgement owing to the limited information available to the reviewer during their rapid assessment. The inclusion of 'PICO relevant' sources in this review is considered in Section 3.5.

The PICO framework was not considered suitable for social science so, following good practice, the social science primary question followed the SPICE (Setting, Perspective, Intervention, Comparison, Evaluation) framework, as shown in Table 4.

Table 4: SPICE for DTC Rapid Evidence Assessment

SPICE Element	**Included**	**Excluded**
Setting	The four DTC catchments	Studies where results are not explicitly DTC (e.g. national studies). Exclude approaches developed in DTCs but applied elsewhere.
Perspective	Land based stakeholders	Studies that do not include empirical data from DTC stakeholders (purely academic perspective)
Intervention	Uptake of agri-environment measures, supported by engagement processes	Studies that do not link engagement processes to specific DTC agri-environment measures; e.g. studies about tools or methodological approaches
Comparison	The counterfactual (what might have happened in terms of agricultural practices without any DTC activities)	Material that is not relevant to the farming systems of the four DTC catchments
Evaluation	Extent and quality of research on farmer engagement that enabled farmer uptake of DTC agri-environment measures; extent and quality of research on uptake of DTC agri- environment measures.	Extent and quality of general farmer engagement practices; extent and quality of general farmer behavioural data

Some evidence sources were not SPICE compliant and were not used in addressing the primary aim of the review (see Section 1.5) as our inclusion criteria were very strict. This meant many relevant pieces of work were excluded. For example, the source discussed engagement of farmers more generally rather than engaging farmers to take up specific DWPA measures in the DTC catchments. Where the sources contained useful information of relevance to the secondary aims of this study (see Section 1.5), these sources were labelled as 'SPICE relevant'. The inclusion of SPICE relevant sources is considered in Section 3.5. In some cases, judging whether something was SPICE compliant or SPICE relevant was difficult, given that it may have been possible to infer insights for the REA from borderline cases. However, the SPICE criteria were strictly applied and insights from SPICE compliant (section 5) and SPICE relevant (Sections 8 and 9) sources were included in the review.

The primary questions for the REA of the DTC initiative were:

Primary Physical: How effective were DTC agri-environment interventions in DTC catchments for improving and maintaining water quality?

Primary Social: How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?

Secondary (supplementary) questions are often used to tease out the diverse issues underpinning the primary questions and ensure there is sufficient context and detail to understand the evidence. Generally, there are only 1-2 secondary questions but here we used more, as our primary questions (and conceptual model) covered a wide range of interconnected issues. Secondary questions are generally non-impact questions and are normally open-ended.

Secondary Physical

SP1. Have DTC monitoring methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

SP2. Have DTC modelling methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

SP3. Based on DTC evidence how effective are specific agri-environment interventions and combinations of measures in mitigating diffuse pollution and improving or maintaining ecology? Are these relevant to or assessed against regulatory standards? Does effectiveness change over time?

SP4. What models were applied during the DTC programme?

SP5. Are models used to represent future scenarios (climatic conditions and land use change) outside the bounds of the DTC dataset?

SP6. What evidence is there from the DTC programme that the effectiveness of agri-environment interventions varied between DTC catchments and was this related to differences in the design and/or management of the interventions?

SP7. What evidence is there from DTC data that confounding factors (e.g. climate, non-agricultural pollution) may be important in the interpretation of the results?

Secondary Social

SS1. What evidence is there that the DTC engagement methodologies appropriately informed, consulted and actively involved farmers and other stakeholders to maximise uptake of interventions?

SS2. What evidence is there that the DTC considered behavioural factors when engaging farmers in their activities?

SS3. What evidence is there of engagement and uptake of DTC interventions varying between catchments?

SS4. What evidence is presented on the cost effectiveness and benefits of DTC interventions, during the initiative and for the five-year period beyond the end of the DTC initiative?

SS5. What evidence is provided that the confounding factors (e.g. existing non-DTC activities, influence of non-participating farmers) were accounted for when reporting on engagement and uptake of DTC interventions?

SS6. What evidence does the DTC data provide about non-participants and why they did not engage in the process?

As well as secondary questions, further criteria, derived from the relevant academic and practice literature, were used during the evidence extraction and evidence assessment phases, to appraise the quality and validity of the evidence claims.

2.3 Scope of the work

The evidence used in this REA has been provided to us and is restricted to the datasets and publications produced as part of the DTC initiative (see establish evidence base below). Therefore, we expect all evidence to meet the category requirements in Table 5 below. However, where there was uncertainty, the criteria below were used to decide if evidence was within the scope of the REA or not.

Table 5: Evidence requirements

Geographical reference	4 DTC catchments (see Figure 2)
Climatic conditions	Present day (all)
Language restrictions	N/A
Date restrictions	DTC defined 2009 to 2019
Population restrictions	DTC defined
Outcome restrictions	Farmer engagement and awareness Uptake of DTC interventions Water quality
Funder	Only work funded by the DTC initiative is included.

Figure 2: Four Demonstration Test Catchments.

This REA was designed to assess the quality of the DTC evidence base. Therefore, data and evidence not produced as part of the DTC initiative were beyond scope of the REA. However, the Steering Group made us aware of non-DTC data and published or unpublished reports funded by other organisations. PI feedback also drew attention to the fact that the PIs were relying on coordination with other projects to analyse and interpret the data collected via the DTC platform, but these connections did not always materialise. Therefore, future pathways for any external work to be brought in to contribute to the overall body of evidence are documented in the workshop report to capture this learning point.

During January 2020 the protocol was finalised taking account of feedback from the Steering Group.

3 Establishing the Evidence Base

All evidence has been provided by Defra in the form of outputs from the DTC initiative. Therefore, no additional searches were undertaken for evidence; removing the need for a protocol to agree search terms, search scope and search boundaries. PI feedback highlighted the fact that documents regarding the DTC archive project were missing, which would explain some issues regarding accessibility of monitoring data.

Supporting references are provided using standard citations. DTC material assessed as part of the rapid evidence assessment are referenced using a unique identifier, which are listed in Appendix A, to allow cross-reference to the evidence tables (see Supplementary material 1 and 2).

The outputs were uploaded into the Mendeley reference manager. An Excel master list of documents was produced and updated following consultation with Defra and the provision of missing information. Missing documents were sought through contacting authors, online searches and requests to the parallel compendium project (Defra, 2020) WT15116. There were also six documents identified via PI feedback, relating to taking a temporal perspective on monitoring signals, that were not supplied to the review. It is unclear if they would have met the criteria for PICO or SPICE review.

This resulted in a total of 173 documents making up our evidence base, of which 143 were assessed for the physical science primary and secondary questions and 77 were assessed for the social science primary and secondary questions (this adds up to more than the total, as some documents contained both physical and social science evidence).

Each document on the excel list was identified as physical, social or both, and assigned a reviewer. The unique document codes were used to reference evidence sources in this report and those cited are included as the DTC evidence reference list. Once a document had been reviewed this was flagged on the list with the reviewer's initials. The information stored in this excel document helped the project manager allocate documents for review and monitor progress. The draft evidence review was shared with DTC principal investigators (PIs); and a questionnaire was used to elicit their views on whether our analysis was factually correct, to draw attention to any gaps and to raise issues that may explain the results of the REA. None of the PI's chose to provide feedback using the questionnaire provided. One PI responded drawing on comments provided by a further four researchers, providing a six-page narrative that gave extremely useful context and background to the DTC from their perspective. The bulk of the response drew attention to where the focus of the review risked, in their opinion, obscuring important learning points. The remaining substantive comments regarding our findings were directed to SP5, SP7 and SS4. Another PI provided six pages of feedback on behalf of their consortium. This response provided useful background on the operational details of the platform, which helped explain why certain evidence was not available. As with the other PI response, the bulk of the material drew attention to research findings generated by the DTC programme that were not addressed by the PICO and

SPICE questions. As much of these narratives drew on unpublished or even undocumented knowledge, this 'evidence' would not be suitable to submit to a REA as it is qualitatively different to formal, published, documents. It has been acknowledged and documented in the synthesis process; and missing evidence acknowledged above. This process of primary research has ethical clearance from the James Hutton Institute (reference 04/2020).

3.1 Screening the Evidence Base

It was agreed with the Steering Group on 21st January 2020 that PowerPoint presentations would be excluded from the review process as they are likely to duplicate information provided in reports and papers.

In order to ensure the evidence extraction was efficient it was necessary to screen and characterise the evidence base to ensure we took account of duplication and redundancy. In terms of redundancy, multiple outputs may report the same findings from the same data but for different audiences; or an annual report may be subsumed into a final project report. A single year of data may also be included in the analysis and reporting of a multi-year dataset. Similarly, there may be duplication where single and then multiple determinants are reported separately (e.g. one report might only consider effects of interventions on N leaching and another might consider N as well as P and sediment). In order to ensure there was no double counting the most comprehensive and contemporary document was used, and the 'duplicates' screened out. In this way, the original evidence base was reduced to those appropriate for answering the primary and secondary questions.

Screening of documents was managed and documented in the evidence extraction tables (see section 3.2). Where documents are screened out basic information was recorded and the reason for rejection was noted. Evidence that was relevant but did not meet the PICO or SPICE inclusion criteria was identified in the extraction table for inclusion in the wider discussion sections 6-11 of this final report. Overall, there were 25 PICO compliant documents (reduced to 12 when documents replicating information were removed) and 15 SPICE compliant documents (although only eight of these 15 had full robustness analysis applied) that were used to generate the findings reported in sections 4 and 5.

Following good practice, screening undertaken by one researcher was cross-checked by another member of the project team who screened a subset of the documents. Where there was a difference of opinion, this was discussed with the PI for the physical or social research and a final decision made. All team members were familiar with physical and social questions and they were briefed to pass on relevant documents to the other team where necessary. PI consultation was used to capture any evidence that may have been missed by either the social or physical researchers.

3.2 Evidence Extraction

Relevant information was extracted from each evidence source and recorded in the Excel extraction tables (Supplementary material 1 and 2).

The first columns of the Excel table gives document information and details regarding whether the evidence meets inclusion criteria and whether it is duplicated evidence. Criteria for evidence extraction are given along the rest of the column headers. The criteria allowed us to extract details of the evidence within each document and record this systematically.

Where appropriate paired columns were used for each criterion – one with pre-set categories to allow us to select a category for each document. This categorical data can be easily summarised and used in critical assessment. A standard category for use with all criterion was 'insufficient information to judge'. The other paired column was used for notes to capture information that did not fit within the categories (e.g. if there was some uncertainty about which category or to capture additional relevant information). This allowed for more qualitative judgements that can help ensure transparency and enrich the narrative commentary in the final synthesis. It was also used to capture information about when an issue was not addressed and why.

The criterion used for data extraction enabled the physical and social science questions to be addressed. The criteria were triangulated with the conceptual model and questions (see section 2.1 above) to ensure that they provided relevant evidence. Three contrasting physical and social science studies were used to test and refine the extraction criteria. As with the screening above, individual researchers were allocated a set of documents to review; and met to compare entries for a sub-set (~10 physical) to ensure consistency in the use of categories and capturing of notes. Further refinements were made to extraction criteria at this stage.

General, physical and social science criteria

For details of extraction criteria please see the associated Excel templates (Supplementary material 1 and 2). Fifteen general criteria are included that give detail on the document type and its relevance. Ten criteria are included to capture information on the geographical context including location and catchment characteristics. For both the physical and social evidence ~50 additional criteria are used to extract evidence. For the physical science these include information on the interventions, research design, and conclusions. For the social science these include information on farm and intervention type, farmer and stakeholder engagement, criteria on non-participation, behavioural considerations, cost effectiveness and benefits, and confounding factors as well as research design. Following tests by the project team, to maximise consistency in the process of data extraction, predefined classifications of information were used where appropriate and additional columns included to capture detail.

To enable a critical assessment of the evidence (Sections 4 and 5) it was important to ensure that the extracted information allowed a number of robustness questions to be answered and scored. Thus, prior to evidence extraction, social and physical robustness questions were identified and checked to see whether they could be answered using the extracted information. This led to several extraction criteria being modified.

3.3 Critical appraisal of Evidence

A key part of a REA is the critical appraisal of the evidence in terms of its relevance and robustness. Thus, the REA does not only answer the primary and secondary questions, but it also reflects on the confidence we might put in these findings. Each piece of evidence is evaluated for its relevance to the question and the robustness of the methods used. Feedback from the PIs noted significant interaction with policy teams in Defra to ensure the DTC was aligned with emerging policy needs in order to maintain relevance. In general, robustness means both accuracy and attempts to minimise bias (Collins et al., 2015).

Given that a sister project (Defra, 2020)[1] was summarising and consolidating the main findings from the DTC initiative, this step became even more important to ensure that there was value added and to illustrate where there may be gaps or contested knowledge.

The REA should tailor this appraisal to the REA questions, in particular to ensure the evaluation was appropriate in light of the method used, the target population, interventions and outcomes measured.

Criteria were identified to assess the accuracy and robustness of each type of evidence. In the extraction table each document and line of evidence was assigned a category under several headings (e.g. study design, publication type) to enable the robustness assessment. As we reviewed a sample of the documents, these categories were refined and adapted.

Collins et al (2015) state that all evaluations should consider whether:

- Specific questions and hypotheses are addressed
- Related existing research or theories are acknowledged
- Sources of funding and vested interests are declared
- The methodology used is clearly and transparently presented
- The degree to which the method reduces bias
- The method is appropriate for the research question and the conclusions reached by the study
- Assumptions are specified
- The geography and context of the study is clear, with a discussion of how relevant findings are to other contexts
- The methods used for measurements and analytical techniques are reliable
- Measurements and analytical techniques have been validated and verified
- Conclusions are backed up by well presented data and findings
- Links between descriptions of existing research, data, analysis and conclusions are clear and logical

---

[1] This Evidence Compendium (WT15116)) covered: summary of the DTC platforms, pollutants, monitoring, catchments interventions; social science and a list of resources used.

- Limitations and quality have been discussed.

These were adapted for our study design as outlined below:

Criteria for Relevance and Robustness of Physical Science Evidence

Relevance: Each evidence source was assessed for its relevance in relation to the primary physical question (Y/N) using the PICO criteria (see Table 2).

Robustness: The robustness of physical evidence was assessed by considering 23 questions focused on the following four areas 1) Reporting; 2) Method (Modelling); 3) Method (Monitoring); and 4) Results. Overall scores were given for each area by averaging component sub scores on a scale of increasing strength from 1 to 3); produced from answering the specific questions. However, the individual sub scores are in many cases extremely important and were preserved for use in discussion. In our evaluation, whilst an individual sub-score of 1 was typically assigned to indicate a satisfactory level of robustness, a study with an average close to 1 raised some provisional concern over its overall robustness. Extraction criteria were carefully chosen to ensure that all robustness questions could be answered using evidence from the extraction table without revisiting documents. The robustness questions included in each of the four areas (also included in the evidence extraction table) along with their scores are given in Appendix C.

Criteria for Relevance and Robustness of Social Science Evidence

Relevance: Each evidence source was assessed for its relevance in relation to the primary social question (Y/N) using the SPICE criteria (see Table 3).

Robustness: The robustness of social evidence was assessed by considering questions focused on the following three areas 1) Data collection; 2) Data analysis; and 3) Reporting, resulting in a total of 22 questions, although only criteria suitable for the type of research (qualitative, quantitative or economic) were applied. Only those documents providing information on their methodologies were assessed for robustness. This differs from the physical sections where the lowest score was given where methods were not presented. Therefore, where robustness scores were not calculated for the SPICE relevant documents, these could be inferred to be low on the basis of lack of methodological reporting (see also Appendix D and E).

Overall scores were given for each area by averaging component sub scores (1 to 3); produced from answering the specific questions. These gave overall judgements of satisfactory, good or excellent. However, the individual sub scores from each of the areas was preserved for use in discussion. The robustness questions included in each of the three areas (also included in the evidence extraction table) along with their scores are given in Appendix D and the results of the robustness analysis per document provided in Appendix G.

As well as providing the numeric score, each social science assessment is accompanied by a text summary (recorded in excel) describing the reasons for the score for transparency. This justification was needed for scoring aspects of the social science as many of the questions will have less well-defined answers relative to those being asked of the physical science.

Scoring

As with screening and extraction, another researcher checked a sample of the work to ensure consistency of scoring (where time permitted). Collins et al. (2015) advise that these scores are combined with relevance scores. However, given our REA was only focussed on material from the DTC initiative, and we screened the evidence base to only retain those of relevance to the primary questions, this step was not required.

This process provided a sense of whether the DTC findings were based on robust methodologies and a sense of whether there was more robust evidence in some areas than others. The point of the evaluation was to improve future research, policy and practice so the assessment was aimed at illustrating where future programmes might need to put more effort to fill research gaps, rather than criticising the existing knowledge base generated by the DTC. Appendix E provides more reflection on these issues.

Summarising the Evidence

Once the evidence extraction had been completed, an initial synthesis of answers to the questions was made before the critical appraisal was completed.

3.4 Synthesising information

Evidence was synthesised under each question. The extracted information provided the necessary granularity to structure answers. For example, physical science findings were structured by intervention type and/or water quality parameter. The objective was to also give a qualitative assessment of the robustness of the findings (Satisfactory (S), Good (G), and Excellent (E)).

Similarly, for the social science appraisal extracted information was also used to structure answers by questions around engagement type and uptake of different interventions. Gaps in knowledge were summarised and, where possible, a commentary was provided about why these data were not collected or sufficiently analysed. Suggestions for designing future DTC type platforms were also included. Cost effectiveness was considered by synthesising information on cost of interventions and their effectiveness. Areas with 'insufficient information to judge' were highlighted.

As this study was a REA a statistical assessment of the results was not appropriate. Findings were summarised in a narrative way. For example, we conclude that 80% of high quality evidence sources reviewed supported X and Y. The sub scores used in our evidence quality assessment were used to enrich the discussion. They helped identify the strengths and weaknesses in the research and provided useful guidance for future catchment initiatives.

3.5 Inclusion of non PICO and non SPICE documents

Many documents were screened out of the REA described above (see Section3.1) as the documents did not specifically provide the information needed to answer the PICO or SPICE questions. However, some of these documents were labelled as 'relevant' in the evidence extraction tables (Supplementary Material 1 and 2) as they contained useful insights into the relevance and robustness of the science undertaken within the three phases of the DTC project.

When addressing the primary aim of this review and evaluating the knowledge gained on intervention effectiveness for water quality, PICO compliant sources were used to answer the primary and secondary questions but where appropriate PICO relevant sources were also described (note that they were clearly separated from PICO compliant evidence using a side heading).

When addressing the secondary aims of this review and evaluating the broader physical and social science methodologies (Sections 6 to 9) and the approach for generalising the findings (Section 10), a more traditional approach was adopted to reviewing the literature provided and evidence sources were not limited to the PICO and SPICE compliant documents. These sections were also enriched by the inclusion of additional literature sources (identified by the reviewers) and preliminary analysis of primary data (e.g. Section 6.5.1).

Although PICO and SPICE relevant documents were entered into extraction tables (Supplementary Material 1 and 2) full data extraction (including relevance and robustness criteria) was not undertaken. However, for many documents researchers captured the main findings, recommendations and limitations of the studies, along with some characterising features. Additional information was kept in working files. The data were then qualitatively analysed using themes generated from the DTC objectives, DTC reviews on the topic and recently published overviews of the topic in the literature (Yanow, 2007). Therefore, the focus of these sections is wider than the primary and secondary science questions used in the REA, precisely as the PICO and SPICE 'relevant' papers did not contain the appropriate information to answer the REA questions.

3.6 Consulting, confirming and communicating findings

The DTC review dissemination workshop was planned to be held face-to-face, but was altered to a virtual workshop due to COVID-19. The results of the review were presented as an evidence compendium which synthesises findings in an easily understandable and useable format. This evidence compendium is published on Defra ScienceSearch: WT15115. The workshop enabled a discussion of the findings of the REA, the reflections on generalisability and the 'value-added'. Attendees included:

- Defra policymakers, analysts and scientists
- Catchment practitioners

- Members of the Steering Group, and
- Academics who worked on the DTC catchments

The workshop involved those with an interest in the quality of the evidence base arising from this initiative and those relevant to thinking of how best to build on the legacy of the DTC programme. The workshop collated supporting evidence sources that might help to fill any gaps or support a DTC-like initiative in another catchment. A full report on this dissemination workshop is provided in Appendix H.

4 Rapid Evidence Assessment: Intervention Effectiveness

In this section responses to our primary and secondary questions are presented. The results used are limited to those generated by the systematic part of the review and thus the evidence presented is limited to those sources that meet our PICO criteria. However, where appropriate additional evidence is included in discussion.

Of the 177 evidence sources provided 12 passed physical science PICO screening, duplicate removal and the removal of specific document types (e.g. PowerPoint presentations). Please see associated table of evidence sources (Supplementary material 1). Of these sources 5 were from the Wensum, 4 were from the Avon and 3 were from the Eden. Four of the studies were model based, 7 were based on monitoring while one represented a combination. Other sections will report on other aspects of the review, including: a review of the transferability of the findings; a review of the monitoring and modelling methods adopted; and a review of the monitoring infrastructure and data collected. These sections are not constrained by the PICO criteria so they will draw on all outputs of the programme.

In this section the primary review question is answered first by combining all evidence in summary form. Distinctions are not made between intervention types, catchments, scales, study types etc. Secondary questions are then used to provide context and clarity on the evidence that underpins the answer to the primary question. These questions provide important information on whether evidence for intervention effectiveness depends on factors such as intervention type, water quality parameter, study method, DTC catchment, and scale.

4.1 Primary Physical: How effective were DTC agri-environment interventions in DTC catchments for improving and maintaining water quality?

The evidence used to answer this question is presented in Appendix F (Table F1).

Answer: The DTC evidence base supports the conclusion that DTC interventions can improve water quality (reported reductions of >80% sediment, >90% for specific pesticides, and 75% for nitrate and 50% for phosphorus) but effectiveness depends on scale and intervention type.

A total number of 12 evidence sources met our PICO criteria and thus provide evidence on whether DTC interventions improve or maintain water quality in DTC catchments (Table F1). Within these sources, evidence was provided for the effectiveness of 24 intervention cases (10 from the Wensum, 9 from the Eden and 5 from the Avon). In 22 (~90%) of the cases a water quality benefit was reported in at least one determinant. However, it important to note that in five of these cases an adverse impact was reported in at least one other determinant. The reported water quality effects included sediments, pesticides, and nitrogen and phosphorus species. It should be noted that no PICO compliant studies reported ecological effects. One pesticide intervention case showed a reduction in concentration. Of 17 intervention cases that reported on nitrogen 13 reported a decrease (~2->75%), three an increase (4.7 - 15%) and two no impact; in some instances reported impacts vary with scale monitored. Of the 22 intervention cases that reported on phosphorus 17 reported a reduction (~2-50%), two an increase (~4-7%) and one no change. Of the 13 intervention cases that reported on sediment 11 reported a decrease (2-82%) and two an increase (8 and 42%). However, a fuller understanding of these changes should take account of the species monitored and the metrics reported (i.e. concentrations or loads); included in Table F1. It is also important to note that in many of the cases where low effects were reported (e.g. A15, A63) interventions related to <20% coverage of the catchment. Where studies cover entire catchment areas or plots (e.g. W4, A67) stronger effects were reported. For sediments, small scale interventions may have a large effect as they treat runoff from a larger contributing area (e.g. roadside wetland in A11 and ponds in A5).

Of the intervention cases reported 10 were modelled and 14 were monitored. Therefore, evidence supporting water quality benefits of interventions is based on both modelled and monitored studies in relatively equal numbers. It is interesting to note that two of the modelling cases reported an increase in N and P (A62). Although outweighed by evidence showing an improvement in water quality, the evidence provided for these two cases was scored as robust so it should not be discounted.

Relevant non PICO study: It is useful to note that a novel study (AC44) used magnetic tracing methods to measure the sediment trapping efficiency of buffers in each of the DTC catchments. Ten buffer strips were monitored (4 in the Avon; 3 in the Wensum; and 3 in the Eden) and all were found to trap sediment. A mean trapping efficiency of 39% was reported.

Robustness of the reported results.

Nine (75%) of the 12 PICO compliant evidence sources included in this evaluation are robust with respect to their reporting given they are peer reviewed articles. The robustness of the results (Satisfactory (1), Good (2), or Excellent (3)) takes account of the detail presented to enable an evaluation of the strength of the findings and their transferability. Of the 24 intervention cases presented 14 (58%) had Good results with rest having satisfactory results. When this is considered alongside the robustness of publication it is clear that 8 of the 10 with satisfactory results have a low publication robustness. This reflects the fact that they related to less formal publications that may include items such as posters and abstracts. It is not always appropriate to document all details of the data gained in these types of publications so these satisfactory scores should not be used to discount the work but to flag that further investigation is needed to obtain the details. However, the robustness of the results of many studies were scored as low owing to the fact that very few were assessed against standards and very few present effectiveness for individual years.

4.2 SP1: Have DTC monitoring methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

The evidence used to answer this question is presented in Appendix F (Table F2).

Answer: Robust DTC monitoring methodologies have allowed the effectiveness of a small range of interventions to be assessed at a range of scales. The majority of studies lacked adequate baseline and post treatment records.

Intervention effectiveness for water quality was only given for a small subset of potential measures with some cases reporting effects of unique combinations of measures (e.g. A21). Although novel observational science was included in the programme for other interventions (e.g. buffer strips in AC44) these results could not be included here as they did not explicitly measure and report water quality changes. Other intervention types were included in model based studies (see SP 2).

Of the 8 PICO compliant monitoring studies (Table F2) in all but one case overall methodologies were scored as 'Good'. It should be noted that the one 'Satisfactory' study had a low publication robustness as it was an abstract (with no full paper available) in which you would not expect to find all the details of the methodology. It is important to note here that monitoring methodologies were specified in contracts to project partners. Although these overall scores can be used to indicate the robustness of the methodologies adopted we acknowledge that it is important to consider the component scores for a fuller understanding. In all cases the experimental design was good or excellent with 'BACI designs', 'before and after' or 'upstream and downstream' experiments being adopted. Three of the studies adopted the most robust BACI experimental design. Furthermore, the location of monitoring with respect to interventions reported in the studies was appropriate in all cases. However, we acknowledge that the ability to detect the effectiveness of measures depended on the extent to which they were implemented over the catchments. Where measures were applied to entire contributing areas effects could be more easily monitored and detected (Table F2) but where area extent was low (<20%) reported effects were lower and more difficult to detect. In many cases in Table F2 the upstream extent of measures was not reported as they related to linear features or a combination of measure types. It is also important to note that the areal extent of some measures may be small but the effect can be large where they treat runoff from a large contributing area (e.g. ponds in A5 and wetlands in A11). Our main concern is that only 3 of the 8 studies reported baseline and post intervention monitoring periods of at least one year. Of these only two studies had 2 years or more of baseline and post intervention monitoring. Given that at least 2 years of monitoring before and after intervention is recommended this will constrain the conclusions of the studies. PI's acknowledged this lack of suitable baseline and post intervention monitoring and in their responses they described how this was discussed in programme meetings but funding constraints did not allow for a sufficiently long programme.

Hydrological measurements were explicitly considered in some but not all studies. This may have added additional insights into the conclusions drawn. Automated sampling and storm event sampling was undertaken in some but not all of the studies. In some cases spot sampling is appropriate but in other cases it may have helped to include automated measurements and to focus on event sampling. However, this consideration must be evaluated alongside the specific question being addressed by the specific studies. We also confirmed our expectation that robust laboratory analytical procedures were used and reported in most studies. Where these were not reported it was due to the nature of the publication.

The PICO compliant evidence relates to studies from plot (part of field), field and sub catchment scale (~1 – 20 km2). Only one study was identified where effects of a specific intervention was reported for plot, field and sub catchment scale. Whereas the effects of several interventions are considered at the subcatchment scale there remains a need to report on the effect of specific interventions at multiple scales. PI's explained how reduced capital funding for measures meant that fewer could be assessed and focus necessarily shifted away from landscape scale assessments.

4.3 SP2: Have DTC modelling methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

The evidence used to answer this question is presented in Appendix F (Table F3).

Answer: The DTC modelling methodologies have assessed a small number of interventions at larger sub catchment and catchment scales (2-1700km^2) but not at plot or field scales.

Intervention effectiveness (Table F1) has only been provided using modelling methodologies for a small number of intervention types at sub catchment or catchment scales (Table F3). PI responses illustrated that in some cases models were specified in Defra contracts and the focus was to model at larger scales (not at the plot or field scale).

Of the 4 PICO compliant studies in 3 cases (A2, A62, and A67) the overall methodologies are scored as 'Good'. Although these overall scores can be used to indicate the robustness of the methodologies adopted we acknowledge that it is important to consider the component scores for a fuller understanding.

Two of the studies (A62 and A67) have undertaken validation whereas the other two have not (A2 and A69). Although in each case peer-reviewed and well-founded models were applied it is clear that there was insufficient data for robust validation. This relates to the shortcoming identified under SP1 with regard to short periods of pre- and/or post- intervention monitoring. Nevertheless, in two cases (A62 and A67), models were run for considerably longer periods under scenario conditions to allow for climatic variability.

Only 1 of the 4 studies (A62) explicitly evaluated uncertainty. Many model applications, especially those using more complex models include uncertainty analysis as part of the calibration and testing process. Undertaking uncertainty analysis is therefore a precursor to assessing the significance of interventions. This process may or may not have been carried out in the case s from the DTC.

4.4 SP3: Based on DTC evidence how effective are specific agri-environment interventions and combinations of measures in mitigating diffuse pollution and improving or maintaining ecology? Are these relevant to or assessed against regulatory standards? Does effectiveness change over time?

The evidence used to answer this question is presented in Appendix F (Table F4).

Answer: Evidence is available from the DTC catchments on the effectiveness of specific interventions on improving or maintaining water quality (but not ecology) but these are only related to quality standards in a couple of cases and information is not reported on the effectiveness of individual years post intervention.

4.4.1 Effectiveness of specific and combinations of interventions

The effectiveness of specific interventions (Table F4) must be considered with respect to their areal extent of implementation (Table F1).

Three studies evaluated the effect of cover cropping on nutrient concentrations (A15, A27 and A62). Concentrations of N were found to reduce by 19.6% in A62 and >75% in A15 (where implementation was 100%). Reductions in P were only found in A62 where a reduction of 1.6% was found. Robustness of these results and their publication was Good and Excellent, respectively. However, no changes in dissolved N2O was observed in response to cover cropping (A27).

Three studies (A15, A27 and A62) investigated cultivation practices (no tillage and conservation tillage). Study A15 and A27 found no impact on N and P concentrations, whereas A62 reported N reductions of 4.7 to 6.3 % and P reductions 3.8 to 7.2%. It is important to note that studies A15 and A27 were based on monitored data with 100% intervention coverage whereas the reported reductions in A62 are model based. Robustness of these results and their publication was Good and Excellent, respectively.

Two studies investigated measures to manage fertiliser application. Study A69 modelled the effects of a maximum reduction scenario (adopting all measures in DEFRA user guidance) and reported a decrease in N (22-34%), P (53%) and SS (66%). Study A67 modelled the effect of reducing fertiliser application by 30% and found P reductions of ~40%. In both cases the robustness of results was Satisfactory whilst publication was Excellent. The satisfactory results scores related in part to the lack of details given on the interventions and the significance of the results.

Two studies (AC21 and A62) reported on combinations of measures. Measures considered in AC21 included track management, ponds, fencing and roofing whilst those considered in A62 included cover cropping, tillage and drainage. In AC21 although N and P reduced (by 5-26% and 34-50%, respectively) sediment was found to increase by 8-42%. The increase in sediment may have been caused by the disturbance created by the implementation of the measures. In A62 N was reported to decreased by 24% and P by 18%. Results and publication robustness for AC21 were Satisfactory but for A62 they were Good and Excellent, respectively. The robustness of results in AC21 were only satisfactory owing to a lack of detail on the interventions and significance of the results.

Two studies (A63 and A2) investigated the water quality benefits of livestock management. Both studies reported a reduction in suspended sediment concentration (2.3 and 6.5%) and one of the studies (A2) reported a reduction in P (4.7%). Both studies were classed as having results and publication robustness of Good and Excellent, respectively.

Two studies (A62 and A5) investigated field drainage measures (including runoff interception features and changes to field underdrainage). Both studies reported reductions in N and P. Study A62 reported reductions in P of ~32% and N of ~60%. In addition to reductions in N and P study A5 reported reductions in sediment.. Results and publication robustness for A5 were Satisfactory but for A62 they were Good and Excellent, respectively. Although intervention details are given in A5 the results were still classed as satisfactory.

Two studies investigated wetlands and ponds but only one (A11) presented quantitative data on the effect of a roadside wetland (~2.7% of the catchment area). Sediment was reported to reduce by 14% (concentration) and >80% (load). There was no reduction in P, and N was reported to increase by 15%. Robustness of results was Good and publication Excellent.

Two studies (A2 and A62) considered buffer strips and field corner management. In both cases P reductions were reported (2-17%). Study A62 reported N reductions of ~2-5% whilst A2 reported a reduction in sediment of 3%. Both studies were classed as having results and publication robustness of Good and Excellent, respectively.

PI's highlighted the fact that budgetary constraints limited the density and variety of on-farm mitigation measures.

Relevant non PICO study: It is useful to note that a novel study (AC44) used magnetic tracing methods to measure the sediment trapping efficiency of buffers in each of the DTC catchments. A high mean trapping efficiency of 39% was reported; providing further support for the effectiveness of buffer strips.

4.4.2 Are results relevant to or assessed against regulatory standards?

Only three of the 12 evidence sources (A11, A15 and A67) related nitrate, phosphorus and suspended sediment to regulatory standards. When evaluating the effect of a wetland in mitigating sediment pollution, although study A11 reported reductions in river sediment loads and concentrations, overall they say that there was no improvement in meeting the WFD (25mg/l) standard during the first 16 months of operation. Study A15 assessed the effects of cover crops and tillage practices on nitrates in soil, drain and river water and related these levels to the EU Drinking Water Directive (98/83/EC) standard of 11.3 mg N L-1. Cover crops reduced soil water nitrate to below the standard whilst concentrations under fallow remained above. Non inversion tillage was not found to reduce nitrate losses in soil water. Both cover cropping and tillage practices were not found to reduce river water nitrate concentrations and the standard was exceeded for 4.5% of the time between September 2012 and August 2015. In the modelling study A67, a combination of agricultural phosphorus reductions together with improved treatment at Waste Water Treatment Plants would reduce SRP levels in rivers to meet EU WFD requirements.

WFD classifications of each DTC subcatchment from 2010 to 2013 are presented in Table 1 (O11). The classifications do not show an improvement in status from 2010 to 2013 for most sites. The Neet (Tamar) and Dacre (Eden) subcatchments are exceptions where water quality appears to have improved. The authors of report O11 suggest that this may have been due to either large changes in water quality being needed to produce ecological responses or other factors such as physical habitat constraining ecology. Furthermore, to better understand ecological response to interventions the authors of E14 emphasised the importance of studying the impact of short term events (e.g storm Desmond) and long term annual variability (A60). PI's also highlighted additional publications (not included in this evidence review) that demonstrate 'legacy' nutrient behaviour and the long period of time that may be needed to see an effect. Thus, the lack DTC evidence of ecological effects should not be taken as confirmation of no effect but as representing the complex and longer term response that may occur.

4.4.3 Does effectiveness change over time?

In all studies changes in effectiveness over time were not reported as results were not given for individual years. With respect to the monitoring studies this probably related to the limited length of record available post intervention. PI feedback illustrated how the focus of the project evolved. Through consultation with the Research Advisory Group, and in response to budgetary constraints limiting on-farm measure implementation, a greater focus was placed on the need for a longer baseline and thus assessing the change in effectiveness over time became more difficult.

With respect to the modelling studies, unless they are highly sophisticated, it is unusual for models to be sensitive to the build-up of the effect of interventions. Therefore typically results are given as a comparison between two "steady state" conditions pre- and post- establishment of the intervention. Often models are applied for long periods to capture the effects of climatic variability. This was the case for the SWAT and INCA applications meeting the PICO criteria. Model applications in PICO compliant DTC studies were unable to discriminate effects for individual sequential years.

4.5 SP4: What models were applied during the DTC programme?

List of models in PICO-compliant papers:

- FARMSCOPER (A69)
- SWAT (A62)
- CRAFT (A2)
- INCA-P (A67)

SWAT and INCA-P are process-based models which, in particular in the case of SWAT, have a long-standing track record of application throughout the world. In principle they are readily applicable elsewhere but require catchment-specific processes of calibration and validation. However, many models of this type (e.g. SWAT) require estimation of numerous parameters that are not readily measurable and this can potentially have an adverse effect on predictive uncertainty which hampers their utility. In contrast CRAFT is a data driven model. It has powerful potential to be applicable in specific catchments nationwide. This has already been demonstrated to some extent. However, it is reliant on detailed monitoring data for calibration. Of the models applied, the FARMSCOPER model is the most readily transferable and applicable to other contexts in the UK. This can be achieved without calibration/validation of the model itself, although in so doing reliance is made on the skill and versatility of other more-detailed models that underpinned its development. The consequence of this is that a detrimental effect on accuracy is likely to be inevitable relative to that achievable by process-based models whose applications are based on catchment-specific calibration and validation.

4.6 SP5: Are models used to represent future scenarios (climatic conditions and landuse change) outside the bounds of the DTC dataset?

None of the four PICO-compliant modelling studies included consideration of future climate or land use change as part of the specified/reported scenarios. However, PICO-relevant studies were identified that take account of climate change scenarios and these are included in Section 7. PI responses also indicated that scenarios were modelled and presented to local stakeholders at local CaBA partnership meetings.

4.7 SP6: What evidence is there from the DTC programme that the effectiveness of agri-environment interventions varied between DTC catchments and was this related to differences in the design and/or management of the interventions?

The evidence used to answer this question is presented in Appendix F (Table F5).

Answer: The evidence from the DTC catchments does not allow a robust assessment of whether the effectiveness of interventions varies across the catchments and whether it is related to their design and/or maintenance.

In total the effectiveness of 24 intervention cases was reported (10 from Wensum, 9 from Eden and 5 from the Avon). Given that the interventions in each of the catchments are diverse it is not possible to compare effect across the catchments (Table F5). To enable this assessment we would need evidence on the effectiveness of a specific intervention type/design from several studies in each catchment.

As an exception the case of buffer strip implementation may be investigated to look at possible differences between catchments; although only two, modelled-based studies are available (A2 and A62). Specifically a comparison is possible for total phosphorus losses in the Eden and the Wensum. In both catchments positive impacts of buffer strips were simulated. In the Eden a 2% decrease in total loss was reported whereas in the Wensum a greater decrease (12.2% or 16.9% for 2m or 6m buffers respectively) was simulated.

Relevant non PICO study: A novel study (AC44) used magnetic tracing methods to measure the sediment trapping efficiency of buffers in each of the DTC catchments. Results were presented for the median trapping efficiencies of 6m wide buffers in the Eden (74%), Wensum (12%) and Avon (16%) catchments. Thus, trapping efficiencies varied across the DTC's and these estimates were used with the Catchment Matcher Tool to upscale results to other parts of England.

PI feedback emphasised the difficulty in comparing performance across sites when case studies represent co-working with framers and the outcome of many diverse decisions, as opposed to a reductionist experimental design.

4.8 SP7: What evidence is there from DTC data that confounding factors (e.g. climate, non agricultural pollution) may be important in the interpretation of the results?

The evidence used to answer this question is presented in Appendix F (Table F6).

Answer: The impacts of confounding factors are minimised through the adopted experimental designs. Although confounding factors are often considered conceptually, many of the DTC studies do not quantitatively account for them.

Confounding factors are known to be important in all catchment experiments. The experimental design adopted by the DTC project was developed to mitigate these as much as possible. BACI1 and BACI2 experimental designs take account of climatic variations and pre intervention monitoring periods provide baseline conditions to allow impacts of mitigation to be robustly assessed. Small headwater subcatchments are also chosen to avoid confounding factors such as non-agricultural diffuse pollution sources. By focusing on small contributing areas the extent and impact of interventions should be relatively large in comparison to effects from other activities in the catchment.

Although confounding factors are minimised through the adopted experimental design they must always be considered in interpreting results. This is discussed further in section 6 where we evaluate the monitoring network and the data that was collected. Load apportionment modelling can be undertaken to ensure no unexpected point sources are present in each catchment. Although confounding factors are often considered conceptually many of the DTC studies do not quantitatively account for them (Table F6). We acknowledge the difficulty in quantifying confounding factors. However, it is important to note that the effects of several confounding factors such as climatic variability (e.g. storm Desmond E1;E14) and the disturbance relating to construction of interventions may have been reduced if longer periods of pre and post intervention monitoring were included. The potential importance of confounding factors was identified in study A15. In this case a reduction in P was observed in the river draining the Blackwater (Wensum) subcatchment but because a reduction was not seen in the drain flow at the intervention level this reduction was attributed to another factor.

Accounting for confounding factors is a complex process when considering model applications. In each case the evaluation entails a dependency on the complexity of the model used. Process-based models implicitly take many confounding factors into account when set up to simulate entire catchments. This is relevant for the models used in the DTC and the studies in which they were applied which met the PICO criteria. For example the SWAT and INCA-P models consider non-agricultural land uses, point sources from STWs and in-stream (and groundwater) legacy effects. However, quantification of these confounding factors is less readily made in reported studies, largely because it is not straightforward to extract and isolate them in a quantifiable way. In the other hand simpler models typically consider fewer or no confounding factors in their structure. These may have been considered external to the model application itself. For example this was the case in the application of CRAFT to the Eden where the effect of interventions was assessed in the context of the characteristics of rainfall storm events.

The DTC project has provided both data and models that could potentially be used in combination to produce a novel quantification of the importance of confounding factors in agricultural environments where interventions have been established. We acknowledge that this exercise will have been beyond the scope of the DTC project.

5 Rapid Evidence Assessment: Farmer engagement and implementation

In this section responses to our primary and secondary questions are presented. The results used were limited to those generated by the systematic part of the review and thus the evidence presented was limited to those sources that met our SPICE criteria. However, where appropriate additional evidence was included in sections 8, 9 and 11. Of the 171evidence sources provided 15 passed social science SPICE screening, duplicate removal and the removal of specific document types (e.g. PowerPoint presentations) and eight were subjected to robustness analysis. Further information on the methodology applied for section 5 can be found in Appendix E.

In this section the primary review question was answered first by combining all evidence in summary form. Secondary questions were then used to provide context and clarity on the evidence that underpins the answer to the primary question. These questions provided important information on whether evidence for engagement and implementation were robust.

5.1 Primary Social Question: How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?

The DTC social science documents shared a common approach to fostering uptake of diffuse pollution mitigation measures. They firstly engaged farmers in a discussion about the issues of diffuse pollution, and their role in mitigation, then behavioural issues around willingness and ability to make changes to their farming practices. Once these discussions have taken place, a change to uptake a measure could occur. This approach was reflected in the conceptual model (see Fig 1) that guided our extraction and evaluation analysis. Therefore, to answer th e primary social science question linking engagement to uptake as agreed by the Steering Group, we needed documents that discussed farmer engagement, farmer behaviour, and then the implementation of measures that were monitored by the DTC programme.

5.1.1 Relevant Evidence Sources:

The evidence used to answer this question is presented in Appendix G (Table G1). There were two SPICE compliant documents (AC26 and O11).

It proved difficult to answer this question using the SPICE documents as very few (AC26 and O11) SPICE documents explicitly linked the specific DTC engagement processes to uptake of measures and no SPICE documents provided an explicit indication of whether these measures will be sustained in the future. AC15 reported that interviews were carried out with farmers who had extensive involvement and implemented interventions, but the results of these interviews were not easily identified in the extraction table. O11 also noted the 'protracted negotiations to ensure buy-in from farmers' (p89), including the need for planning permission for some interventions (PI Feedback). The feedback confirmed that due to changes in focus and budget constraints, very few on-farm interventions were funded within the DTC programme.

The documents where the full conceptual approach, and which mitigation measures were installed, are assessed below. O11 presented information on behaviour surveys and workshops as well as implementation of measures in 12 farms in all four catchments, and AC26 discussed working with a farmer to demonstrate farmyard management and flow attenuation in Avon, and Wensum. Note that AC21 discussed DWPA measures monitored in all four catchments but did not link these measures to stakeholder engagement or wider behavioural issues and O11 also discussed a wetland mitigation intervention on the River Ebble but there was no commentary on how the farmer(s) were engaged. The Physical science data extraction noted additional references to interventions on the Dacre and Pow sub-catchments of the Eden, covering water attenuation and rural SuDS, soil aeration, riparian fencing and woodland creation, however within documents O11 and AC21 these measures were described as being proposed for implementation during phase 2 of the project, rather than reporting on what had been implemented.

5.1.2 How robust are the evidence sources?

O11 was analysed for robustness (overall score was satisfactory). As explained in Section 3.3, only documents providing methodological information were assessed for their robustness. The table of robustness results can be found in Appendix G (Table G1).

5.1.3 Summary - How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?

. It was hard to assess to what extent engagement led to and sustained uptake, particularly without data on non-participants' uptake of DWPA measures. The REA found DTC used a range of methods to engage large numbers of stakeholders, drew attention to the importance of active involvement of farmers in the platform, and suggested the need to supplement farmer discussion groups with effective technical advice. The DTC social science documents were strong on illustrating the complexity of the relationship or series of choices farmers make between becoming aware of DWPA and finally implementing changes on their land. However, it was difficult to find evidence of the full approach (from engagement to sustained uptake) being implemented. Therefore, there was insufficient evidence to fully answer this question.

Feedback from PIs on the REA findings critiqued the narrow SPICE focus, arguing that the DTC was focussed on wider knowledge exchange practices to discuss and understand the nature of the DWPA problem and to build trusted communities of practice. The PI feedback also argued that when there are many small farmers working together in a complex landscape (including commons) it was hard to assess the link between intervention and catchment outcome. These aspects are discussed in Sections 8, 9 and 11, where we could explore these wider issues beyond the constraints of the formal REA evidence extraction process.

5.2 SS1: What evidence is there that the DTC engagement methodologies appropriately informed, consulted and actively involved farmers and other stakeholders to maximise uptake of interventions?

5.2.1 Relevant Evidence Sources:

There were 12 SPICE documents reviewed to answer this question (A35, AV11, W23, AC26, A9, E9, O13, AV8, O11, AC21, AC25, AC15). W42 and W21 did not have any data extracted for SS1. W42 was focussed on workshop messages of what should be done in the future. W21 focussed on explaining the approach to CEA using FarmScoper Tool and A11 focussed on CEA of roadside wetlands, without exploring stakeholder views.

5.2.2 Which stakeholders were engaged?

Explicit stakeholder analysis and mapping techniques were mentioned in three documents (A35, W23, AC15) and could be inferred from O13 which noted the importance of engaging farmers, scientists and advisors. The other documents did not provide information of whether and how the overall stakeholder population was assessed.

All 12 documents discussed engaging farmers. Some farmers actively engaged in implementation of DTC measures were not typical of their peers. For example, AC26 notes DTC Wensum was working with Farmers Weekly Farmer of the Year and Arable Farmer of the Year.

Six documents (AC15, AC21, AC25, O11, A9, O13) also discussed engaging a combination of stakeholders including:

- Farm advisors (AC15, AC21, A9, O11, AC25, O13),
- Catchment Sensitive Farming (CSF) Officers (O13, AC21)
- Environment Agency (EA) (O13, AC21),
- River Trust and other Non-Governmental Organisation staff (AC15, O11, A9, O13),
- Defra (AC25, O13),
- Farm Contractors (O13) and
- Utilities (AC15, O13).

O11 discussed information displays for the general public and how it was essential to work with EA, Natural England and CSF to develop solutions that work for work for farmers and 'land-owners'.

5.2.3 What is the reach of the engagement?

Whilst we tried to extract numbers of farmers informed, consulted and actively involved, we could not find accurate evidence for these numbers. The reason for the difficulty was that sometime numbers were not given for attendance at events or activities, or at other times, the numbers of farmers and other stakeholders were reported together. O11 and O13 gave the most comprehensive overview of the activities in the four catchments, suggesting an impressive range of events and engagement. For example, O11 summarised that the events in DTC catchments had been attended by over 500 farmers, 400 other stakeholder representatives and 1500 members of the general public. However, it was difficult to arrive at a single figure for each category of engagement by stakeholder type. Five documents (AC21, AV8, O13, AC26 and AC15) reported on the proportion of farmers informed, consulted or actively involved but the rest did not. One document reported the overall percentage of farmers in the catchment that were engaged in the DTC activities (AC15), although others e.g. AC26, AC21 gave total farms engaged but did not present this as a proportion of the overall farmers and land managers. Likewise, two documents (A35 and W23) stated whether the farmers and farms they were working with were typical of the catchment and two specified the location in the catchment (AC26, W21).

In most documents, it was not stated how the farms or farmers were selected. Where it was stated it was purposive in six documents (A35, W23, AC15, AC26, O11, AC21) and self-selecting in one case (AC21 Annex farmer survey).

Four spice compliant documents reported working with other existing farmer networks (AC26, O13, O15 and O11). AC26, O13 and O11 mentioned working with CSF, O15 and O11 discussed how all 4 catchments are part of the Catchment-Based (CaBa) approach and have active CSF processes. In addition, this embeddedness in local partnerships was important for DTC success (for example Upstream thinking Payment for Ecosystem Services (PES) in Tamar, Saving the Eden Consortium etc).

AC15 reported that the wider farming community were asked about their awareness of the DTC. Over half of the survey participants had participated in DTC activities (reported in AC16) with the most respondents having participated coming from the Eden and least in the Avon. AC16 also reported on the effects DTC had on their behaviour with the mean failing between ‘a little’ and ‘a fair amount’.

5.2.4 What is the depth of the engagement?

As noted in O11, there was a trend over time to more two-way discussions and more active involvement within the catchments. The conceptual model assumed that information provision will be backed up by consultation with relevant stakeholders on appropriate mitigation measures and behavioural barriers to uptake as well as confounding factors. There would be active involvement of those stakeholders actually engaging in new mitigation interventions.

Very few documents provided an explicit explanation of why they chose to engage as they did (A35, AC26, O13 and O11) with AC26 and O11 explaining how collaborative approaches in the field were needed to move beyond information provision to build trust and form networks.

Overall, the 12 SPICE documents suggested the DTC followed good practice by using a range of approaches including information provision via regular and social media; consultation via workshops, conferences, meetings, engagement events, farmer events, research collaboration events (both DTC badged and presenting at other meetings e.g. CSF etc) and active involvement in negotiating and designing the implementation and maintenance of measures; as well as monitoring. However, as shown in the answer to the primary question, the degree of active involvement was limited compared to the information provision and consultation activities.

5.2.5 How robust are the evidence sources?

Of the 12 SPICE compliant documents reviewed, 6 were reviewed for robustness (the other six did not provide sufficient methodological detail to be rated). These documents reported on social science that were a combination of qualitative, quantitative and/or some economic data (A35, AC15, O13, O11) and quantitative (AC21, A9, W21) research approaches. Of these seven documents, three were rated satisfactory (W21, O11 and O13) and four were rated good (AC21, AC15, A35 and A9). The table of robustness results can be found in Appendix G (Table G2).

5.2.6 Summary: What evidence is there that the DTC engagement methodologies appropriately informed, consulted and actively involved farmers and other stakeholders to maximise uptake of interventions?

The DTC programme engaged a range of appropriate stakeholders using an array of methods to inform, consult and actively involve these stakeholders. However, there was limited evidence available about the overall reach of the engagement, and little information about the active involvement strategies used when engaging farmers about the specifics of why and how they decided to implement DTC measures.

5.3 SS2: What evidence does the DTC data provide about non-participants and why they did not engage in the process?

5.3.1 Relevant Evidence Sources:

A9, AC21, AC26 and O11 provided information about non-participants in their studies. O13 provided information on 13 'non- participant' farmer behaviour types in relation to adoption of 'environmental schemes' but these were not specific to non-participants in the DTC interventions.AC15 noted that as survey "participants were largely drawn from those who had previous engagement with the DTCs there was no special effort to include non-engaged farmers" (p12).

5.3.2 Who did not engage?

All four documents focussed on farmers as non-participants, so we do not know a about other stakeholders (e.g. advisors, contractors, Defra agency staff, water utility staff).

A9 recorded that some farmers in the Eden, Wensum and Avon who responded to the survey were not currently implementing some of the proposed measures and were unlikely to adopt proposed interventions as part of the baseline survey. This did not tell us anything about whether they changed their minds during the DTC interventions. There was no further information provided.

AC21 noted that some farmers in the Eden and Wensum did not engage but did not provide any information about why they did not engage.

AC26 noted that some farmers in the Avon did not want to engage for an unspecified 'range of reasons' but also noted that collaborative approaches can 'bring even the most resistant farmers on-board' (p3).

O11 described how one farmer did not want to engage in the Eden due to tenancy and succession issues. O11 noted that many arable farmers in the Wensum, Tamar and Avon were not very willing to implement buffer strips, whilst 22% of farmers surveyed were not intending to take up any new measures. However, no further information was provided that specifically referred to reasons for their unwillingness.

5.3.3 Why did they not engage?

One document (O11) provided information about reasons not to engage or take up specific DTC measures, this was a farmer who was unwilling to implement DTC measures in the Morland sub catchment on the Eden due to the likelihood of forthcoming change in tenancy. Note that this document was scored as the least scientifically robust of the three documents evaluated.

5.3.4 Did this influence the uptake of DTC interventions in the catchment?

It was unclear why many farmers did not engage, or take up measures, using SS2 criteria alone, but further insights can be gleaned from SS3, which reflects on why farmers may not wish to implement measures to tackle diffuse pollution for agriculture.

5.3.5 How robust are the evidence sources?

We were able to evaluate A9 (Good), AC21 (Good) and O11 (Satisfactory). The types of documents were a peer reviewed paper and two non-reviewed reports. All three reported on a combination of social and physical sciences and within the social science the types of science covered qualitative, quantitative and economic research. The robustness scores can be found in Appendix G (Table G3).

5.3.6 Summary: What evidence does the DTC data provide about non-participants and why they did not engage in the process?

There is very little evidence provided about those who did not participate in the full range of DTC processes and little explanation about why certain stakeholders did not engage with the project.

5.4 SS3: What evidence is there that the DTC considered behavioural factors when engaging farmers in implementing interventions?

5.4.1 Relevant Evidence Sources

13 SPICE documents provided information for SS3 (A35, A9, AC15, AC21, AC31, AV11, AV8, E9, O11, W23, W42, W7, W94).

5.4.2 Were behavioural factors considered?

The following conceptual model of policy interventions and influences on behaviour guided the development of the DTC Phase 2:

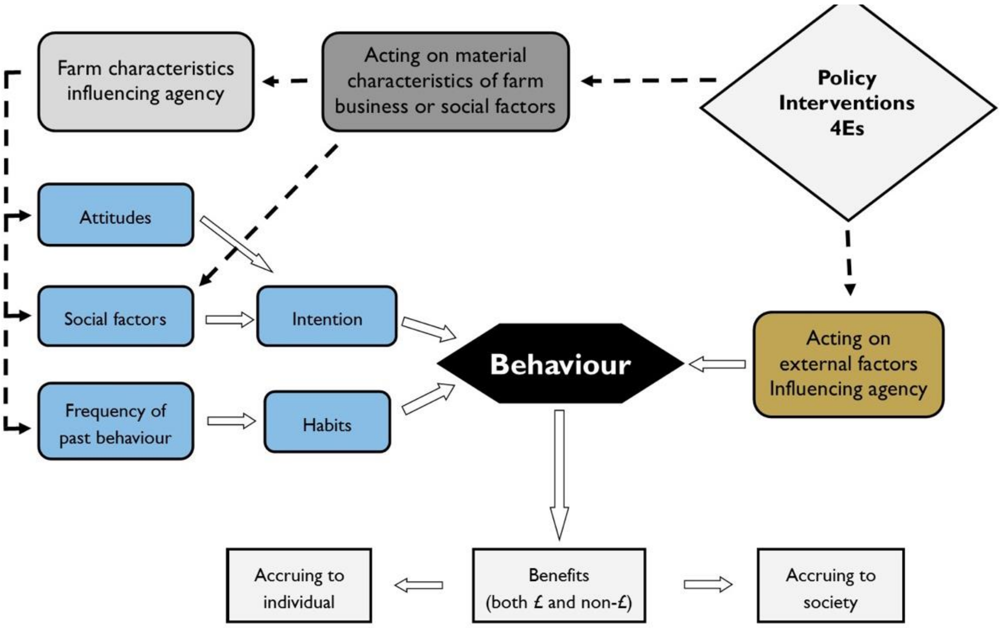

Figure 3: DTC Phase 2 Conceptual Model (DTC, 2016, O11).

The work reported in the documents reviewed clearly showed that data had been collected on the baseline levels of most of the factors influencing behaviour present in this conceptual model. Farm surveys had been undertaken to collect data on baseline farm characteristics (O11), attitudes (O11), and frequency of past behaviour, which had been measured through records of current (pre-DTC) uptake of interventions. Of the 88 farmers surveyed in baseline Farm Survey across Eden, Wensum and Avon, 87% were already enrolled in ELS and 40% in HLS (O11). AC21 reported that the type of farm, land tenancy and enrolment in an existing AES did not influence probability of choosing a new AES (in a stated preferences survey).

In reference to attitudes:

- Two documents reported the level of awareness of diffuse pollution from agriculture was recorded (A35 and AC15). AC15 was the only document that reported changes in the understanding of diffuse pollution within the DTC project.
- Eight documents (A35, AC15, AC21, AC31, AV8, E9, O11, W94) reported farmers' perception of the efficacy of interventions. While farmers reported mixed perceptions for most interventions, the documents also highlighted which interventions appear efficient to farmers (Cultivation and drilling across slopes, Track re-surfacing) and those for which farmers who were reported as being sceptical about their efficacy of the following interventions to address diffuse pollution.
- Two documents reported that farmers' attitudes to the environment were recorded (A35 and AC31).
- Six documents reported that farmers' attitudes to change and innovation were recorded (A35, AC21, AC31, AV11, AV8, E9).

Regarding frequency of past behaviour:

- three documents reported that farmers' agri-environment experience was recorded (A9, AC15, AV8).

Farmers' identities were key to delivery of ecosystem functions by farmers (AC31). Productivist identities did not align with ecosystem service provision, as their focus was on production of food, not ecosystem services. There was a need to foster move from a productivist identity to "multi-functionalist" identity. A35 also provided evidence that identity, beliefs, agency social norms and social network were potential factors affecting uptake of interventions. AC15 noted that pro-DWP was not part of social norms.

Other social factors were found in AC31 (preliminary findings) which reported on the importance of acknowledging agricultural as well as non-agricultural sources of pollutants. Farmers were less likely to engage if other sources were not acknowledged (see also W23). AC31 highlighted the importance of neighbours and consumers as influencers of uptake of mitigation measures, as well as family members and local residents. O11 highlighted the importance of champion farmers and early adopters.

Studies reporting external factors that might influence uptake mostly focused on the role of advisors influencing agency (W42, AC15), the importance of demonstration for uptake (W42, AC31) as local context matters, and on the role of information (AC15). Conflicting advice could be a barrier to uptake (AC15, A9 based on literature review).

5.4.3 Is the connection between behavioural factors and engagement, and subsequent uptake, made?

Less evidence of uptake of interventions was provided in the documents reviewed. The evidence of uptake at the catchment scale was limited to a few targeted farms involved in the physical science research. We could not find evidence that reported diffusion of uptake to other farmers in the catchments as the DTC project progressed.

Two documents provided evidence that behavioural characteristics explain engagement in DTC activities (A35 and AC15), and three documents (A35, AC15 and E9) reported that the understanding of diffuse pollution influenced uptake of interventions, for a wide range of interventions.

Most of the evidence linked behavioural factors to the **intention** to uptake of interventions (e.g. 2014 motivational survey, baseline survey a well as defined choice experiment survey (DCE) (AC21) - most sources only reported findings from the baseline surveys and few results from the DCE are reported). For example, AC15 reported on whether the farmers' perception of interventions, in terms of environmental benefit and, private financial benefits, influences intention to adopt them but not whether this perception actually led to an uptake of the intervention nor whether these benefits were realised.

Different behavioural aspects were reported as underlying the intention of uptake of different interventions or past uptake of interventions (W7, O11 and AC15-16). For example, in the case of cover crops, W94 reported that farmers who were growing cover crops appeared to be innovative and willing to experiment, with one commenting 'It's going to be trial and error to work out what is best'. E9 instead reported from the baseline survey that sediment traps had the lowest uptake rate. There were mixed responses regarding intentions to adopt, and sediment traps were more popular in areas that are wet an unproductive, with farmers' being motivated by environmental (wildlife) and aesthetic preferences as well as by perceived flood risks. W7 assessed eight behavioural aspects (social/cultural, economics, institutional, environment, demographics, automatic and reflective motivation, capability) for 11 mitigation options. The main barriers to uptake were economic (prices of crops) and reflective motivation (waste of land).

Finally, of 21 measures planned to be introduced in 2012, less than half were implemented by 2016 and five of the introduced measures were only partially successful (AC16). Nine farmers found alternatives or no longer needed to do them and two were not introduced due to financial constraints. How robust are the evidence sources?

5.4.4 How robust are the evidence sources?

Five of the 13 SPICE documents providing information for SS3 were assessed for robustness as the remaining 8 did not provide enough information on the method used to be assessed. Four documents were assessed as Good (A9, A35, AC15, AC21), while one was assessed as Satisfactory (O11). Further information can be found in Appendix G (Table G4).

5.4.5 Summary: What evidence is there that the DTC considered behavioural factors when engaging farmers in implementing interventions?

There is evidence that behavioural factors were taken into account when designing the project and planning engagement activities. However, the reporting of these behavioural factors and the evidence that these might have affected uptake of interventions, was limited.

5.5 SS4: What evidence is there of engagement and uptake of DTC interventions varying between catchments?

5.5.1 Relevant Evidence Sources:

Of the 15 SPICE documents, nine documents contained data relevant to more than one catchment:

- Two O11 and AC21 covered all four catchments.
- Four covered the Eden, Wensum and Avon (A35, A9, AC15 and W42)
- One (AV11) covered the Avon and Tamar
- One (AC25) covered the Eden and Wensum. Note AC21 also covers these two catchments
- E9 focussed on the Eden but does provide some comparative data with other DTC catchments.

However, A9, A35, AC25, AV11, O11, W42 did not present comparative data between catchments. O11 reported some differences between farm types that may give some insights into differences between DTC catchments, given their different overall mix of farm types. However, AC21 stated 'The probability of choosing the status quo or any agri-environmental option is not significantly affected by the type of farm, the area of a feature, the land tenancy regime or whether they are enrolled in any agri-environmental scheme'. (p80).

5.5.2 How did engagement processes vary between catchments?

Types of stakeholders

Farmers were engaged in all catchments (O11).

Other stakeholders were engaged in the Avon (O13, A9, A15); Eden (O11, AC25, AC21, A9, A15); Wensum (AC25, AC21, A9, A15). It is unclear how much interaction with other stakeholders occurred in the Tamar although reference is made to the wider catchment activities e.g. PES scheme involving other stakeholders.

Types and depth of engagement

Farmers were engaged by: information provision (O11); consultation via surveys (AC21, AC26, AV8, O13, AV11, AC15, E9, AC15); interviews (AV11, AC15, E9, AV11, A9) and workshops or discussion groups (AV11, E9, A35); and farm visits (AC25, AC21, O11).

Farmers were also actively involved in monitoring in the Wensum (using monitoring kits) and Eden (using farm diaries) (O11). A farm (estate) in the Wensum was used to demonstrate DTC interventions to other farmers (AC26).

Other stakeholders were engaged through farm visits and walks as well as across all catchments in stakeholder workshops (O11). The potential DTC intervention options were often discussed with advisors and farmers in discussion groups before finalising choices with the individual farmer.

Numbers of engagement

As it is unclear exactly how many farmers were informed, consulted or actively involved across the DTC programme, it is impossible to be accurate about any differences between catchments. O11 and O13 do give overviews of the entire suite of DTC engagement or Knowledge Exchange activities but the numbers are not produced for discrete catchments.

Working with farming networks and CSF

O11 notes that all four DTC catchments involve actively coordination with CSF and CaBa processes.

5.5.3 Was there a difference in uptake between catchments?

Buffer strips and Livestock grazing and stocking were taken up in all four catchments. Farmyard management and Traffic on fields was taken up in Avon, Eden and Tamar. They were not taken up in Wensum although these measures were discussed in more general terms in consultation with farmers. Wetlands and ponds were not taken up in Tamar and Wensum. Cultivation measures were not taken up in Avon or Tamar– though these measures were discussed in more general terms in consultation with farmers. There were many specific 'other' interventions that were bespoke to a single catchment.

The PS data extraction noted additional references to interventions on the Dacre and Pow sub-catchments of the Eden, covering water attenuation and rural SuDS, soil aeration, riparian fencing and woodland creation, however within documents O11 and AC21 these measures are described as being proposed for implementation during phase 2 of the project, rather than reporting on what had been implemented. For completeness, the measures have been added. A table illustrating interventions across the catchments can be found in Appendix G (Table G5).

5.5.4 Were there differences in behavioural aspects?

E9 reported that Eden farmers were the least likely to start sharing machinery, join a discussion group or develop joint countryside stewardship agreements compared to other DTC areas. AC15 also found that Eden farmers were less likely than farmers from Avon or Wensum catchments to incorporate manures into soil within 24 hours and Avon farmers particularly stressed financial constraints and weather (confounding factors) as negative impacts on uptake, compared to practical impacts on farming practices in Wensum (within farmer control). Wensum farmers were also more likely to undertake joint activities to control DWPA (AC15).

O11 reported some differences in current uptake (as of 2012-13) and future attitudes to uptake of mitigation measures between farm types (arable, lowland livestock, dairy and mixed) that may give some insights into differences between DTC catchments. Arable farmers already have high uptake (more than 75%) of cultivation, buffer strips, cover cropping, fertiliser application, in field manure management, traffic on fields and other (plant hedges) types of mitigation measures. Lowland farmers already had high uptake of livestock grazing and in-field manure management measures. Dairy farmers shared some existing measures with lowland livestock (livestock grazing measures) but also used field drainage and fertiliser application measures. Finally, mixed farmers shared common interventions (cultivation, in field manure management, traffic on fields, fertiliser management, livestock grazing) with the other farming systems. Overall dairy and lowland livestock farmers had fewer existing measures in place. Dairy farmers reported a positive future attitude to uptake of large range of mitigation measures, compared to arable, mixed and lowland livestock. The latter is the most noteworthy, given how few measures were already implemented.

AC21 reports that Wensum and Eden differ in terms of land use (more arable crops in the Wensum, Eden dominated by grazing) but broadly speaking the views of the farmers on the environmental benefits of DWPM measures were similar. The differences discerned were most farmers are ignorant of the potential environmental benefits of measures in the Eden, in the Wensum farmers associated DWPM measures more with wildlife benefits and not mitigation of diffuse pollution. Note that these data do not specify if these mitigation measures are prompted and installed via the DTC programme, or general DWPA measures (including those installed before the DTC began).

5.5.5 How robust are the evidence sources?

For the five SPICE documents that have robustness scores, two were satisfactory (O11, O13) and three were good (A9, AC15, AC21). The robustness scores can be found in Appendix G (Table G6).

5.5.6 Summary: What evidence is there of engagement and uptake of DTC interventions varying between catchments?

The evidence was often summarised across catchments, making distinctions between catchments more difficult to assess. Engagement does not seem to vary between catchments. There are some differences in uptake of measures and intentions to uptake further measures, but these tend to be related to farming systems and personal attributes than being specific to the geography of individual catchments. PI feedback argues that the Eden differed from the other catchments in terms of the complex social landscape of small, often economically challenged, family farms, increasing the challenge to connect the effect of an intervention on the receiving sub-catchment.

5.6 SS5: What evidence is presented on the cost effectiveness and benefits of DTC interventions, during the initiative and for the five-year period beyond the end of the DTC initiative?

5.6.1 Relevant Evidence Sources:

Nine SPICE documents provided some information on the cost effectiveness of interventions: 1 newsletter (AC25), 3 peer reviewed papers (A9, A11, A35) and 5 reports (AC15, AC21, O11, W7, W21).

5.6.2 Is Cost-effectiveness or Cost-Benefit analysis undertaken?

Looking at the nine SPICE documents, one provided a full cost benefit analysis, four provided full cost-effectiveness analysis (CEA) of interventions, seven some information about the costs and effectiveness of interventions[2], and two provided information on uptake but not on cost-effectiveness analysis. Most reported the cost effectiveness of several interventions.

It seems, from the documents reviewed, that during the first stages of the DTC projects, CEA relied on the FARMSCOPER tool to provide estimates of costs from the adoption of interventions (e.g. A9). PI feedback highlighted that the DTC programme did use research that assessed regional variations in measure costs and that using 5-year average costs captured any regional variations. . In later steps of the project monitored costs from trials on engaged farms are reported (AC21, O11, A11).

The FARMSCOPER tool estimated the average cost of adoption and effectiveness of interventions for a "standard" representative (at the national level) farm for different production systems. This meant that the costs reported are not specific to the DTC catchments.

A11 provided a full cost-benefits analysis of roadside wetlands, providing a monetary value to the reduction of diffuse pollution, which is then compared to the costs of the intervention. The authors concluded that a standard wetland would have a payback time of 8 years, after which the benefits start compensating the costs. AC21 (2018) provided a full cost effectiveness analysis of the different interventions trialled on engaged farms on the Eden and Wensum, some assessment of the costs of the interventions assessed on the Tamar, while no information on the cost effectiveness of interventions trialled on the Avon are reported. O11 provides some estimates of costs for all four catchments. PI feedback noted that there were lengthy debates about how to

[2] Some documents provide a full CEA for some of the interventions while only some assessment of the costs and effectiveness of other intervention, hence the total number is larger than the number of SPICE documents.

calculate environmental damage costs, and that using the FARMSCOPER values help with consistency and ability to compare across catchments. Assessments for a wide range of interventions are summarised in Appendix G (Table G7).

With the exception of A11, the remaining documents did not conclude whether the interventions are cost-effective or not per se, but did provide assessments of how costly each intervention was in regard of the environmental improvements achieved – which will be useful to compare alternative intervention options to achieve diffuse pollution reduction objectives at the lowest cost (A9, AC21, O11, W21). For example, A9 concluded that "a 95% implementation of the 29 preferred measures by farmers would lead to a (median) decrease, relative to BAU, of Nitrates by 10.6%, Phosphorus by 15.2%, Sediments by 19.5%, NH4 16%, CH4 10.5% and N2O 6.9%, for a median cost of £3 per hectare, including 27£ of fixed costs and -17£/hectare of variable costs". The cost-effectiveness assessments provided by the researchers in the documents will have to be assessed in the light of the intervention beneficiary's willingness to pay (benefits) to achieve such a reduction in diffuse pollution. Others relied on qualitative assessments made by farmers on perceived costs and benefits of different interventions, e.g. percentage of farmers who believe d the adoption of interventions would yield benefits or costs to the farm (AC25, A35, W7, AC15).

5.6.3 Are multiple types of costs and benefits and their distribution across stakeholders considered?

There was an extensive focus on costs and measures of effectiveness in physical units for different water pollutants. Looking at the benefits generated by the DTC interventions, one document mentioned private benefits to farmers (AC15), two mentioned societal benefits (A11, W21), and three more mentioned both private and societal benefits (A35, AC21 and AC25). Amongst private benefits to farmers, financial benefits (farm productivity and economic returns) were mentioned (AC15, A35) as well as erosion prevention by fencing, which also limited the spread of livestock diseases (liver fluke), and looking after soil limited weeds and the need for herbicides (AC25). In regards to societal benefits, only the benefits of improved water quality for the general public was mentioned. These benefits were not quantified.

There is a much larger breadth of evidence on the costs of the interventions generated by the DTC interventions. Most documents focus on farmers direct costs (6 documents, A9, A11, O11, W21, W7, AC21) to get a sense of farmers' likelihood of uptake, and AC21 reports a combination of stakeholders bearing the costs of interventions.

Transaction costs are reported as a barrier to agri-environmental scheme uptake in the literature (e.g. Franks 2011), but these do not seem to have been assessed as part of the project, probably as the interventions were not yet implemented under a "scheme" in the DTC project, and hence these transaction costs could not be assessed.

5.6.4 What is the estimated commitment to the intervention? Were payments sufficient?

There was very little evidence related to these questions. A single document (AC31, newsletter) provided evidence on the predicted or estimated commitment to interventions post DTC initiative. It mentioned that the farmers discussion groups were established in the Avon with the objective that they would last beyond the end of the project but without assessing the actual duration. Two documents provided evidence on whether the payments offered to adopt DTC interventions were high enough to cover their net perceived costs of adoption (A35 and AC15), concluding they are not. A35 highlighted the need for fiscal incentives, public environmental payments and payment for ecosystem services to financially empower those wanting to mitigate DWPA.

5.6.5 How robust are the evidence sources?

Six of the nine SPICE documents providing elements of cost-effectiveness analysis were assessed for robustness (A11, A35, A9, AC15, AC21 and W21). The remaining three did not provide information on the methods so could not be assessed for robustness. It is to be noted that AC21 reports CEA for 3 different catchments with different methods used between catchments and different degrees of robustness (from Satisfactory to Excellent). Overall, the robustness of CEA presented in the documents is good (1 excellent (A11), 4 Good (A9, A35, AC15, AC21), 1 Satisfactory (W21). The inclusion of a discussion to justify the choice of costs and benefits flows included or excluded discussed is variable amongst the documents assessed. The reasons for not scoring Excellent are: failing to report sources of economic data used to assess costs and sensitivity analysis not being presented. CEA which rely on a pre-established tool (FARMSCOPER) tend to score lower as the methodology is likely to be more fully described in other documents (outside of the DTC remit) so we are not able to judge its robustness within this review. However, most documents consider flows of costs and benefits over the lifespan of the intervention, which is good practice. The robustness scores can be found in Appendix G (Table G8).

5.6.6 Summary: What evidence is presented on the cost effectiveness and benefits of DTC interventions, during the initiative and for the five-year period beyond the end of the DTC initiative?

There were few specific cost-effectiveness analyses undertaken and a single cost-benefit analyses done for a specific intervention (roadside wetlands). There was a reliance on the FARMSCOPER model for average costs at the beginning of the project, and a tendency to focus on direct private costs, with little information on the wider costs or benefits of an intervention. There was very little information about the potential for measures to be sustained beyond the life of the DTC project, but some sources suggested that payments were not enough to entice farmers to bear the costs of mitigation.

5.7 SS6: What evidence is provided that the confounding factors (e.g. existing non-DTC activities) were accounted for when reporting on engagement and uptake of DTC interventions?

5.7.1 Relevant Evidence Sources:

We have taken a broad perspective on confounding factors that cover: Changes in Regulations/ Markets/Incentives/Information/Support/Technology/policy/Public Opinion/land use or ownership change (see conceptual model, Figure 1). Of the 15 SPICE documents assessed, one reports limitations and how confounding factors affected their results. However, most SPICE documents do discuss the relevance of results to other contexts and propose conclusions that are backed up by findings.

5.7.2 Were confounding factors affecting engagement considered?

Two of the 15 SPICE compliant documents were evaluated to have specifically addressed confounding factors affecting engagement (AC15, E9), which covered the Eden, Wensum and Avon catchments. The confounding factors affecting engagement were the many demands on farmers' time and some farmers going out of business between the survey period, which reduced their ability to engage them in later phases of the research. Strictly speaking, farmers having insufficient time to engage with the DTC project would be a behavioural issue not an external confounding factor. However, the DTC project personnel were not able to control for other demands on farmers' time, and therefore lack of time would affect the ability of the DTC project to engage all relevant farmers. PI feedback draw attention to the impacts of flooding, particularly Storm Desmond, that refocussed the attention of stakeholders – as our SPICE focus screened out any focus on water quantity, we did not reflect on the impacts of flooding on the processes of stakeholder recruitment, nor farmers willingness to uptake measures.

5.7.3 Were confounding factors affecting behaviour considered?

Seven of the documents were evaluated to have addressed confounding factors affecting farmer behaviour, which is a precursor for uptake of the interventions (A35, E9, W23, AC15, A9[3], AV11, O13, O11, AC 21). As with engagement, some confounding factors are not strictly 'external' to the DTC project: lack of time for farmers to learn about the measures (A35, AC15, O11); issues with lack of salient information and advice on mitigation measures (AC15, A9, AV11, E9, O13, AC21); lack of evidence of cost-effectiveness of measures (A9,AV11, E9, O13); succession planning (A9);

[3] The A9 references were made to wider literature on uptake of measures but did not report specific behavioural data from the DTC catchments.

personal preferences and values (E9, AC15, A9); the farm environment not being perceived as suitable for some measures (AC15); and resistance to taking responsibility for DWPA (A9). These were all central to the overall objectives of the DTC itself.

Traditional external confounding factors were also noted such as market prices that were focussing farmers on maximising profitability (especially for dairy and pig producers) (A35, W23, AC15); tenure arrangements affecting decision making (AC15); issues with accessing HLS funding streams for the measures (W23); overly rigid management prescriptions associated with funding schemes (A9) or regulations (O13); and some farmers' perceptions that environmental regulations were not adequately enforced (AC15, A9, O11).

These confounding factors were mainly 'negative' influences on the DTC work, making it more difficult to persuade farmers to consider uptake of the measures. However, there were also positive confounding factors such as high commodity prices making farmers more able to afford to innovate (W23) (although conversely O13 found farmers more likely to seek advice on DWPA measures when profitability decreased); increased availability of grants (O13, E9); improving access to social networks (A9); prior experience of schemes (A9) and increased public concern for food security (W23).These seemed to increase the farmers' interests in learning about potential DWPA mitigation measures.

5.7.4 Were confounding factors affecting uptake considered?

Six documents (A35, AC15, E9, AV11, AC21 and O11) covering all four catchments considered confounding factors affecting uptake of interventions.

Weather was cited by two documents (A35, AC15); and the farm 'environment' by three documents (E9 and AV11, AC21). Although Storm Desmond was referenced in a number of the physical science documents, it was not noted as having an impact on the socio-economic context of the farms or influencing actual or proposed uptake of the DTC measures in the social science extraction table. Economic issues were presented in five documents such as access to finance or grants where large capital investment (AC13, A9, AV11, O11, AC21). The constraints of CAP and associated support schemes were mentioned in three documents (AC15, AV11 O11) although the introduction of Greening measures under the CAP was positive for uptake of cover crops (O11). Social issues also mattered, with the positive influence of Public Opinion and importance of protecting environment contrasted with other views around public ignorance of farming practices (AC15). As noted in SS3 and SS4, there were issues regarding tenants being willing or able to take on mitigation measures (O11) due to tenancy agreement timescales.

As discussed in the engagement (section 5.2) and behaviour (section 5.4), there were also factors that influenced the social learning associated with implanting measures. These included a high turnover of CSF staff who were providing advice and information on measures (AV11) and an insufficient evidence-based advice about the practical elements of measures for farming practices (AC15, E9, AV11). Finally, there was reference to a lack of time for farmers to commit to the implementation of measures (AC15). The transaction costs required to search for, evaluate, oversee and maintain measures were important considerations although rarely reported or quantified (see section 5.6 above).

There was no explicit mention of technological confounding factors affecting engagement, behaviour or uptake in the documents reviewed, although, the cost of advanced technologies involved in some mitigation measures (e.g. subsoiling) is referred to in terms of economic barriers.

5.7.5 Is the additionality of DTC activities recorded?

One paper (A9) accounted for measures that were already in place at the baseline situation to measure additional effects compared to current situation. Of the farmers that engaged in the baseline farm survey across Eden, Wensum and Avon, 87% were already involved in ELS and 40% in HLS (O11, AC21). Documents did provide a baseline of mitigation measures already in use when the DTC proposed further interventions (AV8, O11) but there was not any later explicit reporting of additional uptake or effects compared to this baseline. The documents did not explicitly whether attitudinal change and uptake of measures would have happened with or without the DTC.

Some documents (AC15, O11 and non-SPICE relevant AC12) discussed an evaluation of the Catchment Based Approach in the DTC catchments. However, there was no explicit link made between the roll out of the CaBa approach and the influence on engaging farmers in DTC activities or uptake of DTC interventions.. Likewise there are a number of documents that mention the Catchment Sensitive Farming Initiative (CSF) e.g. (AC26, O11, AC21), but these schemes were not explicitly considered when reporting on results. The role of CSF was also discussed in SPICE relevant documents (AC6a, 6b W12, A64, W32, W41). These schemes were confounding factors for the DTC as they will have affected the DTC results.

It was difficult to disentangle the DTC funded activities from other complementary Defra projects e.g. WQ0225 (AC21), WQ0127, WQ201 and WQ0106 (O11). O11 also notes that the DTC built on evidence from research undertaken as part of previous studies in the catchments (e.g. PARIS in the Avon and CHASM in the Eden). It was also difficult to understand how the DTC added value to the Upstream Thinking Payment for Ecosystem Services Project, being implemented by the West Country Rivers Trust and South West Water in Caudworthy Sub catchment, where it appears that many proposed measures could be or were funded by this scheme (O11).

5.7.6 How robust are the evidence sources?

Engagement: One of the documents (E9) was not reviewed for robustness. The other document (AC15) scored Good for Robustness.

Behaviour: One of the documents (W23) was not reviewed for robustness. There were two satisfactory scores (O11, O13) and four good robustness scores (A9, A35, AC15, AC21).

Interventions: Two of the documents (E9, AV11) were not reviewed for robustness. There was one satisfactory robustness score (O11) and the remaining three documents were good (A35, AC15, AC21).

Additionality: The robustness score was good (A9).

The table reporting the details can be found in Appendix G (Table G9).

5.7.7 Summary: What evidence is provided that the confounding factors (e.g. existing non-DTC activities) were accounted for when reporting on engagement and uptake of DTC interventions?

There was limited evidence to see whether and how confounding factors were considered for engagement, behaviour and uptake. These factors were varied and there were no clear patterns across time or catchments. There was limited evidence illustrating where the DTC project 'added value' to other initiatives but the DTC followed good practice by trying to integrate with existing farming and stakeholder networks as much as possible.

6 Monitoring Methodologies

This section provides an evaluation of the monitoring methodologies adopted by the DTC programme. It was included in recognition of the value of the monitoring and resultant data. Evidence was not restricted to PICO compliant sources as many of the datasets and reports describing the infrastructure did not include analysis of intervention effectiveness or report on outcomes. The evaluation focuses on the monitoring network, monitoring equipment, choice of determinants, sampling frequency and account of non agricultural pollutants. Preliminary analysis of primary data is included to demonstrate how the data collected provide a second opportunity to identify the presence of non agricultural pollutants.

6.1 The DTC Monitoring network

The Demonstration Test Catchment (DTC) project was established in 2009 to test the hypothesis that 'it is possible to cost-effectively reduce the impact of agricultural diffuse water pollution on ecological function while maintaining food security through the implementation of multiple on-farm mitigation measures' (O2). The project established monitoring programmes in four catchments (Eden in Cumbria, Wensum in Norfolk, Avon in Hampshire and the Tamar in Devon/Cornwall) to build on existing infrastructure, datasets, knowledge and farming contacts all developed through previous initiatives. These catchments provide good national coverage and representativeness of different physical and socio-economic factors relevant to diffuse pollution (see section 10 for further detail).

At an early stage reports were commissioned to provide guidance on the experimental design and monitoring strategy (O2) and data management requirements (O5). These reports define the optimal monitoring infrastructure against which we can evaluate the DTC; acknowledging that the implemented infrastructure will have reflected a compromise given budgetary constraints.

All catchments sensibly focused on small headwater streams (~10 km^2) with many subcatchments less than 5 km^2 and overall catchment areas less than 30 km^2 (O11). This enabled interventions to be trialled as intensively as possible over a small areas. The monitoring programme was designed to monitor interventions singularly (often undertaken by specific research projects e.g. W4) and in combination and at scales ranging from plot to catchment scale (e.g. see A15, W4). The monitoring of small subcatchments and at the plot scale meant that in many cases, following recommendations, monitoring was immediately downstream of interventions. Robust BACI experimental designs were used in all catchments. BACI1 experiments adopted an independent control site whereas BACI2 experiments used an upstream site as a control. Subcatchments were carefully selected based on existing knowledge, field reconnaissance, and stakeholder consultation before being signed off by the project's Research Advisory Group.

Hampshire Avon DTC: Three subcatchments were chosen. The Sem subcatchment was monitored at Cool's cottage (2.6 km^2) and Priors farm (4.6 km^2) providing a BACI1 design. The Ebble was monitored at two points at Ebbesbourne Wake (16.7 km^2); providing a BACI2 design. The Wylyle was monitored at Kingston Deverill (25.2 km^2) and Brixton Deverill (50.2 km^2) to provide a BACI2 design.

Tamar DTC: Two subcatchments were chosen. The Caudworthy water subcatchment was monitored at Winnacott (18 km^2) and Caudworthy Ford (26 km^2); providing a BACI2 design. The Neet catchment provided a control catchment and this was monitored at Burracot (10.9 km^2) providing a BACI1 design. The Tamar was added as it included subcatchments with higher densities of on-farm measures.

Eden DTC: Three subcatchments were chosen. The Moorland subcatchment is monitored at Newby Beck (12.5 km^2) and this contains mitigated (1.6 km^2) and control (3.6 km^2) subcatchments. The Pow Beck catchment is monitored at Nabend (10.5 km^2) and this contains mitigated (1.9 km^2) and control (2 km^2) subcatchments. The Dacre subcathcment is monitored at Thackthwaite Beck (10.2 km^2) and this contains mitigated (1.7 km^2) and control (1.3 km^2) subcatchments. All three subcatchments have a BACI1 design.

Wensum DTC: A nested monitoring approach was adopted in the Blackwater Drain subcatchment of the Wensum catchment whereby four first order streams were monitored. A combination of BACI1 and BACI2 designs were used in this catchment. The Merrisons site (3.7 km^2) was monitored as a BACI2 control for the downstream site at Swanhills A (5.3 km^2). The Swanhills B site (1.5 km^2) was monitored as a BACI1 control. A site was also monitored downstream of Swanhills A and B at Stinton Hall farm (7.1 km^2). Monitoring was also undertaken at Brakehills (3.5 km^2) and Black Bridge (6.6 km^2). At the catchment outlet monitoring was undertaken at Park farm (19.7 km^2).

Unfortunately, 2011-12 was very atypical from a hydrological perspective, which compromises the BACI design (see Figure 4). There was a winter drought in 2011, followed by summer flooding in 2012. The years following the interventions (2013-2014) were more typical in rainfall pattern, but had very high winter flows. Ideally, projects should have longer “before” and “after” periods, to ensure that enough data over a full hydrological range is collected.

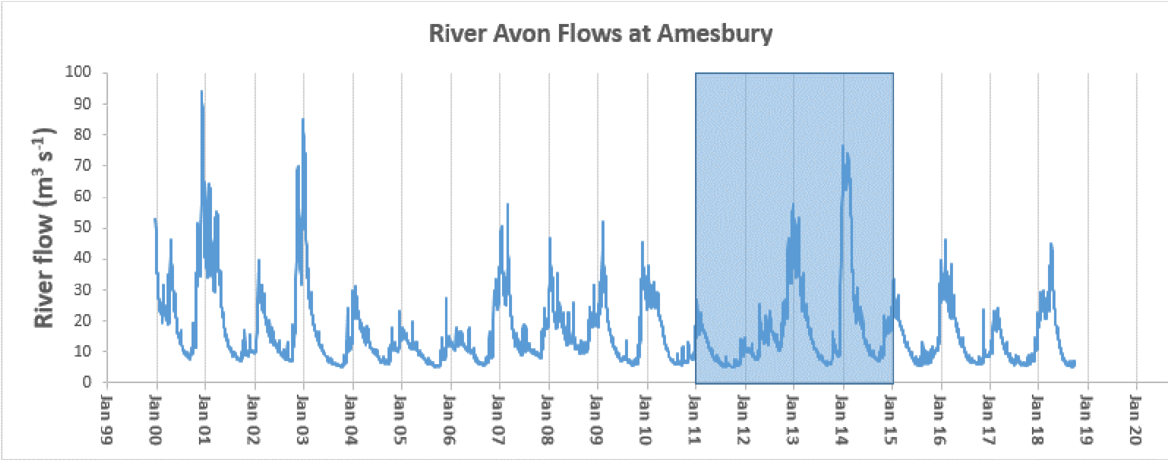

Figure 4: River flow for the Avon at Amesbury from 1999 to 2018.

Groundwater monitoring data have not been provided as outputs from the DTC project although it is clear from O11 that conceptual groundwater models have been produced for each catchment and water quality data exist from other initiatives. Thus, although the importance of groundwater in all catchments has been appropriately identified the limited data are likely to have constrained the full evaluation of interventions. Within both the Hampshire Avon and Wensum DTC, there will be significant groundwater inputs, due to the Chalk geology and high base flow index. Groundwaters in Chalk regions of the UK are usually polluted with nitrate from fertiliser and manure over-applications over the last few hundred years. This high background nitrate enrichment of the study catchments, and how this input will naturally vary under different flow conditions, will make it difficult to detect the impacts of nitrogen mitigation interventions within the DTC project.

This problem is compounded in the Hampshire Avon DTC, due to the presence of Greensand deposits. It is known that Greensands are high in phosphorus, and naturally enrich many of the headwater springs in this region. This naturally-high phosphorus needs to be taken into account when assessing pollution loadings and quantifying sources. It should be noted that no groundwater monitoring was funded in eth Avon DTC.

In the absence of groundwater monitoring in some projects recharge waters were sampled at depth in the soil to evaluate impact (e.g. W4 sampled soil water at 90cm depth in evaluating the impact of measures to mitigate pesticides).

As recommended (O2), in each of the DTC catchments many hydrological, ecological and chemical sampling locations were co-located (see Supplementary material 3). This enabled accurate flux estimation and a fuller understanding of the mechanisms of pressure, impact and response.

Simple rules have been developed in other projects (e.g. WQ0223) to help match sites that may not be co-located.

6.2 The monitoring equipment

Equipment and deployment details were specified in DTC contracts.

Water quality: The project adopted the most appropriate field equipment and used proven kiosk set-ups. Storm sampling was conducted using ISCO auto-samplers; known for their reliability.

The kiosks were equipped with YSI 6600 multi-sondes, which have been proven to be accurate and reliable during long-term mass-deployments by the Environment Agency's National Water Quality Instrumentation Service (NWQIS). In addition, the DTC project decided to deploy these probes within NWQIS-style kiosks, rather than within the streams themselves, to minimise the risk of biofouling, impact of freezing, and reduce the risk of theft / vandalism. These kiosks, incorporating the Phosphax and Nitratax instruments, YSI 6600s with pumping and telemetry systems, have been successfully demonstrated in 2008 – 2011 within the EPSRC-funded LIMPID project (Halliday et al. 2015; Wade et al. 2012). Other existing research projects had also demonstrated the Hach Lange Phosphax to be perhaps the most robust instrument for river research of this type (Jordan et al. 2012). The 30 minute P data compared well with the grab samples at checked sites, such as the Pow outlet (Eden), and at Brixton Deverill in the Hampshire Avon DTC (Figures 5 and 7). Uncertainty in some of the sensor data was evaluated through comparisons with grab samples (O11). Here we observed some discrepancy between the Phosphax TRP data from Brixton Deverill and the corresponding SRP data from the grab samples, which could indicate either that the Phosphax was overestimating the dissolved P load, or that the site has significant soluble organic phosphorus in dissolved unreactive P form. As stated below, total reactive P could have been determined on the grab samples at these Phosphax sites, so this discrepancy could have been investigated. There were also significant gaps in the Phosphax data through the summer – autumn periods at this Brixton Deverill site (64 % of expected TRP data present, Supplementary material 2). It would be useful for each DTC to present the Phosphax P data alongside the grab sample data, so that the accuracy of the Phosphax can be assessed by potential users of the data.

There were a number of choices for providing reliable high-frequency monitoring in nitrate, but the selection of the Hach-Lange Nitratax and YSI Sonde was based on successful applications in previous river research projects. The nitrate data sets checked here show good agreement with grab sample data, although there appears to be a problem with the Nitratax on the Pow Beck, which is out of range above 7 mg NO_3-N l^{-1}, resulting in major nitrate peaks being missed. The selection of a wider calibration range may have helped (Figure 6).

Novel instrumentation was also used in the catchments. Web, motion and time lapse cameras were used in the Eden catchment to engage stakeholders, identify cattle entering watercourses (A63), and to observe landscape changes. Although reliability proved to be an issue these cameras did add value to the monitoring (O11). Sediment fingerprinting and novel magnetic tracing techniques were used to identify the relative important of sediment source contributions and the effectiveness of buffer strips (AC44).

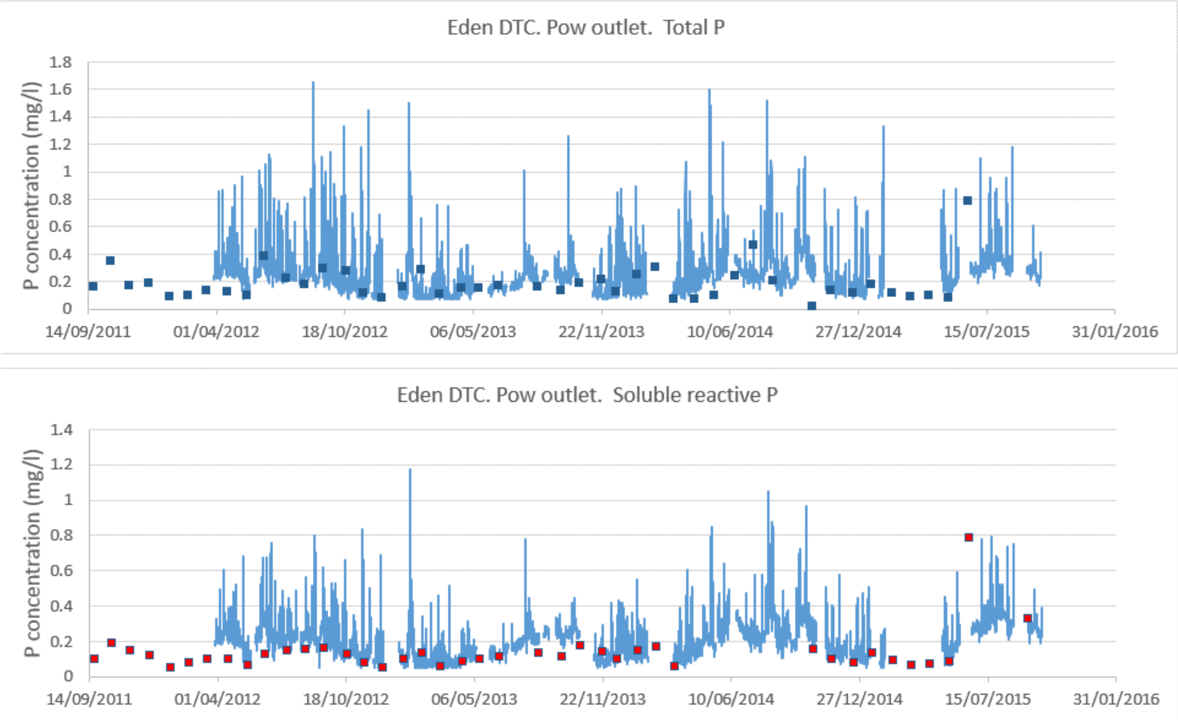

Figure 5: Total Phosphorus (TP; upper graph) and Total Reactive Phosphorus (TRP; lower graph) concentrations produced by the Hach Lange Phosphax, from the Eden DTC, Pow Beck outlet sub-catchment. The blue and red markers are the TP and Soluble Reactive Phosphorus concentrations from monthly grab samples.

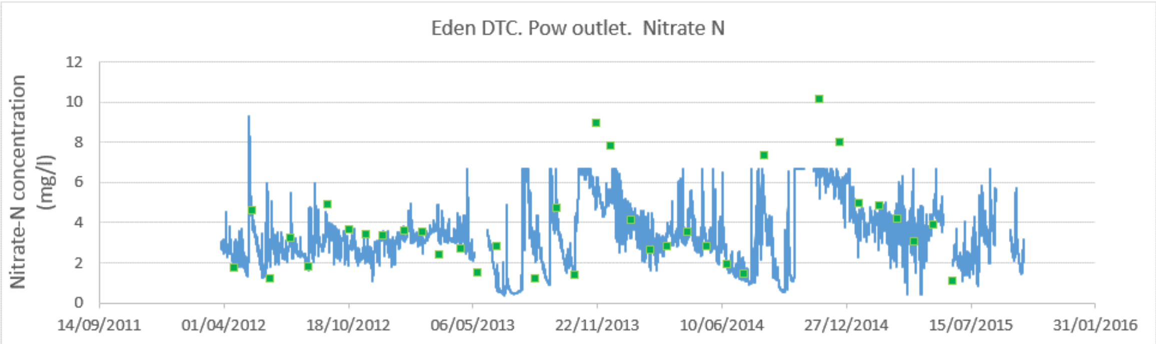

Figure 6: Nitrate-N concentrations produced by the Hach Lange Nitratax sonde, from the Eden DTC, Pow Beck outlet sub-catchment. The green markers are the lab-derived nitrate-N concentrations from monthly grab samples.

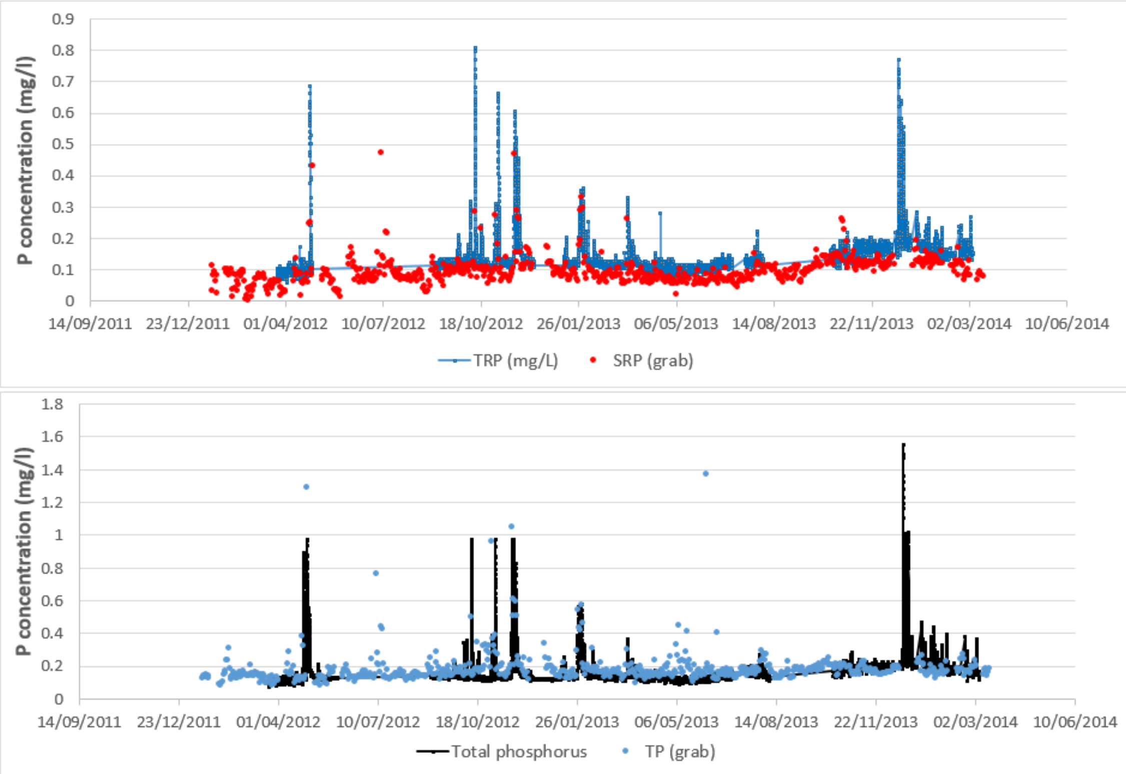

Figure 7: Total reactive Phosphorus (TRP; upper graph) and total Phosphorus (TP; lower graph) concentrations produced by the Hach Lange Phosphax, from the Avon DTC, Brixton Deverill sub-catchment. The red and blue markers are the TP and SRP concentrations from daily grab samples.

6.3 The choice of determinants

The biogeochemical parameters collected by the DTC platform were specified in the contract and were extremely comprehensive. The water quality parameters were selected to identify agricultural diffuse pollution to water bodies, particularly nutrients and sediment. All three DTCs have captured the full range of macronutrient chemical species.

Phosphorus

Phosphorus species included total P, total dissolved P, total reactive P and soluble reactive P. These data can be used to derive specific phosphorus fractions, such as particulate P and dissolved hydrolysable P (equivalent to dissolved organic P). These chosen analytical determinants are standard in nutrient research. TP and TRP were obtained for selected sub-catchments within all three DTCs, using the Hach Lange Phosphax auto-analyser. The lab analysis of the grab samples only seemed to be produced by the Eden and Avon DTCs.

In the sub-catchments that contain a Hach Lange Phosphax monitor, it would have been useful to include total reactive P analysis (SRP analytical method on an unfiltered sample, to mimic the

Phosphax method) in the grab sample lab analysis. This would have allowed the 30 minute TRP data to be more accurately ground truthed, but SRP should provide similar values (Figure 7). It is unclear how the Wensum were able to ground truth their 30 min Phosphax data, as they did not seem to analyse their grab samples for phosphorus.

Nitrogen

A comprehensive range of inorganic nitrogen species were routinely measured across the DTC platform, comprising of total dissolved nitrogen, total oxidisable nitrogen, nitrate, nitrite and ammonium. This range of determinants also allows for organic N fractions to be derived.

Monthly, or even weekly grab sampling is unlikely to catch the peaks of a representative number of storms in these small headwater catchments, resulting in underestimation of N export from the study catchments, but these data were enhanced with storm sampling using automatic water samplers, which covered some of the storms. The inclusion of this will make the load and diffuse agricultural input estimations much more robust.

Carbon

The only carbon fraction that was measured by the DTC project was dissolved organic carbon (DOC). The Eden measured DOC concentrations from the monthly grab samples (54 – 70 % completeness) and also the storm samples (approximately 4 storms per year). The Avon managed to measure DOC (termed NPOC) concentrations at daily frequency, with completeness ranging from 44 to 77 %. The Wensum did not generate any carbon concentration data.

Ecological measurements

Ecological, water quality and hydrology monitoring locations were closely matched in the Avon/Tamar and Eden DTC. Macroinvertebrates, diatoms and macrophytes were monitored as these represent Biological Quality Elements (BQE) that are used in the WFD and respond to specific environmental pressures.

Other determinands

The DTC project has produced data on a wide range of other determinants that could be vital to detect the impact of the interventions. These include some indirect indicators of ecological status, such as chlorophyll, dissolved oxygen, dissolved silicon etc. Other determinants measured by the Eden DTC, such as boron, sodium, potassium etc. could be useful for sewage and fertiliser source apportionment. Unfortunately, these measurements were not sustained throughout the project. Alongside the water temperature, flow and weather station data, these complex data sets could provide a valuable modelling resource to the wider academic community to produce maximum benefit from the DTC programme. One of the major diffuse agricultural inputs to the river network will be from pesticides, both from transfer through the soil and through the air during application. Mitigation measures such as installation of buffer strips is partially designed to minimise this impact. Pesticide analysis was only conducted in the Wensum catchment, rather than all three DTCs, would have been a useful addition to the project, but organic analysis is expensive and there would be cost implications. PI's confirmed that although this was considered in other catchments budgetary constraints meant that it was not possible. Faecal indicator organisms are another key parameter of agricultural diffuse pollution, and this was only carried out in the Eden and Wensum catchment.

Missing determinants

Metals analysis by ICP-OES and ICP-MS may have been a really useful addition to the project. Suites of dissolved metals could have provided conservative markers that could have been used to identifying changes in pollution sources through the seasons and individual storm events. Conservative markers also become essential when trying to detect subtle changes in nutrient concentrations resulting from agricultural interventions throughout a study period of varying weather conditions.

6.4 The sampling frequency

It is well understood that phosphorus concentrations and chemical forms within streams can change rapidly, in response to sporadic inputs (due to farming activities and rainfall events). It is therefore essential that sampling frequency is appropriate to capture these sporadic pulses in phosphorus. The storm sampling that was carried out within each sub-catchment across the DTCs need to identify the impacts of high-flows and how P concentrations are potentially reduced by mitigation measures. The use of Hach Phosphax P auto-analysers and Nitratax sondes at the outlets of each sub-catchment has produced an extremely useful and novel data set that is required to detect changes in phosphorus dynamics through the seasons and the series of storm events.

The three Demonstration Test Catchments adopted different approaches to water quality sampling frequencies. In the Eden DTC, grab sampling and lab analysis was carried out monthly. This provides a useful resource to ground truth the P data from the Hach auto-analysers, but is too low temporal resolution to provide any useful insights or system understanding (Bowes et al. 2009). This is shown in Figure 5, which demonstrates that only one phosphorus peak was possibly captured, but unfortunately this coincided with a period when the Phosphax was not operating (June 2015). The monthly sampling regime was sustained, and most sites produced data sets that were between 80 and 100 % complete. Additional storm sampling using ISCO water samplers was deployed to provide data from approximately 4 storms per year in 2012 to 2014.

The Wensum carried out weekly grab sampling, which will capture more storms and provide a much better quantification of P load exports and understanding of P pollution sources. This was supported by additional storm sampling, which captured approximately 4-5 storms per year between 2011 and 2014. From the supplied data sets, both the storm samples and weekly grab samples did not seem to include any phosphorus analysis, which is a problem as it means that the quality of the Phosphax data cannot be fully verified.

The Avon DTC produced daily samples and lab data, which provides a useful, research-quality data set. Routine storm sampling at this DTC was only undertaken in Phase 1. The daily samples were taken at noon, and include weekends, which implies that they were taken using an ISCO water sampler. It is unclear how often the samples were collected and analysed, which means that there could be issues of accuracy due to the instability of certain nutrient species, such as SRP, ammonium and nitrite. The daily data from Brixton Deverill is relatively continuous, with >80% of days providing data (Lloyd et al. 2016). At daily sampling frequency, almost all peaks were captured to some extent throughout the monitoring period (Figure 7), but it is interesting to note that even at this high sampling frequency, only one peak was sampled at near to its maximum concentration (in Feb 2013). This highlights the strength of using automatic phosphorus auto analysers when investigating diffuse pollution in small, hydrologically-responsive catchments.

The Prior's Farm and Cool's Cottage monitoring sites only provided nutrient data on 70 and 60% of days. Data were near-continuous from early and mid-2013, but data is sparse prior to this. The Ebble site only produced P and N data on 40 to 50 % of days. Gaps in these datasets may be due to gaps in funding and samples failing quality control checks.

Ecological monitoring began in the Avon/Tamar and Eden DTCs during spring 2011 (2 years before most interventions were established- meeting the minimum pre-intervention time required). Annual macrophyte surveys were undertaken. Macroinvertebrates and diatoms were sampled up to 3 times per year.

A summary of the parameters measured at each sub-catchment, sampling frequency and percentage completeness is given in Supplementary material 3.

6.5 The presence of non-agricultural inputs

The study catchments and monitoring sites have been carefully selected to avoid non-agricultural inputs, based on map surveys. The data generated within the project provides a second opportunity to confirm whether other sources are present.

6.5.1 Detecting point source sewage inputs

One simple technique is to examine the relationship between P concentration and flow, as developed in Defra WQ0223, "Developing a field tool kit for ecological targeting of agricultural diffuse pollution mitigation measures" (Bowes et al. 2014; Naden et al. 2015). Diffuse inputs are largely mobilised and transported to streams during rainfall events, and so P concentration and load will increase in response to rainfall (and therefore river flow). In contrast, traditional point sources from sewage treatment works, industrial inputs and septic tank misconnections, will be relatively constant and independent of rainfall. Therefore, rivers dominated by point inputs tend to exhibit a dilution pattern with highest concentrations at low flow. This principal allows the P concentration and flow relationship of a river to be used to quantify loadings from these constant (point) and rain-related (diffuse agricultural) inputs, using Load Apportionment Modelling (LAM) (Bowes et al. 2010). Many of the data sets produced within the Demonstration Test Catchment project would be suitable for this model application.

The P concentration / flow relationships of a few sites were investigated, to determine if (a) there were other non-diffuse inputs, and (b) if these relationships had changed from the first two years and the last two years of the project.

The phosphorus – flow relationships in the Pow Beck outlet (based on 30 minute Phosphax data) shows a strong positive correlation, with both TP and SRP concentrations increasing with increasing flow (Figure 8). There are some very high TP concentrations at very low flow, and the highest TRP concentrations of up to 1200 µg l^{-1} actually occur at flows below 0.07 m^3 l^{-1}. This indicates that there are some intermittent, non-rain-related inputs, but the vast majority of data points are at lowest P concentration at lowest flow. This indicates that this catchment is diffuse dominated.

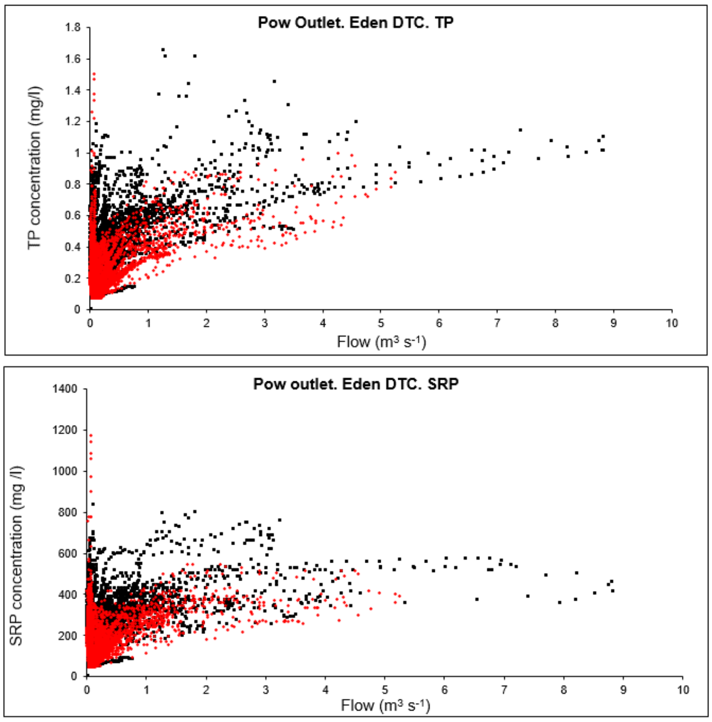

Figure 8: Total Phosphorus (TP) and Total Reactive Phosphorus (TRP) concentration – flow relationships for the Pow Beck outlet, Eden DTC. Based on 30 minute Phosphax data. Black symbols = 2012, Red symbols = 2013.

The River Sem at Cool's Cottage exhibits the pattern of a diffuse dominated catchment, in terms of phosphorus loads (Figure 9). P concentrations are maintained between 0.15 and 0.25 mg TP l^{-1} as flow increases, meaning that an increasing mass of diffuse P is mobilised and transported to the river monitoring point as flow increases. However, the highest P concentrations occur at low flows, which could either indicate that there are large diffuse inputs occurring in response to certain small storm events (possibly after a long dry period), or that there are some non-rain-related intermittent point inputs. Further investigation of the timing of these data points, alongside other determinants, could help to increase our understanding of nutrient source apportionment and dynamics in this sub-catchment.

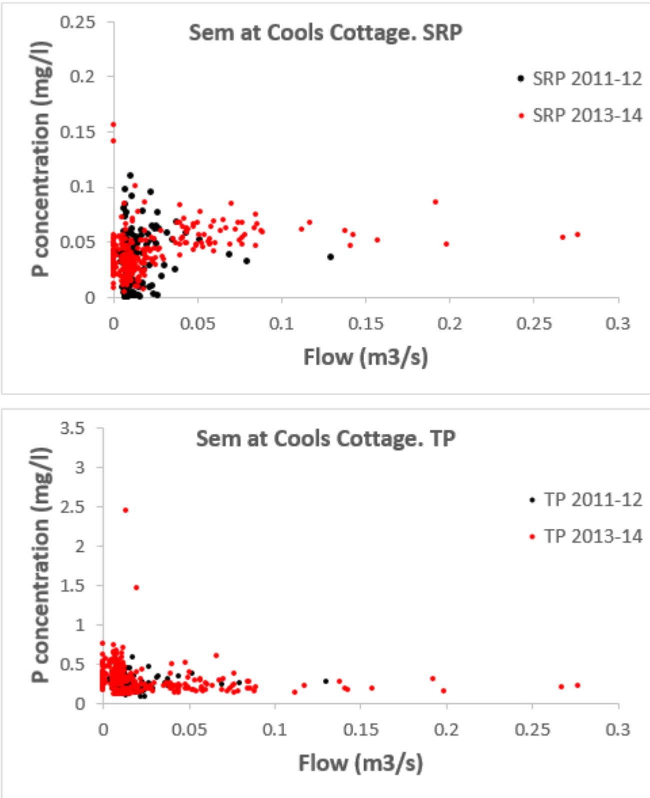

Figure 9: Soluble Reactive Phosphorus and Total Phosphorus concentration – flow relationships for the River Sem at Cool's Cottage, Hampshire Avon DTC. Based on daily laboratory data. Black symbols = 2011-12, Red symbols = 2013 – 14.

The River Sem at Prior's Farm has a very different P concentration flow relationship (Figure 10). Most of the highest SRP and TP concentrations occur at very low flows, indicating that there is a relatively persistent input of P into the river that is independent of rainfall. This suggests that there could be some significant inputs from septic tank misconnections or farming practises that were producing very regular and long-sustained P inputs during low-flow periods. At low flow, SRP concentrations at Prior's Farm are an order of magnitude higher than at the adjacent catchment at Cool's Cottage, which never exceeds 0.16 mg SRP l^{-1}. The Prior's Farm site was monitored during the Defra PARIS project, and had an average boron concentration of 110 µg/l (range 40 – 255 µg l-

[1]), which is very high for such a small stream. (In contrast, the Pow Beck had an average boron concentration of 24 µg l^{-1} within this project). As boron is a constituent of household detergents, this suggests that the high P loadings are caused by sewage inputs, which may prevent this DTC project from detecting the impacts of their interventions.

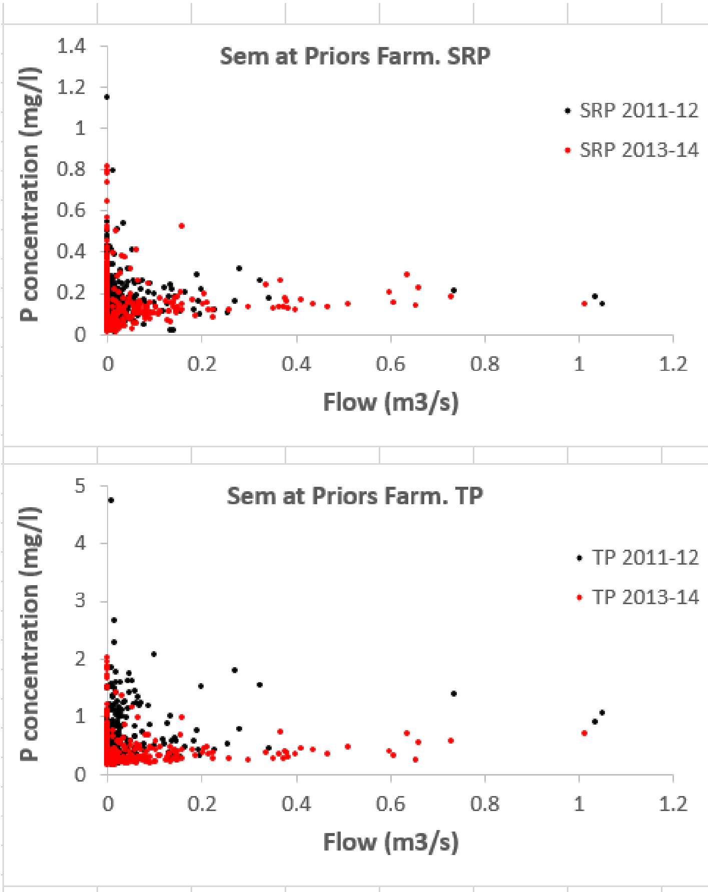

Figure 10: Soluble Reactive Phosphorus and Total Phosphorus concentration – flow relationships for the River Sem at Prior's Farm, Hampshire Avon DTC. Based on daily laboratory data. Black symbols = 2011-12, Red symbols = 2013 – 14.

It is sometimes difficult to identify improvements in a water quality timeseries following an intervention, due to variations in the river hydrology from year to year. This simple approach could be useful to identify any reductions in nutrient concentrations, corrected to river flow. For instance, Figure 8 shows that the P concentration / flow relationship in 2012 and 2013 is similar, but there does seem to be higher TP and SRP concentrations relating to storms between 0.3 and 3 m^{3} s^{-1} (with TP concentrations regularly >1 mg l^{-1}), whereas in 2013, TP concentrations within this flow range never exceeded 0.9 µg l^{-1}. More investigation would be needed (especially as 2012 was particularly atypical in terms of hydrology with a winter drought and summer floods), but this could indicate some success from the DTC interventions.

7 Modelling methodologies

This section provides an evaluation of the models developed and applied in DTC projects. This part of the review is not restricted to PICO compliant documents and includes other relevant sources identified by the reviewers (e.g. model reviews and case studies). This wider review of DTC modelling was needed as not all modelling focused on the assessment of DTC intervention effectiveness.

7.1 Named models

As part of the DTC programme a number of “off-the-shelf” models were used. Of these a number of studies were undertaken using widely-applied physics-based models that describe catchment diffuse pollution mechanisms and include dynamic representation of nutrient and sediment transport in rainfall events. SWAT (applied by UEA and University of Newcastle/Lancaster in Wensum and Eden respectively) (e.g. A62; A32) and INCA-P (applied by University of Oxford in Avon) (A67) represent catchment hydrology and nutrient dynamics in an integrated manner using sub-catchment units. The Wensum has also been modelled using INCA-P (Whitehead et al., 2015) but not as part of DTC activity. In the case of SWAT spatial sub-catchment units are delineated as unique combinations of soil land use and topography which can make for definition of end-members that can be a useful option of a basis for transferability. CRAFT is also a process based model representing hydrology and diffuse pollution and has been newly developed by University of Newcastle and applied in the Eden (A3; A2). In some models the representation of nutrient pollution transfer to rivers requires coupling with a rainfall-runoff model. The INCA model has been linked with HYPE in this respect, and HYPE has been used within DTC programme.

Applications of the PIT and PSYCHIC models have been undertaken to provide stronger emphasis on representing integrated outcomes of diffuse pollution. These typically operate on regular grids. The FARMSCOPER tool uses a number of underpinning models including PSYCHIC to summarise response diffuse pollution losses across farm types at a national level. PSYCHIC and PIT have been applied in the Eden by University of Durham (A24). FARMSCOPER has been applied by Rothamsted (A69) and UEA in the Avon and Wensum respectively.

Other tools have been developed and applied to assess risk and connectivity, and of these SCIMAP (University of Durham) is a risk assessment tool applied in the Eden (A24) for DTC but also to date applied in a number of other catchments. Similarly the DBM model has been applied by Lancaster University across all DTCs (A48). It constitutes a databased mechanistic driven approach (structurally site specific) accounting for local catchment properties to relate rainfall to aggregated P loads.

The use of modelling tools in combination for improving process understanding, for exploring sensitivity to land use and land management and for quantifying effects of interventions has been described in detail by DTC investigators (AC44). Reference should be made to this work, although it should be noted that the report did not include discussion of INCA-P, PIT or HYPE, and only makes passing mention of PSYCHIC.

The model applications made for DTC focused on hydrological and water quality response. We did not find reporting of extension of results to address impact of interventions on aquatic ecology. Whitehead et al. (2014a) mention that INCA-P was calibrated using plant growth and uptake from the water column, although biomass is not reported in the results.

7.2 Other modelling tools developed in DTC

A number of other analysis tools are described as "models" in DTC documentation. A brief narrative of these is provided below. The tools were developed for site-specific analysis. They represent transferable concepts, rather than model codes, but if their detailed data requirements are met they provide powerful approaches for a wider understanding of nutrient pollution across a range of environments.

A substantial body of work in the Eden has focused on empirical modelling phosphorus transfer and the importance of storm event control. These have worked in close concert with research using CRAFT. A site-specific empirical model is presented (A46), relating rainfall intensity to P load, which is then used to look at future trends in event rainfall intensity. A model developed by Hollaway et al in source A31 was used for quality control of continuous monitoring data with the objective of identifying better rating curves for improved estimate of phosphorus loads.

High temporal resolution monitoring techniques have been used in conjunction with molecular and compound specific isotope analysis to develop a suite of Bayesian mixing models for the purposes of apportioning sources of organic matter and stream sediments (A21; A18; 2015b). Other sediment and phosphorus tracing techniques have been reported. A means of estimating geochemical properties from SPM data from autosamplers in the Wensum is reported in A20. Likewise in the Wensum and Avon, empirical sediment fingerprinting models have assisted the process of sediment source apportionment (A10; AC44). Although it is acknowledged here that source apportionment modelling yields highly relevant information the results are not included in our evidence review as the impacts on water quality are not explicitly reported.

Models with a specific focus on the sub-surface have also been applied during the DTC. The Nitrate Time Bomb approach is suited to large scale assessments but has been applied in the Eden (AC44). A set of hydrogeological site-specific conceptual qualitative models support and inform management in the Wensum (A16).

7.3 Summary of capability of model packages

The off-the-shelf process-based catchment-scale models used in the DTC research (e.g. SWAT and INCA-P) are potentially very useful for extrapolating findings to other basins. They are invaluable in quantifying climate variability, the importance of storm events and the influence of confounding factors. These are all aspects that are difficult to cover using other methods. There is considerable knowledge contained in other applications of these models in the UK. A search of journal publications reveals the SWAT model to have been applied in five case studies: West Wales WFD district (Holder et al., 2019), Axe (Glavan et al., 2011), Bedfordshire catchments (Kanna et al., 2007), Great Ouse (Grizzetti et al., 2004) and Yorkshire Ouse (Boorman 2003; Bouraoui et al., 2002). Applications of the INCA model are reported in journal publications for 13 case studies: Trent (Bussi and Whitehead 2020), Herefordshore Wye (Bussi et al., 2018), Conwy (Bussi et al., 2017), Thames (numerous applications including many on sub-catchments such as the Kennet and Lambourn of which the latest is reported in Bussi et al. (2017)), Weaver-Dane (Hankin et al., 2016), Lugg (Lazar et al., 2010; Dean et al., 2009), Twyi, Tamar, Tame and Tweed (Whitehead et al., 2009), Plynlimon and Great Ouse (Whitehead et al., 2014b) and Tillingbourne (whitehead et al., 2002). However, undertaking model applications in other basins is heavily resource dependent. It is recommended that knowledge is pooled from the other existing studies with due consideration to model skill. Criteria of expectation for model performance (e.g. Moriasi et al., 2015) should be met. Report AC44 outlines how these models can be summarised (meta-modelling) and then used in conjunction with nationwide tools such as Farmscoper and Catchment Matcher.

Despite their physical basis, representing a wide range of interventions in process models is not straightforward. Catchment model applications often lack the spatial detail to represent small-scale interventions. Instead this is often captured through pooling of empirical observations, field scale models and expert judgement (e.g. Cuttle et al., 2006). Furthermore models are typically used to represent future scenarios within a catchment system in steady state. This is particularly the case in representing the changes to soil nutrient cycling arising from change in management practice. Development of some interventions are more readily represented (e.g. establishment of riparian canopies) and they have greater capability to represent gradual development of future climate. These challenges are reflected in the modelling undertaken in the DTC projects.

8 General Insights regarding farmer engagement methodology

With regard to engagement, the DTC programme aimed to:

- Investigate approaches to maintain farm productivity whilst delivering improved water quality outcomes; which required negotiating with farmers to implement interventions; and
- Create networks for collaboration and knowledge exchange between researchers, farmers and other stakeholders

Good practice in the academic and practice literature (see Burbi et al., 2016; EC 2018, OECD 2015, Sumane, 2018, WATERLife, 2016) suggests that the following steps are important to have successful engagement with farmers and other stakeholders about agri-environmental interventions. Some of the DTC documents were also reviews of good practice in engagement (A44, AC4).

- Undertake stakeholder analysis for all those with a stake or affected by the proposed activities in order to understand the existing responsibilities, motivations and interactions within these groups including 'new' and 'under-represented' groups
- Ensure all sectors involved in the issue are involved, and the shared responsibility for a problem and the solution is recognised
- Raise awareness of the issues to engage the stakeholders and how it is linking to their concerns
- Stress the benefits of change, in language that is relevant to the different actors, paying attention to what matters to whom
- Use a variety of approaches, and tailor or adapt to local circumstances
- Group learning and demonstration is important building on local knowledge; allow two-way learning between scientific and other actors
- Avoid consultation fatigue by working others on common messages and approaches
- Celebrate and share results so that benefits of changes are understood
- Support long-term relationships as environmental outcomes might take time to show up
- Pay attention to challenges such as consultation fatigue; absence of leadership; lack of resources; consultation capture; resistance to change; information asymmetry
- Use feedback from engagement in wider institutional and governance processes with the brief of DTC

These principles were used to develop the extraction protocol reported in Section 3.2. There was also a great deal of evidence in SPICE relevant, but not compliant, documents evaluated for this project. Therefore, this section illustrates the additional evidence that the DTC documents provided on how they implemented good practice in engagement. Where relevant, SPICE documents have also been included, so that the entire evidence base was summarised. This section complements material answering the Primary Social Science Question "How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?"(see Section 5.1); SS1 on DTC engagement processes (see Section 5.2); and SS2 on provision of information on non-participants (see Section 5.3). The objective of this section is to capture the wider insights from the DTC evidence base on whether the methodologies used in farmer engagement followed good practice; and to capture the additional insights from SPICE compliant documents for the overall recommendations.

8.1 Stakeholder analysis

Many documents stated who they engaged with but there was very little explicit discussion of how stakeholders were selected and whether the interests, motivations, interactions and potential conflicts between these stakeholders were mapped or recorded. W33 and O15 were some of the exceptions that provides this information – O15 identified communities; farmers; investor organisations and planners/politicians as the four main audiences for the DTC information. There was a "low base in terms of community engagement with catchment management issues" in the Wensum in 2010 for example (O14).

In some ways, the stakeholders of relevance to the DTC were self-explanatory, but as observed in the SPICE analysis, it is unclear how many of these were engaged. Information on the overall farming population was given in some SPICE relevant documents (e.g. O15 within the three Eden DTC sub catchments there were over 100 landowners and over 60 active commercial farm businesses), but the proportion of these who were engaged and in which ways was not easily established. There was limited reference to land tenure in many of the documents, yet as O15 observed, implementation of measures needed to take account of influences such as the extent of common grazings in upland catchments. AC6b noted that further DTC research should provide information on why some farmers do not engage; how much time should be spent approaching disengaged farmers and the benefits are of getting them involved but this does not seem to have been addressed in the documents reviewed.

It might be useful to consider 'under-represented' categories e.g. gender, age and potentially other intersectional issues when engaging farmers. The OECD 2015 category of new stakeholders was also pertinent, given the importance of water companies and property development in some catchments; and the policy importance of investing in natural capital. O15 reported on a two-day workshop involving OFFWAT and water utilities. In light of comments about diffuse pollution from roads, rural septic tanks and urban development, it was interesting that the DTC did not engage local authorities in their research networks. Riverkeepers were also suggested as important stakeholders (O13) but it is unclear if they were engaged in the later phases.

Use of 'champions' and knowledge brokers were identified in several documents (e.g. A44, O11) - The Wensum DTC hired a farm liaison officer to help with farmer engagement (O14). AC6b encouraged formal research on farmers who are innovative and early adopters of new ideas/techniques in phase 2. However, it was not clear how this process worked in later reports.

The farmer in Wensum was identified as a champion but this choice was viewed as problematic by some stakeholders as he was atypical of arable farmers in the catchment (W42).

8.2 Cross-sector networks

Building the networks of researchers, agencies, advisors and farmers was important to improve knowledge, pool resources and connect land management and environmental outcomes (A45, O11). Many documents stated the importance of engagement with multiple actors, beyond farmers (A64, W12, AC10, AV9). The establishment of an entity (The Wensum Alliance) (W32, W41, O14) was evidence of the commitment to network development and this Alliance was active throughout the three stages of the DTC project and contained over 180 members. O15 also reported on how the Eden DTC nests into the wider Save the Eden Coalition, which was also a large cross-sectoral partnership.

This wider engagement was important as farmers perceived that diffuse pollution was not only caused by farming, but also by the water industry and the way that roads are managed (W23, W41, AC16, AC31). Furthermore, some documents recognised that non-farming stakeholders might help 'nudge' farmers towards voluntary uptake of measures e.g. AC31 discussed the influence of neighbours and customers. The supply chain was highlighted as an important type of actor in catchment networks (AC15, AC5, AC4) although it was unclear how far down the supply chain the engagement actually went. This is perhaps what was meant by the recommendation to have 'both end' solutions in O11.

Building cross-sector networks was not simple. There was a challenge of the optimal scale, given that networks need to build on 'local expertise' and be context sensitive (A64, W12, O11) leading AC 15 to recommend setting up sub-catchment discussion groups. As discussed further in section 9.5 on advice and section 9.6 on governance, working across sectors required aligning messages and approaches, and/or investing in deliberative spaces to discuss differences (A64, W12).

8.3 Crafting the message

Information provision and raising awareness of the need for action was recognised as an important part of engagement as both monitoring and interventions required the permission of landowners (A45). Some documents were developed to build awareness and encourage participation in co-designing the measures (e.g.W41); and the DTC project was premised on early and active engagement that built ownership of both the DWPA problem, and the potential solutions (O13). The need to tailor the approaches to the local setting was well recognised in the DTC documents (e.g. O13, O11).

Discussions needed to be based on both increasing understanding of DWPA (sources, pathways, treatment trains) and farmers' evaluations of the mitigation measures. Discussion of interventions considered how these interventions fitted with their farming practices and budgets (A35, AC10, AV9). The DTC research contained reflections on how to use tools like FARMSCOPER with farmers and advisors, concluding the tool is useful for discussion and awareness raising but as it can not replicate the details of the specific farm, it cannot be used for bespoke farm planning or 1-to-1 advice (A26; A35, AV9).

8.4 Variety of approaches

As highlighted in SPICE documents (O13, O11) the DTC used a range of approaches to provide passive information e.g. websites and apps, printed materials, press articles as well as activities and events (O15, O14). There was information about digital methods for sharing data on flood risk in the Eden found in A43; on engaging the wider public in catchment management W30 ; and findings from a workshop on how go about utilising data and findings from the DTC platform to influence the messages put across to farmers (AC27). The documents included a 2 minute 30 second animation covering what diffuse pollution is, what are the sources, the pollution pathway, the treatment train and costs to farmer of loss of fertiliser and costs to environment and wider society (AC32), illustrating good practice in using visual methods to engage 'hard-to-reach' audiences. A66 recommended more research on using visualisations (e.g. word clouds) to share messages with stakeholders e.g. advisors and farmers.

Surveys provided good baseline and behavioural data but these were not a sufficient substitute for discussion groups and active citizen science engagement, which also took up time and resources. There were several farm events and visits in the Eden, Avon and Wensum catchments (O14, O15). AC4 concluded that demonstrations were more successful than providing information, echoed in AC31 and W94. The DTC ran demonstration events in the Eden (O15), Avon and Wensum catchments (O11); the Eden Rivers Trust also worked with local further education colleges to demonstration resources (O15). The ability for farmers to see a demonstration of the techniques on conventional commercial farms was very important (O15, AC4).

As AC25 suggested, demonstration was not just about technology - field visits allowed informal knowledge exchange and differences in views to be explored. The approach allowed farmers to express concerns and shared knowledge, reducing information asymmetries (AC5). The 'bus tour' approach used in the Eden in 2012 was useful to not only raise awareness of the different issues and interventions, but to break down barriers between farmers, the Defra family of agencies, and utilities. It is unclear if this approach was repeated in other catchments, or in later phases of the project.

8.5 Group Learning

Several documents noted the importance of local farmer discussion groups to stimulate 'double loop' social learning (A35; E9; W94). Social learning had potential to make environmental stewardship more acceptable amongst farmers (AV11), particularly when farmers felt able to openly discuss issues and solutions. The best way to achieve that was in small groups that were farmer dominated (AC4) and ran over a period (W94), which required resourcing and an external facilitator (AC15). This approach was different from the recommendation to set up cross-sector networks – these are complementary.

Two-way exchange of information over time with scientific results presented or sharing of experience between peers was highly valued (W94, O13, AC15). Recognition of local knowledge held by farmers was recognised from the start of the project (A45, W32, O13, AV9, W41, O14). There were examples of the DTC researchers using local knowledge to identify both the issues and the particularities of how an intervention might fit in the landscape and farming system (e.g. O11). This knowledge was used to improve the DTC modelling capacity (e.g. on field drainage practices in A70); improve the application of the Farmscoper tool (A26), improve mass balance calculations (AC10) and generate scenarios for modelling (A47). O14 discussed recommendations to make data more accessible to other users.

8.6 Working with others

Coordination between organisations was important for good engagement (A44, W32, AC4, O11). Many documents referred to working with Catchment Sensitive Farming (CSF), Catchment Based Approach and other initiatives e.g. Upstream Thinking (e.g. AC9, AC6b). The need to ensure that the DTC was seen as complementary and not competition or duplication to these projects was recognised in A45; W41, O11, O14. AC6b highlighted how information and evidence is needed by CSF Officers and recommended 2-way communication between DTC and CSFOs to share information in future phases, although again it was unclear to what degree this information sharing occurred. W12 found that the CSFOs made a broad range of recommendations that differed between regions, that the CSFO had their own niche compared to other organisations, and were adapting their approaches to the local context as appropriate. The CSFOs were most likely to recommended yard infrastructural changes funded by CSF grants.

O15, for example, stated that the Eden DTC's role was to provide evidence to Saving Eden Catchment Plan (Evidence and Impacts Forum). In the Eden, the engagement was explicitly led by the Eden Rivers Trust (O15), but even here they worked with others (e.g. NFU) when they did not have strong existing networks (O15). O14 reported on how the Wensum DTC moved from DTC-led events, to participating in activities led by others.,. There was an important collaboration between CSF and Avon DTC to look at Phosphorous mitigation measures (AC25) although the outcome of this partnership was not explicitly identified in our research. This coordination is good practice, AC4 recommended using existing mechanisms as farmers are busy. Around 40% of survey respondents recalled hearing about the DTC from other farming events, with highest numbers in the Wensum and lowest in the Eden (AC16).

As shown in AC30 and O11, these wider networks were not only important to gain access to the wider farming community; help with information exchange and share learning about effective interventions; but also as important sources of grant funding for measures being monitored as part of the DTC project. AC30 highlighted the importance of the Upstream Thinking programme that provided £2.2million grants and an average cost saving per farmer of £20k whilst providing 30% less DOC and better summer water levels. There were also references to working with Environmental Virtual Observatory (A43); WQ0106 (A26); WQ202012 (AV4); Defra funded Sustainable Intensification Platform project - May 2014-2017 (AC 26); and ALFA (Adaptive Land use for Flood Alleviation) Interreg III B funded project (O15).

8.7 Celebrate and Share

The DTC aimed to provide accessible and digestible knowledge to users rather than prioritising publications in learned journals (A45). AC15 reported how farmers in discussion groups were keen to learn more about the local research being done in the catchment. The DTC consortia reflected on their findings and took stock of learning, sharing insights with other stakeholders (e.g. W42) although these types of events tended to be attended by stakeholders from research organisations, public bodies and NGOs rather than land managers. They were useful processes but other ways to feedback information to farmers were also required. It was not clear to whom, and how, the DTC newsletters and briefings were distributed. The website and media interactions reported in O11, O14 and O15 will have helped.

A44 (also AC24) highlighted the need for monitoring of progress, evaluation of progress and revision of required actions; and Appendix E notes that the methodology was adaptive and responsive to local enthusiasm (A45, AC5) rather than rigidly experimental. The learning from these adaptive and iterative approaches was not easily captured from the extraction table or summaries of the DTC documents, but there have been some taking stock workshops and conferences (e.g. W42).There was also a ‘mini evaluation’ carries out on the DTC project (AC5) but it was unclear how the findings from this evaluation were shared and with whom.

8.8 Long-term relationships

Several documents drew attention to the need for sustained engagement over a period to time to maintain uptake of voluntary DWPA measures (e.g. A35, AC4). This allowed the development of mutual trust and respect (AV9) and continuity of tailored advice (AC15) that built on shared history and relationships (A66). This was not only about stakeholders. One of the added benefits of the DTC programme is the wealth of expertise contained within the personnel, which would be useful in ongoing knowledge exchange in these catchments (AC6b). Plans for ongoing discussion groups and sharing future monitoring results were not easily detected in the documents reviewed. PIs noted one achievement of the DTC was to develop co-operative working relationships between previously competing consortia, which improved their science and support to policy. They also highlighted how the DTC provided a useful training programme, exposing young researchers to a wide range of stakeholders and real-world science-policy interactions. These researchers now occupy policy and industry positions.

8.9 Challenges

There was very little information extracted on consultation fatigue; absence of leadership; consultation capture or information asymmetry, although these are well documented challenges to engagement.

8.10 Summary

The DTC body of evidence clearly illustrated a good understanding of the practical aspects of engagement and advanced the knowledge in this field academically. The material was strong on the importance of: planning engagement to build up trust and social networks, moving move from information provision to active involvement, and the embedding farmer discussions in wider existing farmer and stakeholder networks. The gaps in the evidence corresponded to how stakeholder analysis was undertaken, and decisions made about how and where to focus effort. The dynamics of stakeholder interactions within discussion groups, particularly regarding 'capture' by dominant interests or effects of information asymmetry, could have had more attention.

9 General Insights regarding uptake of measures by farmers

With regard to uptake, the DTC programme aimed to:

Investigate approaches to maintain farm productivity whilst delivering improved water quality outcomes; which required negotiating with farmers to implement interventions; and

AC35 showed that the DTC project was structured on the Theory of Change in order to:

1. Understand the nature of the problem;
2. Design improved interventions;
3. Understand barriers to uptake; and
4. Improve research and monitoring capacity.

In the literature, Mills et al. (2017) summarize and apply several conceptual models of behaviour, the Theory of Planned Behaviour and the Value-Belief-Norm theory, to the adoption of pro-environmental behaviours by farmers. They summarize the factors affecting farmers' environmental decision making (and hence uptake of interventions) by the following figure.

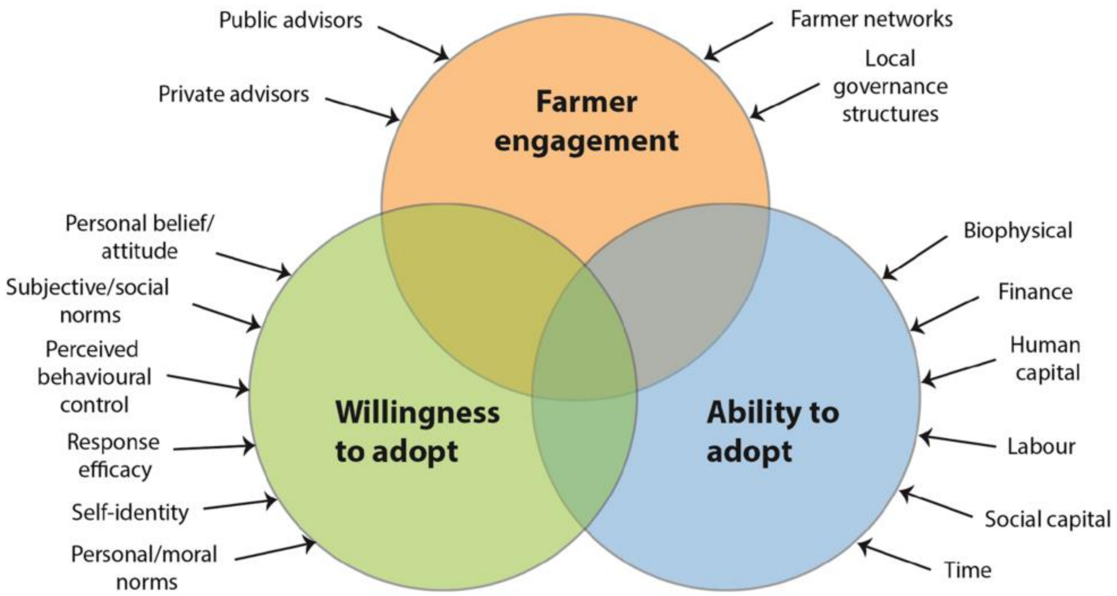

Figure 11: Factors influencing farmer environmental decision-making.

This section builds on the DTC planning (AC35) and Mills et al. (2017) to consider how the DTC documents addressed:

- To what extent farmers understood and felt responsible for DWPA (underlying their willingness to adopt interventions)
- To what extent personal attributes affected their decision to adopt interventions
- To what extent farmers felt able to adopt interventions

- Influence of Advice on farmer choices
- How farmers decided to act at the individual level, given their ability, willingness and level of engagement
- Impact of Governance and wider catchment processes (building on farmers' engagement in Figure above) on uptake of interventions

These principles were used to develop the extraction protocol reported in Section 3.2. There was also a great deal of evidence in SPICE relevant, but not compliant, documents evaluated for this project. Therefore, this section illustrates the additional evidence that the DTC documents provided on how the DTC implemented good practice in promoting uptake. Where relevant, SPICE documents have also been included, so that the entire evidence base was summarised. This section complements material answering the Primary Social Science Question "How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?"(see Section 5.1); SS3 on behaviour factors (see Section 5.4) and SS5 on costs and benefits (see Section 5.6). The objective of this section is to capture the wider insights from the DTC evidence base on whether the methodologies used in understanding farmer uptake of DWPA measures followed good practice; and to capture the additional insights from SPICE compliant documents for the overall recommendations.

9.1 Understanding of the problems

The overall premise of the DTC was to raise awareness of DWPA to encourage farmers to voluntarily take up mitigation measures. A11 reported that monitoring and reporting data on the state of catchments improves engagement (see also W94, W32). W41 provided a good example of DTC attempts to provide farmers with Information on current state of water quality in the catchment, sources of diffuse pollution, how diffuse pollution is related to food production, and how it harms the aquatic ecosystem. There was some attempt to practice 'citizen science' in catchments (e.g. farmer diaries in Eden (AC6b) and water quality testing in the Wensum, (AC30, O14) to build awareness. O15 noted that the increased access to smartphones during the project meant that farmers could capture storm and other events more easily. A35 stressed that double loop learning required farmers to be involved in evaluating diffuse pollution measures for themselves.

As noted in SS3, there were some changes in understanding (see also AC28). However, as AC15 notes, whilst 95% of the respondents agreed diffuse water pollution from agriculture (DWPA) was an important environmental issue and 82% agreed there was a link between farm activities and DWP, there remained confusion over scale and severity, source and probability of interventions making a difference. A70 also found that farmers had believed land drainage to be better than it was. AC15-16 illustrated the level of environmental improvement perceived by farmers as a results of putting in measures. As well as considerable heterogeneity across the catchments; it appeared that the measure with most improvement was cultivating compacted tillage soils, with more mixed results for cover crops or excluding livestock from watercourses. AC21 also found that farmers were more aware of wildlife benefits from DWPA mitigation measures, than any impact on diffuse pollution.

W23 highlighted that many farmers did want to see improved water quality and accepted that farm activities are polluters. However, they argued that DWP does not only arise from agriculture, but wastewater treatment and road run-off were perceived as important contributors (see also A67 and A11). AC15-16 also noted differences in perceptions of agriculture's contribution to DWP with public highways being seen by survey respondents as the highest contributor and farming only a medium contributor.

A66 argued that the link between DWPA and cost savings needed more work to be convincing, echoing E9 which recommended that the benefits of mitigation to farm business needs to be conveyed to farmers. AC25 advised that these links are better achieved when there are multiple benefits (e.g. reduce liver fluke via riparian fencing) and these are pointed out to farmers. AC10 suggested that further work is required to help characterise the damage costs of water pollution in order that economic assessments for cost-benefit are more reliable. This affected advice (see below). A66 recognised that farmers need advice on new management techniques and infrastructure adjustments (e.g. cover crops, subsoiling, bio beds) and this advice needs to include costs, impacts on whole farm plans, and benefits for the business.

These findings explained why AV9 (and others e.g. A35) highlighted that not only do farmers need to understand DWPA but there is also a need for scientists to understand farming systems, farming practices and how the mitigation measures impact on profitability. This was an objective of the DTC but seems to have remained challenging for the project teams. Feedback from PIs notes that the DTC aimed to develop open discussion and assessment of evidence by multiple users, not just farmers.

9.2 Personal attributes

There were references to co-design of measures being undertaken with farmers in the first phase of the project (A45) which ensured the measures are adapted to suit the individual circumstances. The DTC project recognised the importance of understanding personal preferences, attributes and values when engaging farmers - DTC baseline surveys have collected a range of data in behaviour and attitudes (AC10). This information was also intended to help understand when and why individuals might resist participation in agri-environmental schemes (e.g. O13 suggests 13 types of non-participants). The DTC documents highlighted some ingrained resistance to change (O13), where negative attitude towards some measures and taking action to protect environment were key barriers to the uptake of measures. Barriers expressed by farmer quotes included: 'Waste of time', 'Not convinced it works' and 'Damages drains' etc (AV11). This was despite the survey sample being biased towards farmers who were engaged with DWPA and had existing environmental concerns (AC15-16). This is contrary to the findings in AC21, where farmers were more concerned with economic benefits of measures and less concerned about environmental impacts (see also A70). Findings reported in AV8 show that the majority of farmers indicated that they would never "establish permanent woodlands", or "allow drainage systems to deteriorate". The majority of farmers also indicated that they were unlikely to carry out "arable reversion to low input extensive grazing", "grow biomass crops", or "irrigate crops to achieve maximum yield", as these would all reduce food production and/or increase their costs. Therefore, as noted in several document, increasing uptake of DWPA requires a shift in identities and behavioural beliefs including ideas about farmers' roles in society including a shift from productivist to multi-functionalist (A 35, AC15, A9, AC31).

9.3 Ability to Act

AC10 highlighted the diversity within the DTC farming population – "Farm businesses and farmers are not alike; leaving aside both the physical differences in farm type and location and the consequent cultural differences, there is a huge range of capabilities, resources and infrastructure. This results in differing motivators and responses to requirements to change to address the pollution problem" (p6). AV3 also highlighted the importance of the social and economic context for measure uptake. W7 summarised these contextual issues using a spider diagram of multiple barriers to uptake (Capability, Social/cultural, Economics, Institutional, Environment, Demographics, Automatic Motivation, Reflective Motivation), providing information about agreement or strength of feeling about barriers. These documents illustrated that whilst economic barriers are important (e.g. W23 reports that some farmers want easier access to HLS funding to help pay for mitigation measures), constraints might not be solely about finance but also but succession and choices about the long term future of the farm (O15, AC16). Indeed, O13 found that reduced profit margins might stimulate farmers to seek advice on changing farming practices (including taking up DWPA measures). Furthermore, it was important to understand the impact of collective or individualistic approaches within a catchments – for example O15 noted the diversity of land ownership in the Eden (owner occupied farms, institutional and private estates with a mixture of tenants and tenancy agreements, and a large number of actively grazed commons which have an implication for downstream land use).

9.4 Decision to Act

The documents suggested that measures requiring land use change were less likely than measures that address farm infrastructure, and measures most likely to be adopted in future were those that decreased costs and demonstrate compliance with regulations (W11, O11, AC10, O13). For example, measures with the highest current or future planned uptake across Avon, Eden and Wensum were all concerned with fertiliser or manure management and form part of the cross-compliance requirements for the receipt of CAP Pillar 1 Single Farm Payment (AC10). Also, sediment traps were more popular in areas that were already wet and unproductive, where they could improve wildlife habitat, aesthetics and manage flood risks (E9). A47 found that farmers tended to be more likely to uptake measures that are already established in area/farming systems (e.g. soil conservation measures). There were no obvious differences in the uptake of measures that address source management, pathway interception or receptor protection (AC10). AC15 recommended having tailored interventions, using multi-step process that means starting with simple interventions to build trust and building up to more complex measures that might require land use change. There was limited information on why individual farmers did adopt DWPA mitigation measures, although some documents (e.g. AC5) recommended later stages of the DTC should look at tweaking existing measures and designing measures focussed on behavioural change variables. The relatively low uptake rates reported in section 5.4 echo those found by Burbi et al (2016) where only half the farmers adopted changes in farm management: the main obstacles to innovation were limited financial capital, lack of trust in government action and confusion over the effectiveness of farm advice on mitigation. Although the DTC objective was focussed on farm productivity, our review did not pick up much evidence of quantified impacts on farm productivity or benefits realised through adopting mitigation measures.

9.5 Importance of Advice

Many documents reflected on the importance of advice as part of the connection between engaging farmers and seeing farmers taking up and sustaining DWPA measures. It was clear from the material that trusted individuals providing tailored information were useful to supplement and extend the peer learning and collaborative relationships developed through the DTC (AC4, O11). Discussions needed to be complemented by 1-to-1 advice to get from interest to actual uptake (A35, AC15, AV11). Advisors could illustrate multiple benefits and potential cost-savings for farmers (AC4, AV11). It was important to understand the connection between message and messenger, which means that the DTC needed to consider not only which mitigation measures, but also which organisations and individuals were best placed to deliver the advice (A64, W12, AC4, E9). For example, A66 recommends environmental organisations promote themselves to farmers by emphasising their local knowledge and evidence to encourage uptake of advice.

Whilst most farmer respondents used an advisor and felt well informed (O13), DTC documents illustrated that there were skills gaps in the provision of DWPA advice to farmers on the benefits of implementing measures or how the measures might fit with their farming system (A35, A65, A64, AC15, AV11, E9, W94) that need addressing. These seemed to vary geographically (A64, W12). AV11 recommended that multiple agencies need to communicate the same messages to reduce farmer apathy or confusion. For example, more information was needed on soil management and the benefits of cover crops in the Wensum; and the lack of continuity of advisors was noted (W94). There were areas that were farmers sought advice on issues not covered by the DTC objectives slug control, rainwater harvesting, and accessing agri-environmental grants (AC15) and these were passed onto CSF and other advisory networks. CSF officers felt that there was often an overload of information available to them and they needed current scientific data synthesised into key messages (AC6b).

9.6 Governance and fit with policy

Although many documents contextualised the work of the DTC within the need to meet Water Framework Directive targets, the DTC engagement processes were not part of a statutory process like River Basin Management Planning. Informal mechanisms are generally more open and flexible which can improve the quality of deliberation and attract more participants. However, they often lack a pathway from these deliberations to final decisions and require follow-up work to turn concerns into final decisions (OECD, 2015). AC24 and A44 also noted that adaptive management at a catchment scale was also likely to require institutional change. It was unclear from the documents to what extent DTC researchers were able to use the feedback from engaging farmers in wider institutional and governance change processes, beyond presenting their findings to steering groups compromised of Defra and relevant agencies. However, there were several documents that do consider how the research had implications for wider agri-environmental and water quality policies. PI feedback explained how the researchers had regular discussions with policy teams, providing potential pathways for change.

The need for the DTC approach to fit with wider policy mechanisms such as regulatory guidance, agri-environmental schemes and general advice was mentioned in A64 and W23, although there was no discussion of whether or how the DTC evidence might be used to adapt and improve these mechanisms. Often newsletters linked DTC activities to wider policy processes such as measures for WFD or the importance of SAC and SSSI designations (e.g. W41). AC5 noted that approaches needed to adopt more realistic timescales regarding when environmental benefits might accrue. The DTC documents drew attention to the need to show farmers what WFD meant for their sector (O15). There was also mention of CAP payments and the coordination of agri-environment payments to support WFD outcomes and fund mitigation measures (e.g. AC23, O15). Some documents highlighted problems with the current payment systems (A35 , AC10 O13). It was not only the amount of payments that mattered; given the heterogeneity of farmers and farming systems, O15 and O13 recommended that voluntary and incentive type mechanisms must be given greater flexibility to take account of these varied circumstances.

Indeed, AC10 summarised how the Upstream Thinking initiative in the Tamar DTC “illustrates the level of capital investment (approaching £1 million in an area of ~25 km2) required to support significant uptake of farm infrastructure measures”. Although increased payments might help uptake particularly when recommending expensive machinery or farm infrastructure (W11, AV11, W94, W23, E9), these documents noted that financial support needed to be combined with changes in regulation (W11) or more demonstration and advice (AV11, AC15). The need to resource the human capital as well as built or technological capital required for catchment management of DWPA, via coordinators, demonstrators, advisors or facilitators for farming discussion groups, was a recurrent theme (O13, AC15, AC24, W94). This funding needed to be long-term to attract and secure skilled staff who could build relationships of trust with farmers (W94, AC24). However, tweaks to the regulatory approach were also highlighted – farmer participants in DTC workshops argued that a ‘big stick’ is needed when other approaches have been exhausted (O13) in order to address 'free riding' by other farmers (AC15). These farmers argued that motivation to adopt mitigation measures would increase if other farmers were unable to continue to pollute without sanctions.

9.7 Summary

The DTC evidence illustrated that the researchers understood theories of behavioural change and invested time in understanding the many dimensions of farmer behaviour in the catchments. Recognising that individual decisions are influenced by others, including trusted advisors, was also in keeping with current social science perspectives. However, there was more evidence around how farmers understood DWPA and their existing personal circumstances than evidence about why and who made a decision to take on new mitigation measures. Likewise, connecting individual farmer's identities to wider institutional drivers (both policy and the supply chain) was important. However, the evidence was limited about to what extent these linkages influenced the outcomes, and with the liaison with CaBa and CSF processes in the later phases of the DTC project.

10 Methodology for generalising physical science findings

10.1 Generalising findings

There are many existing methods for categorising land in GB, based on environmental, physical and social variables. Such methods have the potential to aid in assessing the representativeness or the DTCs and the probable transferability of specific environmental interventions. Three methods of particular interest are the existing WFD catchment typology, the Sustainable Intensification Research Platform Landscape Typology tool and DTC Catchment Matcher tool (Lovett et al., 2018).

The EU WFD system A catchment typology is based on a comparatively limited subset of categorical factors (ecoregion, altitude, geology, area). The various factorial combinations of these categories then drives the typology. In England and Wales there are around 24 possible catchment types (all GB catchments falling within the same ecoregion). However, because these typologies are designed at European scale, much of England and Wales is dominated by comparatively few types (Fig. 11), so using these data to make assessments on the similarity of the DTCs to other catchments would be inadvisable without substantial additional data on other factors affecting catchment characteristics, although they do at least show that DTCs are not highly exceptional in their typology (Fig. 11).

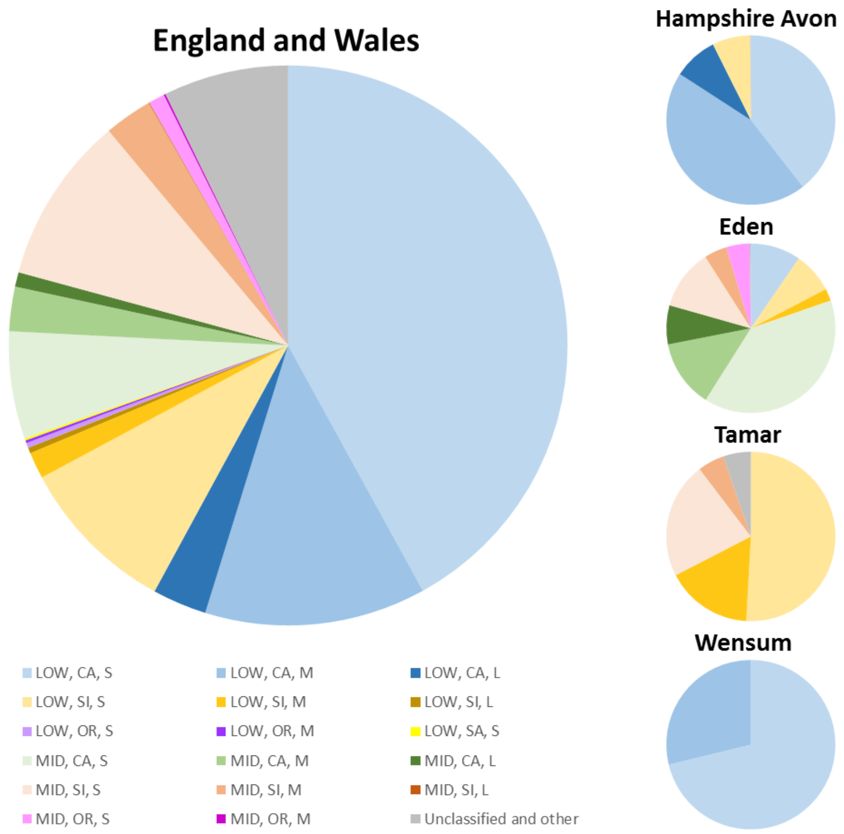

Figure 12: Proportion by area of WFD System A catchment types in England and Wales, and in each of the four DTCs. Catchments are typed by combinations of altitude (LOW = lowland, <200m, MID = moderate, 200-800m), geology (CA = calcareous, SI = siliceous, OR = organic) and catchment size (S = small, 10 - 100 km^2, M= medium, 100 - 1000 km^2, L = large 1000 - 10,000 km^2).

The Sustainable Intensification Research Platform (SIP) Landscape Typology Tool was funded by Defra and Welsh Government, with the aim of exploring risks and opportunities for sustainable intensification of agriculture (i.e. increasing farm output whilst simultaneously maintaining or increasing other environmental and social outcomes). Whilst the SIP tool explicitly includes goals around water quality and flood management it has several limitations in the context of assessing the representativeness of the DTCs. First of these is a matter of spatial scale, as the SIP tool runs at 10km x 10km grid resolution rather than discrete catchments, so making direct comparisons between catchments becomes challenging. Secondly, the aim of the SIP tool is to identify areas of land that provide greatest opportunity for specific sustainable intensification goals at national to regional scales. Whist it is possible to identify areas with a similar level of suitability for a given objective to those cells within the DTCs, the factors driving this suitability are complex (such that areas may be equally suitable for different reasons) and the ability of the SIP tool to alter the weightings of multiple goals makes the maps difficult to interpret in terms of assessing transferability.

The Catchment Matcher tool was funded under the DTC programme specifically to address the challenge of extrapolating evaluations in trial areas to regional or national scales (Lovett et al.

2018). The tool assesses catchment typology and similarity on the basis of over 20 variables, broadly divisible into physical environmental variables (e.g. soils, topography), land cover and use (e.g. forest extent, crop types) and policy designations (e.g. nitrate vulnerable zones, designation as AONB). The tool works at the scale of WFD Cycle 2 Operational Catchments, which corresponds well to the DTCs and the spatial resolution of the input datasets. The tool uses principal components analysis to reduce dimensionality amongst variables across the three broad groupings. The component scores from these analysis then form the basis of clustering algorithms to group catchments into clusters of similar type and of the assessments of catchment similarity based on their proximity in ordination space. This latter use allows assessment of pairwise similarity of each DTC to each other catchment and of average pairwise similarity to all other catchments (a measure of general 'representativeness'). The tool can be used to map the similarity of other catchments to each DTC (Fig. 12) or to assess the general degree of representativeness of the DTCs (i.e. proportion of green vs yellow catchments in Fig. 12). Results for these assessments are presented in Lovett et al. (2018), with the general conclusion that the DTCs are representative of a wide range of catchments in England and Wales with the exception of highly urbanised catchments in the South East and upland-dominated catchments in Wales (as demonstrated by the fact that these areas have high distance scores across all plots in Fig. 12). In the DTC report on upscaling and extrapolation (AC44) the Catchment Matcher Tool was successfully used to extrapolate to the national scale the efficiencies of buffer strips in mitigating fine sediment pollution. National estimates of fine sediment loads from agricultural (from Defra project WQ0223) were used in conjunction with buffer strip efficiencies to estimate potential reductions in sediment yields.

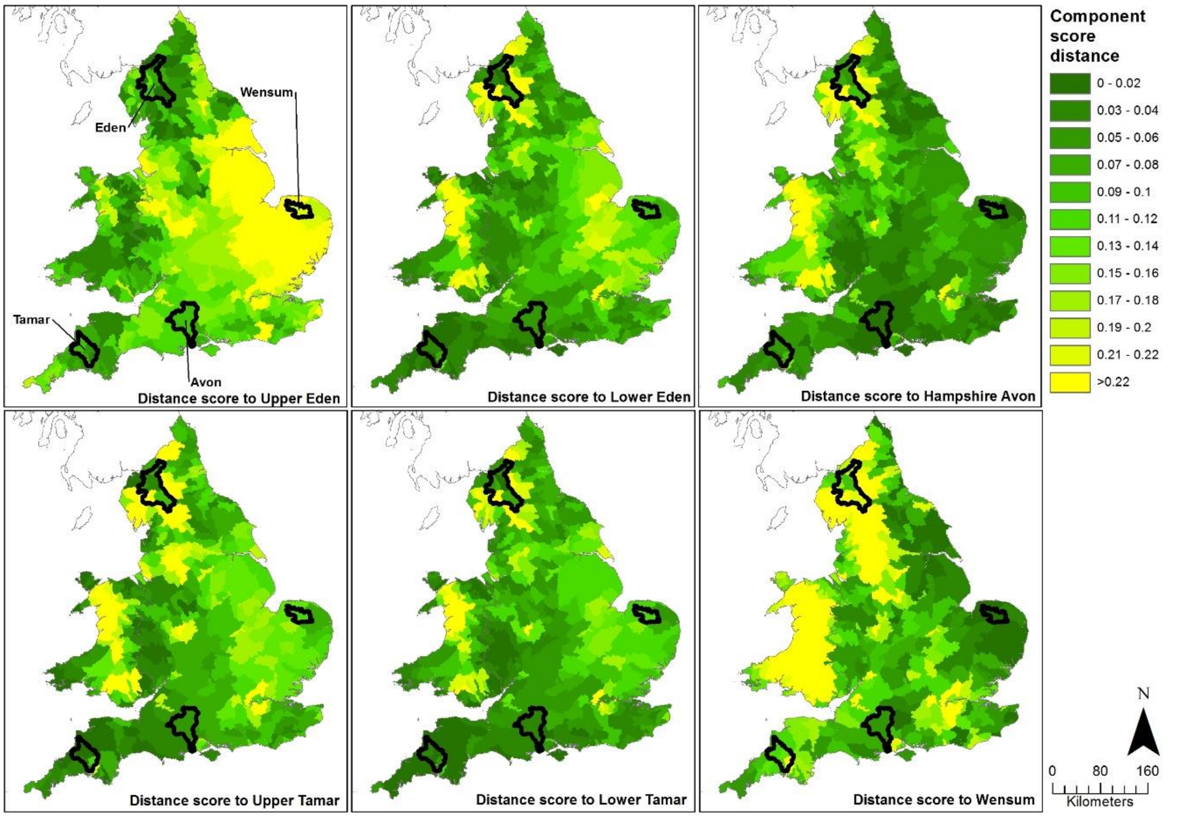

Figure 13: Similarity to each DTC catchment (or sub-catchment) of all WFD Cycle 2 Operational Catchments in England and Wales. A higher component score distance indicates a lower similarity.

The Catchment Matcher tool currently consists of a series of Excel worksheets giving the calculated principal components scores and clusters. Whilst this gives the raw information required to assess catchment similarity in a number of ways, it is not necessarily intuitive for a decision maker to use, and graphical outputs must be created manually by appending summary data to catchment GIS files. A simple user interface for the tool (for example allowing interactive selection of a catchment and visualization of relative similarity to all other catchments as defined by a user-selected subset of variables) would greatly increase its usability.

As with any decision support tool, the Catchment Matcher tool has a number of limitations. Analyses based on ordination space can be potentially misleading. A high degree of average pairwise similarity will be best obtained by a catchment with intermediate values of all variables, rather than one showing extremes. However, if the majority of catchments show extreme values of one sort or another, a catchment with intermediate values may be atypical even though it is the best compromise. Examining similarity over national scales may also mask important differences. For example whilst the Hampshire Avon and South Chilterns are relatively close in ordination (and physical) space, when compared to all other catchments in England and Wales, the way the landscapes have historically been managed is very different, with the former being characterised by extensive grassland and large areas of arable agriculture, and the latter by more mixed agriculture and woodland, with complex field boundaries. Therefore care should be taken when using the Catchment Matcher scores that comparisons are made within a sufficiently restricted set of candidate catchments (e.g. those in the same Catchment Matcher clusters) to make meaningful assessments of similarity. It may also be possible to incorporate additional variables which are important in driving the transferability of findings from the DTCs and feasibility of implementing specific interventions, including additional information on landscape character and history. This may be an important driver of the willingness of stakeholders to take up specific intervention options, even if the physical and political characteristics of a catchment show a high degree of similarity to situations where these options has proven effective. For example, the modification of traditionally open landscapes with tree planting may be unpalatable to land owners or recreational stakeholders. The current Catchment Matcher tool also does not take into account other activities which may limit the feasibility of environmental interventions such as existing modification of the river channel. The need to maintain particularly levels of flow in order to meet the needs of abstraction for drinking water, hydropower, industrial or irrigation purposes may also constrain the transferability of findings from the DTCs beyond the environmental constraints currently in the tool.

As acknowledged by Lovett et al. (2018), the current tool uses only a single year of cropping data from CROME. Longer time-series of crop data (e.g. the CEH Land Cover Plus: Crops datasets) may help elucidate rotations and degree of agricultural intensity. Other data on agricultural intensity exist, including estimated loadings of pesticides and fertilisers (e.g. Jarvis et al. 2019), although these may show strong correlations with other indictors of agricultural intensity (e.g. robust farm type) already in the tool.

10.2 Extending DTC model applications to provide nationwide guidance

The DTC research community has provided detailed and valuable guidance regarding the wider applicability of models applied at DTC sites (WQ0225 WP4: approaches to extrapolate and upscale DTC outputs, 2017). Notably the report highlights the utility of summarising the process models that were applied (citing SWAT CRAFT and PIM) at catchment level to provide a library of export coefficients that can then be linked to FARMSCOPER. This is valuable in many respects, three main capabilities being (i) to look at broad scale sensitivity of water quality to land cover as for example studied in the Eden using SWAT by Yumei Huang (PhD study – as referred to in the WP4 WQ0225 2017 report) (ii) to further provide quantified evidence of the benefits of interventions directly from their representation in process models and (iii) to link with the wide range of measures represented in FARMSCOPER to build in detail of how effectiveness is sensitive to climatic variability in particular the occurrence of extreme events. It is recommended that these types of activities be pursued, and due consideration given to accounting for model uncertainty and performance.

Conceptually, the Catchment Matcher tool provides a valuable means to facilitate extension of model results to a nationwide level. Analysis of results from the tool demonstrate that with some exceptions (in upland Wales and peri-urban southeast England) the DTC catchments are sufficiently representative of the majority of England and Wales. However, currently the evidence base from DTC is insufficient to interpolate the quantified effects of all the interventions considered in DTC to other catchments. This can be done partly. Capability to do this more comprehensively would require two activities: (i) running all interventions through model applications in all DTCs, (ii) collating outputs from the models used in DTC in other catchments in England and Wales, and appraising model outcomes and performance. This would require literature review and/or contacting lead authors for summary results. Brief details of other England/Wales applications of off-the-shelf models used in DTC are provided in Section 7

Empirical models developed in DTC using continuous monitoring data or detailed field/laboratory analysis are also potentially useful for extrapolation in conjunction with other tools such as Catchment Matcher and FARMSCOPER. However, the capability of this approach is restricted by the unique and detailed nature of the DTC monitoring and experimental work which has not been widely repeated elsewhere.

11 Pathways to added value

The social science review has provided some insights for wider integrated catchment management processes such as the Catchment Sensitive Farming processes (reviewed in AC6b) and the Catchment Based Approach (reviewed in AC12); both of which were established to support the delivery of the Water Framework Directive and associated environmental legislation (EU Natura 2000 SACs and SPAs and National designations like SSSIs). The findings are also relevant for other forms of landscape level partnerships (AC9, A45), which are increasingly important in England and Wales, and often overlap or explicitly interact with catchment partnerships (such as the links between the Hampshire Avon Catchment Partnership, and living landscape initiatives in Dorset, Green Space initiatives in Wiltshire and protected habitat initiatives in the New Forest). Despite the selection of the Tamar for its participation in a PES scheme, the documents reviewed provided little commentary on the utility of PES schemes for national level adoption of DWPA mitigation measures. This is something that could be further developed in future, given the increasing interest in harnessing private sector funding for environmental improvements. There were also limited references to integrating WFD with the Floods Directive or combining diffuse pollution mitigation with natural flood management measures (see A43, E9, O15 and W19 for exceptions). However, we believe many of the engagement and behavioural aspects covered in the DTC assessment would be relevant for such processes (Waylen et al., 2018). PI feedback suggests that there was considerable attention to the interaction between flooding and DWPA, although these insights were mainly focussed on how flooding distorts the signal when trying to monitor pollutant delivery pathways.

The CSF and CaBa reviews illustrated the need for investment of time and resources into coordination, relationships and longitudinal engagement (see also A45). This combines deliberation with action, working with stakeholders to demonstrate measures, trying them and adapting them where needed. Despite the ten-year DTC project, the documents suggested that there was insufficient time to work through multiple adaptive management cycles – by the time the social relationships were established, and measures implemented, the monitoring processes were winding up. Feedback from PIs notes that the size of budget and time available was not realistic to develop evidence to reduce DWPA at the sub-catchment scale. The DTC demonstrated the catch-22 that farmers wanted evidence of cost-effective mitigation methods before they would agree to trial them, yet the DTC needed mitigation methods to be implemented on farms to provide cost-effectiveness data, hence the reliance on FARMSCOPER data. Once farmers were convinced, it took more time to identify, apply and negotiate the blended funding mechanisms required, and even more time to implement the measures, where unforeseen practical issues (floods, drought, issues with contractors) created further delays. PI feedback suggests more resource and time is needed if trying to address pollutant pathways in catchments with multiple small farms. Therefore, the data envisaged during phase 2 was only sparsely available, and there were no documented plans to re-survey farmers to track ongoing changes in attitudes or their experiences with measures. This would be a valuable addition to supplement the missing data for SS5 (section 5.6) about likelihood that measures are being sustained post-DTC support. It would be useful to know how the DTC engagement processes were wound down or passed onto other, ongoing, networks. Considering how to 'exit' something like the Wensum Alliance, or how to evolve the consortium post-project funding, is something for all partnerships to consider in future.

Overall, although the DTC has been, and will continue to be, evaluated, it would be useful to explicitly consider evaluation approaches as part of the overall project design for future research platforms. There are vigorous debates about which type of evaluation to adopt, based on the purpose of the evaluation and beliefs about the need for experimental rigour or more interpretivist approaches that prioritise learning (Patton, 2018). For example, Castano et al (2019) argue for counterfactual modelling methodologies when assessing wide programmes of public investment, and the use of indicators selected partially based on agricultural and environmental data already collected. Although integration with existing data sets on farm production and economics was raised by AC5, only the annex of AC21 research seemed to explicitly engage with these data; but this could be built into future project designs. An experimental approach would require monitoring (baseline and follow up) with both participants and non-participants in the DTC catchments as well as outside the DTC catchment areas – ideally by randomly allocating farmers to a control or a treated group. This may not always be possible, as participation in research is voluntary and it may be difficult, or too resource intensive, to recruit the ideal sample. We have shown (SS1, section 5.2 and SS2, section 5.3) that it is not clear how representative the DTC farming sample was. The limited attention to counterfactual aspects of the research (SS6, section 5.7) would restrict the generalisability of the findings despite the fact the arguments are theoretically grounded and the catchments were selected on the basis of national robust farm types. There was no discussion of social research ethics in our extraction table. Some of the farmers implementing DTC measures were named (e.g. the Salle Estate manager) but others were not. One assumes that individuals chose to waive anonymity given they took on the role of champion to demonstrate measures to their peers. However, not all research participants will be willing to do so, which might limit the ability to provide the detail about specific farm intervention uptake processes. This should be considered and potentially these specific lessons learnt might be shared on a restrictive basis to inform policy but not made publicly available. This need to learn lessons whilst understanding stakeholder sensitivity can be inferred from the PI feedback.

There was a tendency for the, data collection and analysis to be either physical or social (which is also reflected in our approach to the REA) rather than an approach that emphasised an explicit interdisciplinary approach. Indeed, AC5 observed that there was more effort spent on monitoring the physical environment than invested in farm economics, agricultural regulation and links to the wider institutional context. Our evaluation of the DTC data records illustrated that they are comprehensive. They are clearly a valuable resource for use in testing/driving catchment water quality models, supporting studies that link hydrochemistry to ecology and evaluating the effect of extreme events. PI feedback suggested that there was insufficient budget to allow the multi-disciplinary consortium to pursue in-depth analysis across all disciplines and study sites. The lessons learnt from this initiative will be taken into account during the current NERC scoping study over plans for investing in a Flood and Drought Research Infrastructure. In Section 6.5 the potential value in further analysis of the data was highlighted. Furthermore, there is a clear opportunity to publicise data for use in other initiatives such as the NERC Strategic Priority Fund project AGLAND where ecosystem services are being modelled and validated. There is a growing literature on the additional demands that interdisciplinary and transdisciplinary research place on researchers (e.g. Lynch et al., 2015). The benefits of the DTC platform such as pooling resources and investing in networks of policy makers, supply chain actors, farmers and researchers (A45) are best realised if challenges in building and sustaining inter- and transdisciplinary research teams are acknowledged; and sufficient time and resources are dedicated to these issues. Projects such as the DTC shape researchers and enrich their science through exposing them to the practicalities of farming systems and the politics of a crowded institutional landscape; generating unique social capital and personal expertise. These attributes are invisible in project outputs but often make valuable contributions to project outcomes – exploring how these networks persist beyond the DTC may also yield useful information when planning future platforms.

Finally, a couple of practical observations. Many databases (e.g. Huddle and Mendeley) do not allow titles of more than a certain length, which can introduce inconsistencies in these databases, particularly where there may be multiple similar titles or citations. It might be useful to use a unique identifier for all material help keep track of documents and make avoiding gaps/duplications easier. Secondly, not all reports had executive summaries, which was surprising given the policy and stakeholder audience. Whilst there were many useful summaries developed for these audiences, it would be good practice to include both policy and technical summaries for long reports; and provide information of how the longer reports related to complementary outputs such as academic publications or briefings. This information may well have been available on the DTC project website but it is always useful to have it repeated within documents, given that project website are not always sustained when the project ends.

Summary: The DTC linked their work to CSF and CaBa processes but insights may also be relevant to landscape partnerships; partnerships focussed on natural flood management and conservation finance or payment for ecosystem services projects. Despite being a ten year programme, more time was needed to identify, agree and implement measures and to provide evidence of ongoing uptake and impacts on water quality. Thought is needed how to report learning when dealing with commercial sensitivities. The costs and benefits of interdisciplinary working are hard to evaluate but the development of human capital and social networks may be important legacies of the DTC process. The DTC has clearly resulted in the collection of a comprehensive dataset and there is great potential for further insights if the data are publicised for further analysis.

12 Concluding Recommendations

The overall aim of the review was to evaluate the knowledge gained on intervention effectiveness and farmer engagement from the DTC programme. The specific objectives of the REA were to assess the volume and characteristics of the DTC evidence base; synthesise what the DTC evidence base indicates; and to offer a critical appraisal of the DTC evidence base. The sections below provide some recommendations based on the knowledge gained, and the gaps identified, from the REA and insights from the wider evaluation of the PICO and SPICE relevant documents.

12.1 Recommendations arising from reviewing the physical science aspects of the DTC

The breadth of understanding gained across a wide range of interventions is undeniably valuable but much more important is to build on this activity to understand how the significance of their environmental effect is influenced by confounding factors such as downstream dilution and climatic variability. From the range and volume of published outputs, many of them not falling within the constrained remit of the PICO analysis, it is clear the knowledge arising from the course of the DTC programme provides a uniquely strong foundation towards a better understanding of these issues. The following recommendations are made, which would involve a combination of monitoring and modelling:

- Continue assessments of waterbodies affected by interventions (and comparable control sites) to better characterise temporal variability and the progressive establishment of effects of the interventions (See Sections 4.2, 4.4.3 and 6.1).
- Undertake further analysis of the existing high quality comprehensive datasets that have been collected in the DTC catchments to investigate the effects of the measures that were implemented (See Section 6).
- Continue to strive for integrated assessments of ecological impacts alongside those of water quality response (See Section 4.4 and 7.1). This is important for future focusing of effort as it is likely that any ecological response to interventions will be prolonged and delayed.
- Assess the extent to which the signatures of interventions are seen to persist downstream to provide better understanding of the significance of processes acting on pollutants such dilution and attenuation (See Section 4.2).
- In the study catchments assess effectiveness of DTC interventions in the context of improvements in point source treatment or changes in point source pollution load (See Sections 4.8 and 6.5).
- Make efforts to assess intervention types across all DTC catchments to improve between-DTC comparisons (See Sections 4.7 and 10). This will help enable a more-comprehensive

extrapolation of findings to other catchments on a national scale, as the catchment matcher tool and other approaches are showing promise for this purpose.

- Some of the models used in the DTC programme have been applied in other non-DTC catchments. To widen the context of programme findings, where possible it would be valuable to extend these non-DTC applications to include the interventions assessed within the DTC (See Section 7).

12.2 Recommendations arising from reviewing social and economic aspect of the DTC:

12.2.1 Robust Engagement Evidence

The recommendations below are based on what was learnt about the DTC evidence on farmer engagement. The recommendations reinforce the positive findings in the DTC documents but also consider how to design future research projects to provide evidence where the REA struggled to identify clear answers to the questions posed.

Record the resource, the specific interactions required to get engagement from individual landowners (or chain of landowners covering intervention to monitoring stations) to access sites for monitoring and possible interventions, which may be time-consuming yet difficult to report on in scientific publications (see Section 11).

Provide clear information on stakeholder analysis and participant sampling protocols, to allow judgements about generalisability and reach of activities (see Section 5.2).

Ensure that data on non-traditional farming voices (new entrants, institutional landowners, women in farming) is captured (see Section 5.2).

Collect data on common challenges (e.g. consultation fatigue, consultation capture, information asymmetry) and whether good design processes could mitigate these (see Section 8.9).

Invest time and resources to collect data on non-participation (as far as possible) (see Section 5.3).

Evaluate the benefits of a diversity of methods throughout the project, allowing time for networks to emerge and evolve but recognising that not all farmers enjoy collective interaction (see section 5.2 and 8.5).

Collect data on the social dynamics of farmer-only safe spaces, cross-sector networks; and the opportunities for these different networks to interact (see Section 8.5).

Plan for, and adapt to, opportunities offered by interacting with other networks and landscape scale interventions (see Section 8.6).

Illustrate how the project is adding value to previous initiatives and respecting the input of prior participants (see Section 8.8).

Collect data on the financial and other resources needed to coordinate and support these networks (see Section 8.4).

Consider when and how scientists end partnerships, or exit the field, when funding ends, including building in time to set up post-project arrangements within the project funding period (see Section 11.

12.2.2 Robust Uptake of Measures Evidence

The recommendations below are based on what was learnt about the DTC evidence on farmer uptake of interventions or measures. The recommendations reinforce the positive findings in the DTC documents but also consider how to design future research projects to provide evidence where the REA struggled to identify clear answers to the questions posed

Ensure that sufficient resource is budgeted to understand changes to the baseline, including plans to resample after the formal project has ended (this may need careful planning in terms of ethical and GDPR agreement) (see Section 5.4).

Collate evidence, and capture the time invested in discussing said evidence with the participants. Plan to work through raising awareness, agreeing the problem framing, agreeing possible solutions and empowering participants to enact the solutions. This may take several years. Capture the time that this process takes to inform future projects (see Section 9.1).

Invest not only in monitoring the biophysical parameters in specific sites, but also the socio-economic parameters for the individual landowner. This may involve sharing sensitive personal and commercial information, requiring delicate negotiation and data sharing agreements, and may not be possible to share the data or outcomes publicly due to the difficulty of anonymising the material. This should be recognised and innovative ways to share learning whilst protecting participants should be considered at the start of the project (see Section 5.6 and 11).

Collect data on monetised costs and benefits to the farm business, but also collect evidence of how measures impact on farming practices, beyond the final productivity and profitability calculations (see Section 5.6 and 9.4).

Collect data on the full range of costs (including transaction costs of setting up and participating in funding schemes) and the distribution of these costs, as well as the distribution of benefits (see Section 5.6).

It may be better to trial fewer measures, over more farms with different farming practices or preferences, to allow more complete socio-economic comparisons (see Section 5.5 and 5.7).

Understand, and allow for, confounding factors that affect the research but are outside the researchers' control and capture these influences to allow for robust interpretation of results. As noted, this suggests an adaptive management approach using social learning rather than strictly experimental research design. Alternatively, consider using an explicitly counterfactual research design, using indicators from ongoing time series data that pre- and post-date the research; and cover a suitable control population (see Section 5.7 and Section 11)

12.2.3 Developing Robust Transdisciplinary Platforms

The recommendations below are based on what was learnt about the operationalisation of the DTC platform form the DTC documentation and PI feedback. The recommendations reinforce the positive findings in the DTC documents but also consider how to design future research projects in light of the considerable experience gained across the three consortia.

Understand the complexity of sampling processes that try to combine the heterogeneity of biophysical catchments, farming systems and socio-economic typologies of land managers, and recognise that due to the voluntary nature of participation, there are likely to be gaps. Allowing redundancies and duplication at the start may be useful given the potential for participants to drop out over time (see Section 5.1 and Section 5.5).

Citizen Science approaches need to be approached with a clear vision of their purpose – to raise awareness and/or to expand data points and/or to empower farmers to self-evaluate their practices – each requires a different strategy and will achieve different, though equally valid, outcomes (see Section 9.1).

Consider how the research will feed into the wider institutional landscape in which it sits – academics increasingly develop pathways to impact as part of proposals, but the pathway also depends on non-academic partners, who are not always involved in planning these pathways, or able/willing to engage in developing the implications of the research. This may be particularly true of commercial partners, who are not funded to participate in public good science. Changing policy and stakeholder priorities can be challenging to manage over a ten year period (see Section 1.5 and 9.6).

13 References

Boorman, D.B. Climate, Hydrochemistry and Economics of Surface-water Systems (CHESS): adding a European dimension to the catchment modelling experience developed under LOIS. *Science of the Total Environment* 2003; 314: 411-437.

Bouraoui, F., Galbiati, L., and Bidoglio, G. Climate change impacts on nutrient loads in the Yorkshire Ouse catchment (UK). *Hydrology and Earth System Sciences* 2002; 6(2): 197-209.

Bowes, M. J., H. P. Jarvie, P. S. Naden, G. H. Old, P. M. Scarlett, C. Roberts, L. K. Armstrong, S. A. Harman, H. D. Wickham & A. L. Collins, 2014. Identifying priorities for nutrient mitigation using river concentration–flow relationships: The Thames basin, UK. J Hydrol 517(0):1-12 doi:http://dx.doi.org/10.1016/j.jhydrol.2014.03.063.

Bowes, M. J., C. Neal, H. P. Jarvie, J. T. Smith & H. N. Davies, 2010. Predicting phosphorus concentrations in British rivers resulting from the introduction of improved phosphorus removal from sewage effluent. Sci Total Environ 408(19):4239-4250 doi:10.1016/j.scitotenv.2010.05.016.

Bowes, M. J., J. T. Smith & C. Neal, 2009. The value of high-resolution nutrient monitoring: A case study of the River Frome, Dorset, UK. J Hydrol 378(1-2):82-96 doi:10.1016/j.jhydrol.2009.09.015.Burbi, S., Baines, R.N., and Conway J.S. (2016) Achieving successful farmer engagement on greenhouse gas emission mitigation. *International Journal of Agricultural Sustainability* 14:4, p 466-483.Bussi, G., Janes, V., Whitehead, P.G., Dadson, S.J., and Holman, I.P. Dynamic response of land use and river nutrient concentration to long-term climatic changes. *Science of the Total Environment* 2017; 590: 818-831.

Bussi, G., and Whitehead, P.G. Impacts of droughts on low flows and water quality near power stations. *Hydrological Sciences Journal* 2020; 1-16.

Bussi, G., Whitehead, P.G., Gutiérrez-Cánovas, C., Ledesma, J.L.J., Ormerod, S.J., and Couture R.M. Modelling the effects of climate and land-use change on the hydrochemistry and ecology of the River Wye (Wales). *Science of the Total Environment* 2018; 627: 733-743.

Bussi, G., Whitehead, P.G., Thomas, A.R.C., Masante, D., Jones, L., Cosby, B.J., Emmett, B.A., Malham, S.K., Prudhomme, C., and Prosser, H. Climate and land-use change impact on faecal indicator bacteria in a temperate maritime catchment (the River Conwy, Wales). *Journal of Hydrology* 2017; 553: 248-261.

Castano J, Blanco M and Martinez P (2019) Reviewing Counterfactual Analyses to Assess Impacts of EU Rural Development Programmes: What lessons can be learned from the 2007-2012 Ex-Post Evaluations? *Sustainability*, 11, 1105: DOI: 10.2290/su11041105

Collins, A., Coughlin, D., Miller, J. and Kirk, S. (2015) The production of Quick Scoping Reviews and Rapid Evidence Assessments. A How to Guide. Defra, UK, p.1-63.

Cuttle, S.P., Macleod, C.J.A., Chadwick, D.R., Scholefield, D., Haygarth, P.M., Newell-Price, P., Harris, D., Shepherd, M.A., Chambers, B.J. and Humphrey, R. An Inventory of Methods to Control Diffuse Water Pollution from Agriculture (DWPA). 2007; User Manual (Defra Project ES0203), UK.

Dean, S., Freer, J., Beven, K., Wade, A.J., and Butterfield, D. Uncertainty assessment of a process-based integrated catchment model of phosphorus. *Stochastic Environmental Research and Risk Assessment* 2009; 23(7): 991-1010.European Commission (no date) *Participatory approaches for agricultural innovation,* https://ec.europa.eu/eip/agriculture/sites/agri-eip/files/eipagri_brochure_participatory_approaches_2015_en_web.pdf [accessed 3rd September 2019]

Defra (2020) WT15116 The Demonstration Test Catchments Evidence Compendium, 78 pages [available online, http://sciencesearch.defra.gov.uk/.]

Defra (2020) WT15115 Evidence Compendium Evaluation of knowledge gained from the National Demonstration Test Catchment (DTC) Project, 28 pages [available online, http://sciencesearch.defra.gov.uk/.]

Franks, J. R. (2011). The collective provision of environmental goods: a discussion of contractual issues. *Journal of Environmental Planning and Management*, 54(5), 637–660.

Glavan, M., White, S., and Holman, I.P. Evaluation of river water quality simulations at a daily time step–Experience with SWAT in the Axe Catchment, UK. *CLEAN–Soil, Air, Water* 2011; 39(1): 43-54.

Grizzetti, B., Bouraoui, F.D.M.G., and De Marsily, G. Modelling nitrogen pressure in river basins: A comparison between a statistical approach and the physically-based SWAT model. *Physics and Chemistry of the Earth, Parts A/B/C* 2005; 30(8-10): 508-517.

Halliday, S. J., R. A. Skeffington, A. J. Wade, M. J. Bowes, E. Gozzard, J. R. Newman, M. Loewenthal, E. J. Palmer-Felgate & H. P. Jarvie, 2015. High-frequency water quality monitoring in an urban catchment: hydrochemical dynamics, primary production and implications for the Water Framework Directive. Hydrol Process 29:3388-3407 doi:10.1002/hyp.10453.

Hankin, B., Bielby, S., Pope, L., and Douglass, J. Catchment-scale sensitivity and uncertainty in water quality modelling. *Hydrological Processes* 2016; 30(22): 4004-4018.

Holder, Amanda J., Rowe, R., and McNamara, N.P. Soil & Water Assessment Tool (SWAT) simulated hydrological impacts of land use change from temperate grassland to energy crops: A case study in western UK. *GCB Bioenergy* 2019; 11(11): 1298-1317.

Jarvis, S.G.; Henrys, P.A.; Redhead, J.W.; Da Silva Osório, B.M.; Pywell, R.F. (2019). CEH Land Cover plus: Pesticides 2012-2016 (England and Wales). NERC Environmental Information Data Centre. https://doi.org/10.5285/a72f8ce8-561f-4f3a-8866-5da620c0c9fe

Jordan, P., A. R. Melland, P. E. Mellander, G. Shortle & D. Wall, 2012. The seasonality of phosphorus transfers from land to water: Implications for trophic impacts and policy evaluation. Sci Total Environ 434(0):101-109 doi:http://dx.doi.org/10.1016/j.scitotenv.2011.12.070.

Kannan, N., White, S.M, and Whelan, M.J. Predicting diffuse-source transfers of surfactants to surface waters using SWAT. *Chemosphere* 2007; 66(7): 1336-1345.Lazar, A.N., Butterfield, D., Futter, M.N., Rankinen, K., Thouvenot-Korppo, M., Jarritt, N, Lawrence, D.S.L., Wade, A.J., and Whitehead, P.G. An assessment of the fine sediment dynamics in an upland river system: INCA-Sed modifications and implications for fisheries. *Science of the Total Environment* 2010; 408(12): 2555-2566.

Lloyd, C. E. M., J. E. Freer, P. J. Johnes, G. Coxon & A. L. Collins, 2016. Discharge and nutrient uncertainty: implications for nutrient flux estimation in small streams. Hydrol Process 30(1):135-152 doi:10.1002/hyp.10574.

Lovett, A., Sünnenberg, G., Burke, S., and Collins, A. 2018. Catchment Matcher. DEFRA, WQ0225. 52pp.

Lynch, A. J. J., et al. (2015). "Transdisciplinary synthesis for ecosystem science, policy and management: The Australian experience." *Science of the Total Environment* 534: 173-184.

Mills, J., et al. (2017) Engaging farmers in environmental management through a better understanding of behaviour. *Agriculture and Human Values*, 34, 283–299.

Naden, P. S., A. L. Collins, M. J. Bowes, F. Edwards & e. al, 2015. Developing a field tool kit for ecological targeting of agricultural diffuse pollution mitigation measures. Defra Report WQ0223.

OECD (2015) *Stakeholder Engagement for Inclusive Water Governance*, OECD Studies on Water, OECD, Paris. http://dx.doi.org/10.1787/9789264231122-en

Patton MQ (2018) Evaluation Science, *American Journal of Evaluation*, 39 2 183-200 https://journals.sagepub.com/doi/10.1177/1098214018763121

POSTnote (2014) Diffuse Pollution of Water by Agriculture, Parliamentary Office of Science and Technology, POSTnote 478, 5 pp.

Randall, N.P., Donnison, L.M., Lewis, P.J. and James, K. 2015. How effective are on-farm mitigation measures for delivering an improved water environment? A systematic map. Environmental Evidence, 4 (18).

Šumane, S. et al. (2018) Local and farmers' knowledge matters! How integrating informal and formal knowledge enhances sustainable and resilient agriculture. *Journal of Rural Studies* 59, p 232-241.

Wade, A. J., E. J. Palmer-Felgate, S. J. Halliday, R. A. Skeffington, M. Loewenthal, H. P. Jarvie, M. J. Bowes, G. M. Greenway, S. J. Haswell, I. M. Bell, E. Joly, A. Fallatah, C. Neal, R. J. Williams, E.

Gozzard & J. R. Newman, 2012. Hydrochemical processes in lowland rivers: insights from in situ, high-resolution monitoring. Hydrol Earth Syst Sci 16(11):4323-4342 doi:10.5194/hess-16-4323-2012.

WaterLIFE (2016) *Encouraging Farmers To Get Involved In Agri-Environment Initiatives*, 2 page briefing: https://catchmentbasedapproach.org/wp-content/uploads/2018/07/07_complete_engaging_farmers.pdf

Waylen K A et al. (2017) Challenges to enabling and implementing Natural Flood Management in Scotland, *Journal of Flood Risk Management*, 11, 2, 1078 – 1089.

Whitehead, P.G., Futter, M.N., Comber, S., Butterfield, D., Pope, L., Willows, R., and Burgess, C. Modelling impacts of seasonal wastewater treatment plant effluent permits and biosolid substitution for phosphorus management in catchments and river systems. *Hydrology Research* 2015; 46(3): 313-324.

Whitehead, P.G., Hill, T.J., and Neal, C. Impacts of forestry on nitrogen in upland and lowland catchments: a comparison of the River Severn at Plynlimon in mid-Wales and the Bedford Ouse in south-east England using the INCA Model. *Hydrology and Earth System Sciences* 2004; 8(3): 533-544

Whitehead, P.G., Lapworth, D.J., Skeffington, R.A., and Wade, A.J. Excess nitrogen leaching and C/N decline in the Tillingbourne catchment, southern England: INCA process modelling for current and historic time series. *Hydrology and Earth System Sciences* 2002; 6(3): 455-466

Whitehead, P.G., Wade, A.J., and Butterfield, D. Potential impacts of climate change on water quality and ecology in six UK rivers. *Hydrology Research* 2009; 40(2-3): 113-122.

Yanow, D., 2007. Qualitative-Interpretive Methods in Policy Research, in: Fischer, F., Miller, G., Sidney, M. (Eds.). CRC Press, Boca Raton.

Appendix A: List of Sources and their IDs

A1 Adams, R., and G. Owen. 2015. "Developing a Framework to Link Catchment Modelling tools to Decision Support Systems for Catchment Management and Planning." In *Geophysical Research Abstracts, EGU General Assembly 2015*. p. 1p.

A2 Adams, R., P. Quinn, N. Barber, and S. Reaney. 2018. "The role of attenuation and land management in small catchments to remove sediment and phosphorus: A modelling study of mitigation options and impacts." *Water* 10(1227).

A3 Adams, R., P.F. Quinn, M. Perks, N.J. Barber, J. Jonczyk, and G.J. Owen. 2016. "Simulating high frequency water quality monitoring data using a catchment runoff attenuation flux tool (CRAFT)." *Science of the Total Environment* 572:1622–1635. Available at: http://dx.doi.org/10.1016/j.scitotenv.2016.01.045.

A4 Allen, D.J., W.G. Darling, J. Davies, A.J. Newell, D.C. Gooddy, and A.L. Collins. 2014. "Groundwater conceptual models: Implications for evaluating diffuse pollution mitigation measures." *Quarterly Journal of Engineering Geology and Hydrogeology* 47(1):65–80.

A5 Barber, Reaney, Barker, Benskin, Burke, Cleasby, Haygarth, Jonczyk, Owen & Snel. 2016. "The Treatment Train approach to reducing non-point source pollution from agriculture" *AGU Fall Meeting Abstracts, 2016,* 1p.

A6 Biddulph, M. 2016. *Testing the efficacy of mitigation measures for reducing fine sediment and associated pollutant delivery to and through rivers in agricultural catchments of England*. Doctoral Thesis. University of Northampton. Available at: https://www.researchgate.net/publication/316994083.

A7 Biddulph, M., A.L. Collins, I.D.L. Foster, and N. Holmes. 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.

A8 Collins, A.L., J.P. Newell Price, Y. Zhang, R. Gooday, P.S. Naden, and D. Skirvin. 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(February):1499–1511. Available at: https://doi.org/10.1016/j.scitotenv.2017.10.078.

A9 Collins, A.L., Y.S. Zhang, M. Winter, A. Inman, J.I. Jones, P.J. Johnes, W. Cleasby, E. Vrain, A. Lovett, and L. Noble. 2016. "Tackling agricultural diffuse pollution: What might uptake of farmer-preferred measures deliver for emissions to water and air?" *Science of the Total Environment* 547:269–281. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.12.130.

A10 Collins, A.L., Y.S. Zhang, R. Hickinbotham, G. Bailey, S. Darlington, S.E. Grenfell, R. Evans, and M. Blackwell. 2013. "Contemporary fine-grained bed sediment sources across the River Wensum Demonstration Test Catchment, UK." *Hydrological Processes* 27(6):857–884.

A11 Cooper, R.J., Z.M. Battams, S.H. Pearl, and K.M. Hiscock. 2019. "Mitigating river sediment enrichment through the construction of roadside wetlands." *Journal of Environmental Management* 231(October 2018):146–154. Available at: https://doi.org/10.1016/j.jenvman.2018.10.035.

A12 Cooper, R.J., B.G. Rawlins, B. Lézé, T. Krueger, and K.M. Hiscock. 2014. "Combining two filter paper-based analytical methods to monitor temporal variations in the geochemical properties of fluvial suspended particulate matter." *Hydrological Processes* 28(13):4042–4056.

A13 Cooper, R.J., P. Fitt, K.M. Hiscock, A.A. Lovett, L. Gumm, S.J. Dugdale, J. Rambohul, A. Williamson, L. Noble, J. Beamish, and P. Hovesen. 2016. "Assessing the effectiveness of a three-stage on-farm biobed in treating pesticide contaminated wastewater." *Journal of Environmental Management* 181:874–882. Available at: http://dx.doi.org/10.1016/j.jenvman.2016.06.047.

A14	Cooper, R.J., T. Krueger, K.M. Hiscock, and B.G. Rawlins. 2015. "High-temporal resolution fluvial sediment source fingerprinting with uncertainty: A Bayesian approach." *Earth Surface Processes and Landforms* 40(1):78–92.
A15	Cooper, R.J., Z. Hama-Aziz, K.M. Hiscock, A.A. Lovett, S.J. Dugdale, G. Sünnenberg, L. Noble, J. Beamish, and P. Hovesen. 2017. "Assessing the farm-scale impacts of cover crops and non-inversion tillage regimes on nutrient losses from an arable catchment." *Agriculture, Ecosystems and Environment* 237:181–193. Available at: http://dx.doi.org/10.1016/j.agee.2016.12.034.
A16	Cooper, R.J., K.M. Hiscock, A.A. Lovett, S.J. Dugdale, G. Sünnenberg, N.L. Garrard, F.N. Outram, Z.Q. Hama-Aziz, L. Noble, and M.A. Lewis. 2018. "Application of high-resolution telemetered sensor technology to develop conceptual models of catchment hydrogeological processes." *Journal of Hydrology X* 1:100007. Available at: https://doi.org/10.1016/j.hydroa.2018.100007.
A17	Cooper, R.J., F.N. Outram, and K.M. Hiscock. 2016. "Diel turbidity cycles in a headwater stream: evidence of nocturnal bioturbation?" *Journal of Soils and Sediments* 16(6):1815–1824. Available at: http://dx.doi.org/10.1007/s11368-016-1372-y.
A18	Cooper, R.J., N. Pedentchouk, K.M. Hiscock, P. Disdle, T. Krueger, and B.G. Rawlins. 2015. "Apportioning sources of organic matter in streambed sediments: An integrated molecular and compound-specific stable isotope approach." *Science of the Total Environment* 520:187–197. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.03.058.
A19	Cooper, R.J., S.K. Wexler, C.A. Adams, and K.M. Hiscock. 2017. "Hydrogeological Controls on Regional-Scale Indirect Nitrous Oxide Emission Factors for Rivers." *Environmental Science and Technology* 51(18):10440–10448.
A20	Cooper, R.J., B.G. Rawlins, T. Krueger, B. Lézé, K.M. Hiscock, and N. Pedentchouk. 2015. "Contrasting controls on the phosphorus concentration of suspended particulate matter under baseflow and storm event conditions in agricultural headwater streams." *Science of the Total Environment* 533:49–59. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.06.113.
A21	Cooper, R.J., T. Krueger, K.M. Hiscock, and B.G. Rawlins. 2014. "Sensitivity of fluvial sediment source apportionment to mixing model assumptions: A Bayesian model comparison." *Water Resources Research* 50:9031–9047.
A23-24	Dixon, S.C. 2012. *The Effects of Land Management and Predicted Climate Change on Hydrological Connectivity and Diffuse Fine Sediment Pollution Risk Within the River Eden Catchment*. Durham University. Masters Thesis. 171pp.
A25	Forber, K.J., M.C. Ockenden, C. Wearing, M.J. Hollaway, P.D. Falloon, R. Kahana, M.L. Villamizar, J.G. Zhou, P.J.A. Withers, K.J. Beven, A.L. Collins, R. Evans, K.M. Hiscock, C.J.A. Macleod, and P.M. Haygarth. 2017. "Determining the effect of drying time on phosphorus solubilization from three agricultural soils under climate change scenarios." *Journal of Environmental Quality* 46(5):1131–1136.
A26	Gooday, R.D., S.G. Anthony, D.R. Chadwick, P. Newell-Price, D. Harris, D. Duethmann, R. Fish, A.L. Collins, and M. Winter. 2014. "Modelling the cost-effectiveness of mitigation methods for multiple pollutants at farm scale." *Science of the Total Environment* 468–469:1198–1209. Available at: http://dx.doi.org/10.1016/j.scitotenv.2013.04.078.
A27	Hama-Aziz, Z.Q., K.M. Hiscock, and R.J. Cooper. 2017. "Dissolved nitrous oxide (N2O) dynamics in agricultural field drains and headwater streams in an intensive arable catchment." *Hydrological Processes* 31(6):1371–1381.
A28	Hama-Aziz, Z.Q., K.M. Hiscock, and R.J. Cooper. 2017. "Indirect Nitrous Oxide Emission Factors for Agricultural Field Drains and Headwater Streams." *Environmental Science and Technology* 51(1):301–307.
A29	Heppell, C.M. 2016. "Hampshire Avon: Vertical head gradient, saturated hydraulic conductivity and pore water chemistry data from six river reaches". *NERC Environmental Information Data Centre*. 8pp.

A30 Heppell, C.M., A. Binley, M. Trimmer, T. Darch, A. Jones, E. Malone, A.L. Collins, P.J. Johnes, J.E. Freer, and C.E.M. Lloyd. 2017. "Hydrological controls on DOC nitrate resource stoichiometry in a lowland, agricultural catchment, southern UK." *Hydrology and Earth System Sciences* 21(9):4785–4802.

A31 Hollaway, M.J., K.J. Beven, C.M.W.H. Benskin, A.L. Collins, R. Evans, P.D. Falloon, K.J. Forber, K.M. Hiscock, R. Kahana, C.J.A. Macleod, M.C. Ockenden, M.L. Villamizar, C. Wearing, P.J.A. Withers, J.G. Zhou, N.J. Barber, and P.M. Haygarth. 2018. "A method for uncertainty constraint of catchment discharge and phosphorus load estimates." *Hydrological Processes* 32(17):2779–2787.

A32 Hollaway, M.J., K.J. Beven, C.M.W.H. Benskin, A.L. Collins, R. Evans, P.D. Falloon, K.J. Forber, K.M. Hiscock, R. Kahana, C.J.A. Macleod, M.C. Ockenden, M.L. Villamizar, C. Wearing, P.J.A. Withers, J.G. Zhou, N.J. Barber, and P.M. Haygarth. 2018. "The challenges of modelling phosphorus in a headwater catchment: Applying a 'limits of acceptability' uncertainty framework to a water quality model." *Journal of Hydrology* 558:607–624. Available at: https://doi.org/10.1016/j.jhydrol.2018.01.063.

A33 Huang, Y., P. Quinn, Q. Liang, and R. Adams. 2017. "Impacts of Cropland Changes on Water Balance, Sediment and Nutrient Transport in Eden River, UK." In *Geophysical Research Abstracts, EGU General Assembly 2017*. p. 1p.

A34 Huang, Y., P. Quinn, Q. Liang, and R. Adams. 2017. "The benefits of daily data and scale up issues in hydrologic models-SWAT and CRAFT." In *Geophysical Research Abstracts, EGU General Assembly 2017*. p. 1p.

A35 Inman, A., M. Winter, R. Wheeler, E. Vain, A. Lovett, A. Collins, I. Jones, P. Johnes, and W. Cleasby. 2018. "An exploration of individual, social and material factors influencing water pollution mitigation behaviours within the farming community." *Land Use Policy* 70(December 2016):16–26. Available at: https://doi.org/10.1016/j.landusepol.2017.09.042.

A36 Janes, V.J.J., A.P. Nicholas, A.L. Collins, and T.A. Quine. 2017. "Analysis of fundamental physical factors influencing channel bank erosion: results for contrasting catchments in England and Wales." *Environmental Earth Sciences* 76(7):1–18.

A37 Lloyd, C.E.M., J.E. Freer, A.L. Collins, P.J. Johnes, and J.I. Jones. 2014. "Methods for detecting change in hydrochemical time series in response to targeted pollutant mitigation in river catchments." *Journal of Hydrology* 514:297–312. Available at: http://dx.doi.org/10.1016/j.jhydrol.2014.04.036.

A38 Lloyd, C.E.M., J.E. Freer, P.J. Johnes, and A.L. Collins. 2016. "Using hysteresis analysis of high-resolution water quality monitoring data, including uncertainty, to infer controls on nutrient and sediment transfer in catchments." *Science of the Total Environment* 543:388–404. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.11.028.

A39 Lloyd, C.E.M., J.E. Freer, P.J. Johnes, and A.L. Collins. 2016. "Technical Note: Testing an improved index for analysing storm discharge-concentration hysteresis." *Hydrology and Earth System Sciences* 20(2):625–632.

A40 Lloyd, C.E.M., J.E. Freer, P.J. Johnes, G. Coxon, and A.L. Collins. 2016. "Discharge and nutrient uncertainty: Implications for nutrient flux estimation in small streams." *Hydrological Processes* 30(1):135–152.

A41 Lloyd, C.E.M., P.J. Johnes, J.E. Freer, A.M. Carswell, J.I. Jones, M.W. Stirling, R.A. Hodgkinson, C. Richmond, and A.L. Collins. 2019. "Determining the sources of nutrient flux to water in headwater catchments: Examining the speciation balance to inform the targeting of mitigation measures." *Science of the Total Environment* 648:1179–1200. Available at: https://doi.org/10.1016/j.scitotenv.2018.08.190.

A42 Lovett, A., K. Hiscock, F. Outram, R. Cooper, S. Dugdale, J. Stevenson, G. Sunnenberg, Z. Hama-Aziz, T. Dockerty, L. Noble, J. Beamish, and P. Hovesen. 2015. "Experiments with cover crops and cultivation techniques in the Wensum DTC." *Aspects of Applies Biology* 129:85–90.

A43 Mackay, E.B., M.E. Wilkinson, C.J.A. Macleod, K. Beven, B.J. Percy, M.G. Macklin, P.F. Quinn, M. Stutter, and P.M. Haygarth. 2015. "Digital catchment observatories: A platform for engagement and knowledge exchange between catchment scientists, policy makers, and local communities." *Water Resources Research* 51(6):4815–4822.

A44	Marshall, K. 2011. "Critical review of international literature relating to integrated and adaptive natural resource management, with particular reference to effective stakeholder engagement in catchment management." *Report, the James Hutton Institute,* 21p.
A45	McGonigle, D.F., S.P. Burke, A.L. Collins, R. Gartner, M.R. Haft, R.C. Harris, P.M. Haygarth, M.C. Hedges, K.M. Hiscock, and A.A. Lovett. 2014. "Developing Demonstration Test Catchments as a platform for transdisciplinary land management research in England and Wales." *Environmental Sciences: Processes and Impacts* 16(7):1618–1628.
A46	Ockenden, M.C., C.E. Deasy, C.M.W.H. Benskin, K.J. Beven, S. Burke, A.L. Collins, R. Evans, P.D. Falloon, K.J. Forber, K.M. Hiscock, M.J. Hollaway, R. Kahana, C.J.A. Macleod, S.M. Reaney, M.A. Snell, M.L. Villamizar, C. Wearing, P.J.A. Withers, J.G. Zhou, and P.M. Haygarth. 2016. "Changing climate and nutrient transfers: Evidence from high temporal resolution concentration-flow dynamics in headwater catchments." *Science of the Total Environment* 548–549(January):325–339. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.12.086.
A47	Ockenden, M.C., M.J. Hollaway, K.J. Beven, A.L. Collins, R. Evans, P.D. Falloon, K.J. Forber, K.M. Hiscock, R. Kahana, C.J.A. MacLeod, W. Tych, M.L. Villamizar, C. Wearing, P.J.A. Withers, J.G. Zhou, P.A. Barker, S. Burke, J.E. Freer, P.J. Johnes, M.A. Snell, B.W.J. Surridge, and P.M. Haygarth. 2017. "Major agricultural changes required to mitigate phosphorus losses under climate change." *Nature Communications* 8(1). Available at: http://dx.doi.org/10.1038/s41467-017-00232-0.
A48	Ockenden, M.C., W. Tych, K.J. Beven, A.L. Collins, R. Evans, P.D. Falloon, K.J. Forber, K.M. Hiscock, M.J. Hollaway, R. Kahana, C.J.A. Macleod, M.L. Villamizar, C. Wearing, P.J.A. Withers, J.G. Zhou, C.M.W.H. Benskin, S. Burke, R.J. Cooper, J.E. Freer, and P.M. Haygarth. 2017. "Prediction of storm transfers and annual loads with data-based mechanistic models using high-frequency data." *Hydrology and Earth System Sciences* 21(12):6425–6444.
A49	Outram, F.N., R.J. Cooper, G. Sünnenberg, K.M. Hiscock, and A.A. Lovett. 2016. "Antecedent conditions, hydrological connectivity and anthropogenic inputs: Factors affecting nitrate and phosphorus transfers to agricultural headwater streams." *Science of the Total Environment* 545–546:184–199. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.12.025.
A50	Outram, F.N., C. Lloyd, J. Jonczyk, C.M.H. Benskin, F. Grant, S.R. Dorling, C.J. Steele, A.L. Collins, J. Freer, P.M. Haygarth, K.M. Hiscock, P.J. Johnes, and A.L. Lovett. 2013. "High-resolution monitoring of catchment nutrient response to the end of the 2011–2012 drought in England, captured by the demonstration test catchments." *Hydrology and Earth System Sciences Discussions* 10(12):15119–15165.
A51	Outram, F.N., C.E.M. Lloyd, J. Jonczyk, C. McW. H. Benskin, F. Grant, M.T. Perks, C. Deasy, S.P. Burke, A.L. Collins, J. Freer, P.M. Haygarth, K.M. Hiscock, P.J. Johnes, and A.L. Lovett. 2014. "High-frequency monitoring of nitrogen and phosphorus response in three rural catchments to the end of the 2011-2012 drought in England." *Hydrology and Earth System Sciences* 18(9):3429–3448.
A52	Owen, G.J., M.T. Perks, C.M.W.H. Benskin, M.E. Wilkinson, J. Jonczyk, and P.F. Quinn. 2012. "Monitoring agricultural diffuse pollution through a dense monitoring network in the River Eden Demonstration Test Catchment, Cumbria, UK." *Area* 44(4):443–453.
A53	Perks, M.T., G.J. Owen, C.M.W.H. Benskin, J. Jonczyk, C. Deasy, S. Burke, S.M. Reaney, and P.M. Haygarth. 2015. "Dominant mechanisms for the delivery of fine sediment and phosphorus to fluvial networks draining grassland dominated headwater catchments." *Science of the Total Environment* 523:178–190. Available at: http://dx.doi.org/10.1016/j.scitotenv.2015.03.008.
A54	Quinn, P., J. Jonczyk, G. Owen, N. Barber, R. Adams, G. Odonnell, and E. Team. 2015. "The role of high frequency monitoring in understanding nutrient pollution processes to address catchment

management issues." In *Geophysical Research Abstracts, EGU General Assembly 2015*.

A55	Reaney, S.M. 2016. "Mitigating Agricultural Diffuse Pollution: Learning from The River Eden Demonstration Test Catchment Experiments." In *AGU Fall Meeting*. San Francisco, p. 1p.
A56	Snell, M.A., P.A. Barker, A. Ashar, N. Barber, C. Benskin, S. Burke, W. Cleasby, P. Haygarth, J.C. Jonczyk, G.J. Owen, M.T. Perks, P.F. Quinn, and S.M. Reaney. 2016. "Spatio-temporal dynamics in phytobenthos structural properties reveal insights into agricultural catchment dynamics and nutrient fluxes." In *AGU Fall Meeting*. San Francisco, p. 1p.
A57	Riley, W.D., E.C.E. Potter, J. Biggs, A.L. Collins, H.P. Jarvie, J.I. Jones, M. Kelly-Quinn, S.J. Ormerod, D.A. Sear, R.L. Wilby, S. Broadmeadow, C.D. Brown, P. Chanin, G.H. Copp, I.G. Cowx, A. Grogan, D.D. Hornby, D. Huggett, M.G. Kelly, M. Naura, J.R. Newman, and G.M. Siriwardena. 2018. "Small Water Bodies in Great Britain and Ireland: Ecosystem function, human-generated degradation, and options for restorative action." *Science of the Total Environment* 645:1598–1616. Available at: https://doi.org/10.1016/j.scitotenv.2018.07.243.
A58	Roberts, E.J., and R.J. Cooper. 2018. "Riverbed sediments buffer phosphorus concentrations downstream of sewage treatment works across the River Wensum catchment, UK." *Journal of Soils and Sediments* 18(5):2107–2116.
A59	Roberts, W.M., R.A. Matthews, M.S.A. Blackwell, S. Peukert, A.L. Collins, M.I. Stutter, and P.M. Haygarth. 2013. "Microbial biomass phosphorus contributions to phosphorus solubility in riparian vegetated buffer strip soils." *Biology and Fertility of Soils* 49(8):1237–1241.
A60	Snell, M.A., P.A. Barker, B.W.J. Surridge, C.M.W.H. Benskin, N. Barber, S.M. Reaney, W. Tych, D. Mindham, A.R.G. Large, S. Burke, and P.M. Haygarth. 2019. "Strong and recurring seasonality revealed within stream diatom assemblages." *Scientific Reports* 9(1):1–7. Available at: http://dx.doi.org/10.1038/s41598-018-37831-w.
A61	Snell, M.A., P.A. Barker, B.W.J. Surridge, A.R.G. Large, J. Jonczyk, C.M.W.H. Benskin, S. Reaney, M.T. Perks, G.J. Owen, W. Cleasby, C. Deasy, S. Burke, and P.M. Haygarth. 2014. "High frequency variability of environmental drivers determining benthic community dynamics in headwater streams." *Environmental Sciences: Processes and Impacts* 16(7):1629–1636.
A62	Taylor, S.D., Y. He, and K.M. Hiscock. 2016. "Modelling the impacts of agricultural management practices on river water quality in Eastern England." *Journal of Environmental Management* 180:147–163.
A63	Terry, J.A., C. McW.h. Benskin, E.F. Eastoe, and P.M. Haygarth. 2014. "Temporal dynamics between cattle in-stream presence and suspended solids in a headwater catchment." *Environmental Sciences: Processes and Impacts* 16(7):1570–1577.
A64	Vrain, E., and A. Lovett. 2016. "The roles of farm advisors in the uptake of measures for the mitigation of diffuse water pollution." *Land Use Policy* 54:413–422. Available at: http://dx.doi.org/10.1016/j.landusepol.2016.03.007.
A65	Vrain, E., and A. Lovett. 2019. "Using word clouds to present farmers' perceptions of advisory services on pollution mitigation measures." *Journal of Environmental Planning and Management* 0(0):1–18. Available at: https://doi.org/10.1080/09640568.2019.1638232.
A66	Wang, Y. 2015. *Concentration And Distribution Of Organic Phosphorus Through A Grassland Catchment Transfer Continuum*. Lancaster University. 200p.
A67	Whitehead, P.G., L. Jin, J. Crossman, S. Comber, P.J. Johnes, P. Daldorph, N. Flynn, A.L. Collins, D. Butterfield, R. Mistry, R. Bardon, L. Pope, and R. Willows. 2014. "Distributed and dynamic modelling of hydrology, phosphorus and ecology in the Hampshire Avon and Blashford Lakes: Evaluating alternative strategies to meet WFD standards." *Science of the Total Environment* 481(1):157–166.
A68	Yates, C.A., and P.J. Johnes. 2013. "Nitrogen speciation and phosphorus fractionation dynamics in a lowland Chalk catchment." *Science of the Total Environment* 444:466–479. Available at:

http://dx.doi.org/10.1016/j.scitotenv.2012.12.002.

A69	Zhang, Y., A.L. Collins, and R.D. Gooday. 2012. "Application of the FARMSCOPER tool for assessing agricultural diffuse pollution mitigation methods across the Hampshire Avon Demonstration Test Catchment, UK." *Environmental Science and Policy* 24:120–131. Available at: http://dx.doi.org/10.1016/j.envsci.2012.08.003.
A70	Zhang, Y., A.L. Collins, and R. Hodgkinson. 2016. "Use of farm survey returns from the Demonstration Test Catchments to update modelled predictions of sediment and total phosphorus loadings from subsurface drains across England and Wales." *Soil Use and Management* 32(June):127–137.
A71	Zhang, Y., A.L. Collins, P.J. Johnes, and J.I. Jones. 2017. "Projected impacts of increased uptake of source control mitigation measures on agricultural diffuse pollution emissions to water and air." *Land Use Policy* 62:185–201. Available at: http://dx.doi.org/10.1016/j.landusepol.2016.12.017.
A72	Zhang, Y., A.L. Collins, J.I. Jones, P.J. Johnes, A. Inman, and J.E. Freer. 2017. "The potential benefits of on-farm mitigation scenarios for reducing multiple pollutant loadings in prioritised agri-environment areas across England." *Environmental Science and Policy* 73(April):100–114. Available at: http://dx.doi.org/10.1016/j.envsci.2017.04.004.
A73	Lloyd, C.E.M., J.E. Freer, P.J. Johnes, G. Coxon, and A.L. Collins. 2016. "Discharge and nutrient uncertainty: Implications for nutrient flux estimation in small streams." *Hydrological Processes* 30(1):135–152. [DUPLICATE of A40]
A74	Whitehead, P.G., L. Jin, J. Crossman, S. Comber, P.J. Johnes, P. Daldorph, N. Flynn, A.L. Collins, D. Butterfield, R. Mistry, R. Bardon, L. Pope, and R. Willows. 2014. "Distributed and dynamic modelling of hydrology, phosphorus and ecology in the Hampshire Avon and Blashford Lakes: Evaluating alternative strategies to meet WFD standards." *Science of the Total Environment* 481(1):157–166. DUPLICATE of A67
A75	Yates, C.A., P.J. Johnes, and R.G.M. Spencer. 2016. "Assessing the drivers of dissolved organic matter export from two contrasting lowland catchments, U.K." *Science of the Total Environment* 569–570:1330–1340. Available at: http://dx.doi.org/10.1016/j.scitotenv.2016.06.211.
AC1	Price, P.N., Y.S. Zhang, and A. Collins. 2013. "Identification of basic measures to address agriculture's impact on water. Final report." *Environment Agency*, 104pp.
AC4	DTC. 2012. "Creating and applying robust information resources for catchment management." *Learning from the Demonstration Test Catchments, Note No1* (01). 6p.
AC5	Joyce, J. 2013. "DTC Programme, Mini Ex Post Evaluation and possible Ex Ante Indicators." *DTC Research Advisory Group.* 10p.
AC6a	Williamson, A. 2013. "Helping knowledge exchange between the Demonstration Test Catchments and CSF." Defra Programme Management Group (paper 13 for Meeting 130220). 2p.
AC6b	Grant, F., A. Collins, R. Fish, M. Winter, M. Lobley, and C. Nolden. 2013. "Catchment Sensitive Farming and Demonstration Test Catchments Research Link Project." 21p.
AC7	DTC. 2013. "What part are the Demonstration Test Catchments playing in addressing water pollution?" *Learning from the Demonstration Test Catchments, Note No2* (02).
AC8	DTC. 2013. "What part are the Demonstration Test Catchments playing in addressing water pollution?" *Learning from the Demonstration Test Catchments, Note No2* (02) DUPLICATE of AC7.
AC9	DTC. 2014. "Specification of requirement for DTC Phase 2." *DTC Research Advisory Group.* 21p.
AC10	DTC. 2014. "Summary report on Phase 1." *DTC Research Advisory Group,* 9pp.
AC11	DTC. 2014. "Summary of emerging evidence from the Demonstration Test Catchments (DTC) Platform: Phase 1 Final Report." *DTC Research Advisory Group,* 265pp.
AC12	Rees, Y., D. Kingsley-Rowe, P. Cryle, and K. Conlan. 2015. "Evaluation of the Catchment Based

	Approach: Phase 2 Final report: WT1559." A report of research carried out by CASCADE Consulting, on behalf of the Department for Environment, Farming and Rural Affairs (DEFRA), 44pp. Available at: http://randd.defra.gov.uk/Default.aspx?Menu=Menu&Module=More&Location=None&Completed=0&ProjectID=19337.
AC13	DTC. 2015. "Demonstration Test Catchments: Improving Water Quality in Rural Areas." *National DTC Conference 10th June 2015, Stoneleigh Park, Warwickshire. DEFRA, DTC*. 16 p.
AC14	DTC. 2015. "Summary of emerging evidence from the Demonstration Test Catchments (DTC) Platform: Phase 1 Final Report." 265p.
AC15	Vrain, E., A. Lovett, A. Inman, R. Wheeler, M. Winter, A. Collins, W. Cleasby, and C. McIlwraith. 2017. "LM 0304: Demonstration Test Catchments Phase 2, Work Package 3: Working with Stakeholders and Influencing Behaviour Change. Final Report." 61p.
AC16	DTC. 2017. "Supplementary document to Work Package 3: Working with Stakeholders and Influencing Behaviour Change. Final Report. Phase 2."
AC17	Lovett, A., E. Vrain, L. Noble, R. Wheeler, M. Winter, W. Cleasby, and S. Burke. 2016. "Anaerobic Digestion and Water Quality, A Demonstration Test Catchments Response." 15pp.
AC18	Palmer, R.C. 2016. "Preliminary soil investigation into runoff, soil erosion, landslides and flooding on the Cumbrian Fells triggered by storm Desmond in December 2015." 40pp.
AC19	Vrain, E. 2016. "Barriers to land use change: From a DTC survey with 53 farmers in the Eden, Wensum and Tamar catchments." 3p.
AC20	DTC. 2012. "Creating and applying robust information resources for catchment management." *Learning from the Demonstration Test Catchments, Note No1* (01).
AC21	Blackwell, M., C. Hodgson, A. Collins, Y. Zhang, B. Surridge, N. Barber, C. Benskin, G. Mcshane, R. Cooper, K. Hiscock, A. Lovett, S. Burke, A. Aftab, and D. Harris. 2018. "LM 0304: Demonstration Test Catchments Phase 2 Work Package 2: Planning and Implementing Mitigation Interventions; WQ 0225: Implementing component 2 of Demonstration Test Catchments – to test integrated diffuse pollution mitigation measures for arable." 151p.
AC22	DTC. 2013. "Newsletter February 2013." *DTC Newsletters.* 6p.
AC23	DTC. 2013. "Newsletter June 2013." *DTC Newsletters*. 10p.
AC24	DTC. 2013. "Newsletter December 2013." *DTC Newsletters*. 15p.
AC25	DTC. 2014. "Newsletter June 2014." *DTC Newsletters.* 11p.
AC26	DTC. 2014. "Newsletter October 2014." *DTC Newsletters.* 10p.
AC27	DTC. 2014. "Newsletter December 2014." *DTC Newsletters.* 16p.
AC28	DTC. 2015. "Newsletter March 2015." *DTC Newsletters.* 17p.
AC29	DTC. 2015. "Newsletter June 2015." *DTC Newsletters.* 20p.
AC30	DTC. 2015. "Newsletter December 2015." *DTC Newsletters.* 16p.
AC31	DTC. 2016. "Newsletter June 2016." *DTC Newsletters.* 15p.
AC32	Truner, C. 2016. "What is the cost to a farm of diffuse pollution?". Video available at: https://vimeo.com/153293245
AC34	DTC. 2016. "Demonstration Test Catchments Platform Programme Plan." DEFRA report. 46p.
AC35	DTC. 2015. "DTC impact pathway."
AC36	DTC. 2013. "Programme Context, Drivers, Objectives, Approaches and Outputs." 9p.

AC44	DTC. 2017. “Phase 2: WP4: Approaches to upscale and extrapolate DTC outputs.” *DTC Research Advisory Group Report*. 80p.
AV1	Tamar DTC. 2013. “Tamar DTC Sub-catchments Water Quality: 2012 summary.” 2p.
AV2	Avon DTC. 2014. “Mitigation Measures Summary, Hampshire Avon DTC, The River Sem, Priors Farm sub-catchment.” D*emonstrating Catchment Management,* 2p.
AV3	Eden DTC. 2014. “Mitigation Measures Summary, Eden DTC, Morland mitigation sub-catchment.” 2p.
AV4	Wensum DTC. 2014. “Mitigation Measures Summary, Wensum DTC, Blackwater Drain (Mini-catchments A, B and E).” 2p.
AV5	DTC. (n.d.). *131015_Avon_Wylye_Quality_Summary_2012*. 2.
AV6	DTC. (2013). *140117_Avon_Ebble_quality_Summary_2012*.
AV7	DTC. (2013). *140115_Avon_Sem_Quality_summary_2013*.
AV8	Avon DTC. 2014. “Farm Baseline Survey.” 1p.
AV9	Avon DTC. 2013. “Working together to mitigate diffuse pollution.” *Avon Factsheet,* 1p.
AV10	Bidulph, M. 2013. “EFFECTIVE MITIGATION OF POLLUTION IN AGRICULTURAL CATCHMENTS.” *Avon Factsheet*:2p.
AV11	DTC. 2016. “The Avon and Tamar - Influencing farmer uptake of water pollution mitigation measures.” *Policy Brief, November 2016,* 2p.
E1	Ockenden, M., N. Barber, and P. Haygarth. 2016. “Storm Desmond in the Eden DTC catchments : 4th – 6th December 2015.” 9p.
E2	Eden DTC. 2012. “Morland Catchment Water Quality: Hydrological Year 2012.” *Eden DTC InfoSheet*: 2p.
E9	Eden DTC. 2016. “The Eden: Influencing farmer uptake of water pollution mitigation measures.” *Eden DTC Policy Brief, Phase 2: Work Package 3*:2p.
E10	Suslovaite, V., S. Reaney, and T. Burt. 2019. “Temporal dynamics and projected future changes in the nitrate leaching in a small river catchment dominated by under-drained clay soil grasslands: analysis of high-frequency monitoring data.” In *Poster presented at the Eden DTC Conference*. p. 1.
E11	Jonczyk, J.C., G.J. Owen, M.A. Snell, N. Barber, C. Benskin, S.M. Reaney, P. Haygarth, P.F. Quinn, P.A. Barker, A. Aftab, S. Burke, W. Cleasby, B. Surridge, and M.T. Perks. 2019. “A New Approach To Soil Sampling For Risk Assessment Of Nutrient Mobilisation.” In *Poster presented at the Eden DTC Conference*. Available at: http://adsabs.harvard.edu/abs/2016AGUFM.H33J1695J.
E12	Reaney, S.M., P. Barker, P. Haygarth, P. Quinn, A. Aftab, N. Barber, C. Benskin, S. Burke, W. Cleasby, J. Jonczyk, G. Owen, M. Perks, M. Snell, B. Surridge, and E. Team. 2019. “Mitigating Agricultural Di use Pollution: Learning from The River Eden Demonstration Test Catchment Experiments.” In *Poster presented at the Eden DTC Conference*. DUPLICATE of A55
E13	Reaney, S.M., P. Barker, P. Haygarth, P. Quinn, A. Aftab, N. Barber, C. Benskin, S. Burke, W. Cleasby, J. Jonczyk, G. Owen, M. Perks, M. Snell, B. Surridge, and E. Team. 2019. “Mitigating Agricultural Di use Pollution: Learning from The River Eden Demonstration Test Catchment Experiments.” In *Poster presented at the Eden DTC Conference*. DUPLICATE of A5
E14	Snell, M.A., P.A. Barker, P.M. Haygarth, P. Quinn, A. Aftab, N. Barber, C.M.H. Benskin, S. Burke, W. Cleasby, J. Jonczyk, G.J. Owen, M.T. Perks, S.M. Reaney, B.W.J. Surridge, and EdenDTC team. 2019.

	"Spatial-temporal dynamics in phytobenthos structural properties reveal insights into agricultural catchment dynamics and nutrient fluxes." In *Poster presented at the Eden DTC Conference*.
O2	Davey, A. 2010. "Demonstrating Test Catchments: An Experimental Design and Monitoring Strategy." 42p.
O4	DTC. 2010. "Demonstration Test Catchments: Development of a Data Model. Overview statement." 2p.
O5	Lawrence, B., A. Stephens, A. Woolf, G. Rees, and J. Doughty. 2010. "Demonstration Test Catchments – Data Management Requirements." 53p. Available at: http://home.badc.rl.ac.uk/lawrence/static/2010/06/22/dtc-report.pdf.
O7	Robinson, M., and H. Gibson. 2011. "Assessing the status of drainage in the UK - A Case Study of the Avon, Eden and Wensum DTCs and their representatives with regional and national patterns." 12p. Available at: http://nora.nerc.ac.uk/id/eprint/14118/1/ADAS-Defra_Drainage_in_DTCs.pdf.
O11	DTC. 2016. "Summary of emerging evidence from the Demonstration Test Catchments (DTC) Platform: Phase 1 Final Report." 265p.
O12	Demonstration Test Catchment. 2015. "Demonstration Test Catchments: Improving Water Quality in Rural Areas. Phase 1 final report." 16p.
O13	Winter, M. 2014. "Extending Knowledge Exchange in the Avon DTC: Report on Activities 2010-2014." Avon DTC, 19p.
O14	Dockerty, T., A. Lovett, K. Hiscock, and L. Noble. 2014. "Extending Knowledge Exchange in the Wensum DTC: Report on Activities 2010-2014." Wensum Alliance. 17p.
O15	Cleasby, W. 2014. "Extending Knowledge Exchange in the Eden DTC: Report on Activities 2010-2014." Eden DTC, 18p.
W1	Cooper, C.R., K. Hiscock, A. Lovett, and S. Dugdale. 2019. "Investigating the impacts of cover crops and reduced tillage regimes on water quality, soil health and crop yields, Winter 2018/19 update." *Report for DEFRA (Department for Environment, Food and Rural Affairs) by Wensum Demonstration Test Catchment Project.* 24p.
W2	Wensum DTC. 2017. "River Wensum DTC, Research Update 1." *Wensum Research Summaries*. 4p.
W3	Wensum DTC. 2017. "River Wensum DTC, Research Update 2." *Wensum Research Summaries*. 4p.
W4	Wensum DTC. 2017. "River Wensum DTC, Research Update 3." *Wensum Research Summaries*. 3p.
W5	Wensum DTC. 2017. "River Wensum DTC, Research Update 4." *Wensum Research Summaries*. 3p.
W6	Wensum DTC. 2019. "River Wensum DTC, Research Update 5." *Wensum Research Summaries.* 3p.
W7	Vrain, E. 2016. "Barriers to land use change." *Wensum Research Summaries*. 3p.
W8	Wensum Alliance. 2014. "Understanding water quality issues using high- frequency nitrate and phosphate data in the Wensum Demonstration Test Catchment." *Wensum Research Summaries*. 2p.

W9	Wensum Alliance. 2014. "Operational testing of a biobed for pesticide removal in the Wensum Demonstration Test Catchment." *Wensum Research Summaries*. 2p.
W10	Wensum Alliance. 2014. "Experimentation with cover cropping in the Wensum Demonstration Test Catchment." *Wensum Research Summaries*. 2p.
W11	DTC. 2014. "Attitudes of Farmers Towards Diffuse Pollution Mitigation Measures in the Demonstration Test Catchments." *Wensum Research Summaries*. 2p.
W12	DTC. 2014. "The Niche of Catchment Sensitive Farming in the Provision of Advice to Farmers." *Wensum Research Summaries*. 2p.
W13	DTC. 2014. "The Role of Farm Advisors in Improving the Uptake of Water Quality and Other Environmental Measures." *Wensum Research Summaries*. 2p.
W14	Outram, F. 2013. "Water quality summary for the Blackwater: The Wensum Demonstration Test catchment hydrological year 2011-12." *Wensum Research Summaries*. 2p.
W15	Cooper, R., T. Krueger, K. Hiscock, and B. Rawlins. 2013. "Fluvial Sediment Tracing in the Wensum DTC." *Wensum Research Summaries*. 13p.
W16	Vrain, E., A. Lovett, L. Noble, F. Grant, P. Blundell, and W. Cleasby. 2013. "Attitudes of Farmers Towards Diffuse Pollution Mitigation Measures in the Demonstration Test Catchments." *Wensum Research Summaries*. 1p.
W17	Hama-aziz, Z., C. Adams, and K. Hiscock. 2013. "Assessment of the effects of reduced cultivation practices on soil nutrient and nitrous oxide losses in arable soils." *Wensum Research Summaries*. 1p.
W 18	Environment Agency. 2013. "River Wensum Restoration Strategy, Swanton Morley Restoration Scheme – Reach 14a." *Factsheet*. 7p.
W19	Carrick, J. 2013. "Environmental Stewardship at Castle and Park Farms, Swanton Morley, Norfolk." *Wensum Research Summaries*. 1p.
W20	Collier, S. 2012. "Application of the FARMSCOPER Decision-Support Tool to the Wensum Catchment." *Wensum Research Summaries*:1p.
W21	Wensum Alliance. 2012. "Cost-Effectiveness Analysis (CEA) using FARMSCOPER." *Wensum Research Summaries*. 1p.
W22	Bird, E., and S. White. 2012. "Evaluating the Accuracy and Usability of Low-Cost Phosphorus Testing Kits." *Wensum Research Summaries*:1p.
W23	Vrain, E., and L. Noble. 2012. "Farm Business Survey." *Wensum Research Summaries*. 1p.
W24	Wensum Alliance. 2012. "High-Resolution Monitoring of Stream Water Quality." *Wensum Research Summaries*. 1p.
W25	Howson, T. 2012. "Hydrograph Separation using Stable Isotopes." *Wensum Research Summaries*:1p.

W26	Read, T., V. Bense, and A. Guéron. 2012. "Identifying Flow Pathways with Fibre Optic Distributed Temperature Sensing (FO-DTS)." *Wensum Research Summaries*. 1p.
W27	Albaggar, A. 2012. "Investigation of bacterial community composition and abundance of the River Wensum catchment." *Wensum Research Summaries*. 1p.
W28	Cooper, R., and Rawlins. 2012. "Investigating Organo-Mineral Suspended Sediment Controls on Catchment Phosphorus Export." *Wensum Research Summaries*. 1p.
W29	Wensum Alliance. 2012. "River Wensum Water Quality Monitoring." *Wensum Research Summaries* (March). 1p.
W30	Taigel, S., and K. Appleton. 2012. "Visioning Catchment Futures: Bringing the Landscape to Life." *Wensum Research Summaries*:1p.
W31	Al-Yami, M. 2012. "Visualization of Digital Elevation Data for Catchment Modelling." *Wensum Research Summaries.* 1p.
W32	Wensum Alliance. 2010. "Fact Sheet 1: THE WENSUM DEMONSTRATION TEST CATCHMENT PROJECT." *Wensum Factsheets*. 2p.
W33	Wensum Alliance. 2010. "Fact Sheet 2: About the River Wensum." *Wensum Factsheets*. 2p. Available at: http://www.wensumalliance.org.uk/wensum.html.
W34	Wensum Alliance. 2010. "Fact Sheet 3: Pollution issues in the Wensum catchment." *Wensum Factsheets*:2p.
W35	Wensum Alliance. 2010. "Fact Sheet 4: WHAT MEASURES CAN BE TAKEN TO REDUCE DIFFUSE POLLUTION?" *Wensum Factsheets*. 2p.
W36	Wensum Alliance. 2010. "Fact Sheet 5: Monitoring: What Is Involved?" *Wensum Factsheets*. 2p.
W37	Wensum Alliance. 2010. "Fact Sheet 6: Precision farming." *Wensum Factsheets*. 2p.
W38	Wensum Alliance. 2012. "Fact Sheet 7: NITRATE AND PHOSPHATE: a problem for water quality?" *Wensum Factsheets*. 2p.
W41	Wensum Alliance. 2011. "Protecting the River Wensum: have your say." *Wensum Leaflet*. 1p.
W42	Wensum Alliance. 2016. "Sharing Experience of Integrated Catchment Management." In *Demonstration Test Catchment Project National Conference*.
W94	Wensum DTC. 2016. "The Wensum - Influencing farmer uptake of water pollution mitigation measures. Phase 2: Work Package 3." *Wensum Policy Brief*.

Appendix B: DTC policy-relevant questions

A1.1 Phase 1

The following 15 policy-relevant questions, being asked of the DTC researcher community, were presented in the Phase 1 summary report:

1) Which parts of the country are currently failing WFD targets due to agriculture and what are the main reasons for failure to achieve good water quality and ecological status?

2) How does ecology respond to seasonal pollutant fluxes? How important are short term pollution concentration 'spikes' and seasonal trends (i.e. pollutant losses during winter when little is growing compared with the, more biologically active, summer)?

3) How long does it take for pollutants to travel from their source to their point of impact? How will this affect our chances of meeting WFD targets for different catchments? This includes: nutrient cycling (interactions with biota), sedimentation and remobilisation, attenuation, and transformations (e.g. denitrification).

4) To what extent are pollutants removed from the environment e.g. by burial in sediment, loss to the atmosphere or chemical breakdown (e.g. of pesticides)?

5) How important are storm events in affecting water quality in different parts of the country?

6) What are the most cost-effective measures for a given catchment?

7) How can measures be effectively targeted within a catchment?

8) Which combinations of measures work best and how do they interact?

9) What is the attitude of farmers towards acceptance of different measures? What would motivate farmers to undertake such measures and what are the main things likely to put them off?

10) What support do they require (financial, technical, guidance) to undertake such work to protect the environment?

11) What effects does increasing self-monitoring have on farmers' attitudes to taking up measures?

12) Which are the most cost-effective monitoring technologies and investigative techniques?

13) How can we make monitoring and modelling approaches easier to use by catchment managers?

14) What tools are required by the new generation of catchment managers, what is available and what technical skill is required to use them?

15) How can we make better use of emerging knowledge to inform decisions by policy makers, catchment managers and farmers?

A1.2 Phase 2

The specification for Phase 2 of the DTC project specifies the following questions organised under 4 workpackages:

Work package 1: Understanding the nature of the problem (catchment function and response)

1.1 Sources: Identification and quantification of agriculture pollution sources

What are the main agricultural sources that bring potential polluting substances into the catchments (nutrients, FIOs, sediment, [pesticides])

1.2 Mobilisation: Identification of solubilisation, detachment and incidental mobilisation from the soil

How can we best assess the risk of solubilisation, detachment and incidental mobilisation from the soils in the DTCs?

Can we determine the conceptual basis upon which these are scaled up, to other catchments and soils?

1.3 Delivery: Pollutant transport pathways and transformations

What are the most important pathways of pollutant transfer from agricultural sources to water bodies and how can they be identified?

What is the effect of natural physical, chemical and biological processes on the timing of pollutant delivery to water bodies and the overall impact of diffuse pollution?

What are the main diffuse pollution risk factors and how do they vary spatially and temporally in terms of the likelihood of pollutant loss from farming?

How can the key pollution hot-spots (often called critical source areas or CSAs) be identified within catchments and how are they spatially distributed?

1.4 Impacts: Water body (receptor) response and other socio economic impacts

What is the contribution of different agricultural pollution sources on water quality in surface and subsurface water bodies?

What is the ecological response to such pollutant inputs in receiving surface water bodies? (this should take into account interactions with other factors, such as the morphology of the river channel)

How resilient are water quality and ecological response of water bodies to temporal variations in pollution losses (from storm-event to seasonal trends)?

What is the social and economic cost of diffuse agricultural pollution for individual water bodies in the sub-catchment?

1.5 Extrapolating to the wider catchment and nationally

How much diffuse agricultural pollution in the DTC headwater catchment impact the overall economic cost of not achieving good water body status at the river basin scale?

What is the economic cost of the transfer of pollutants from the DTC headwater catchments further down gradient ?

What is the relevance of DTC findings to other catchments and how can they be applied?

Work Package 2 - Planning and implementing mitigation interventions

2.1 Measure cost, design and maintenance

What is the lifespan of, and what are the maintenance requirements for, different mitigation measures?

What are the direct and indirect costs of measures, including implementation, maintenance, impact on productivity/ profitability and savings due to more effective resource management, and to whom do these costs fall?

2.2 Environmental outcomes of mitigation

How can you extrapolate mitigation efficacy from a single measure to a series of interconnected measures?

How much of an improvement in terms of receptors can we attain by implementing mitigation?

To what extent does targeting only part of the anthropogenic pressures on water bodies allow us to achieve significant improvement in status?

How long will it take for mitigation interventions to: (i) meet a set pollutant threshold (WFD related targets) at a given point in a catchment and (ii) achieve an ecological response?

2.3 Cost-effectiveness of mitigation measures

What is the benefit:cost ratio for the various measures applied in the DTC catchments? (This should account for all downstream benefits of improved water quality from improvements in the DTC headwater study areas)

What is the comparative cost-effectiveness of implementing measures on a targeted basis compared with blanket implementation? (i.e. how does benefit:cost ratio vary spatially for different measures?)

What is the scale of land use and land management change that is needed to achieve EU-WFD targets of good water body status at the DTC catchment scale?

How might such improvements contribute to achieving good water body status and reduce environmental damage costs at the river basin scale?

2.4 Designing an agricultural pollution mitigation strategy at the catchment scale

What are the logical steps to identify pressures and plan a programme of measures within a catchment? (Developing a treatment-train approach)

How can you determine the level of measure coverage needed to achieve EU-WFD objectives in different catchments?

How can you optimize cost-effectiveness when implementing measures (i) at the catchment scale, (ii) at the river basin scale and (iii) nationally?

Work Package 3 - Working with Stakeholders and Influencing Behaviour Change

3.1 Current Practices

What is the current baseline level of practice in terms of diffuse pollution mitigation?

3.2 Behaviour/Attitudes, Support and Collaboration

What are farmers' current attitudes towards diffuse pollution measures and how likely are they to adopt them?

Which factors motivate farmers to adopt measures and what are the main things likely to put them off?

What are the practical constraints to implementing measures within the context of a farm business?

What consequences such patterns have for policy levers?

Subsidiary questions:

What level of technical support do land managers require to adopt measures?

What level of financial support do farmers need to implement different types of measure?

Which policy interventions are best adapted to encourage the uptake of different types of measures?

How can farmers be encouraged to collaborate to implement measures strategically at a catchment or sub-catchment scale?

3.3 Developing catchment scale stakeholder groups

What governance arrangements are needed to implement a catchment-based approach?

DTC Work package 4: Developing improved monitoring and research techniques to inform, monitor and evaluate policy and extend DTC outcomes to other catchments

4.1 Improving focus and approaches to monitor and quantify agricultural diffuse pollution and impact on water bodies

What best knowledge exchange frameworks and models allow for ensuring research answers policy and other catchment manager key needs?

Which are the most cost-effective monitoring technologies and investigative techniques for (i) identifying pressures in a catchment, (ii) undertaking source apportionment, and (iii) detecting the effects of pollution mitigation?

How can monitoring be optimally deployed, and what is the minimum amount of data needed to detect a reduction in diffuse pollution and an ecological response at given spatial and temporal resolutions?

How can monitoring and modelling tools be practically used in combination to build 'weight of evidence' to inform catchment management?

4.2 Developing approaches to up-scale and extrapolate DTC outputs

How can results of plot-scale research be reliably up-scaled and applied to inform decision making in heterogeneous landscapes?

How can we infer impact of processes monitored in headwater catchments to those occurring downstream (i.e. integrating mid-catchment and lowland areas)?

How can we translate outputs of catchment-specific studies to other catchments nationally?

Appendix C: Scoring robustness of physical science

Reporting:

R1_What type of publication is used to report results?

Score	Criteria
3 (Excel)	Peer reviewed
2 (Good)	Report with signed off
1 (Satis)	Other: Not peer reviewed or signed off

Method (Modelling)_Mo:

Mo1_Has uncertainty been reported?

Score	Criteria
3 (Yes)	Yes
2	
1 (No)	No consideration of uncertainty

Mo2_Have authors undertaken calibration and or validation?

Score	Criteria
3 (Excel)	Model not calibrated but validated
2 (Good)	Model calibrated and validated
1 (Satis)	Model not validated but may or may not be calibrated

Mo3_Is there peer reviewed evidence for the validity of the model outside of the DTC?

Score	Criteria
3 (Yes)	Yes
2	
1 (No)	No

Mo4_Is the model time step sufficiently fine to represent key processes?

Score	Criteria
3 (Excel)	Hourly or less
2 (Good)	Hourly – Daily
1 (Satis)	>daily

Mo5_Is the testing period of sufficient duration?

Score	Criteria
3 (Excel)	>2.5 years
2 (Good)	1.5 – 2.5 years
1 (Satis)	<1.5 years

Mo6_Is the period of scenario simulation of sufficient duration?

Score	Criteria
3 (Excel)	>2.5 years
2 (Good)	1.5 – 2.5
1 (Satis)	<1.5 years

Method (Monitoring)_M:

M1_Are hydrological measurements considered in conclusions?

Score	Criteria
3 (Excel)	Yes and co-located
2 (Good)	Yes but not co-located
1 (Satis)	Not considered

M2_Is automated monitoring used and is it ground truthed (at least 2 weekly)?

Score	Criteria
3 (Excel)	Yes and ground truthed 2 weekly or more frequent
2 (Good)	Yes but ground truthed less frequently than 2 weekly or not specified
1 (Satis)	Yes but not ground truthed / automated monitoring not used / ground truthing not mentioned

M3_Are laboratory analytical procedures used and reported?

Score	Criteria
3 (Excel)	Yes and reported
2 (Good)	Yes but not reported
1 (Satis)	No - only field determinations used

M4_Is regular monitoring of sufficient temporal resolution?

Score	Criteria
3 (Excel)	hourly or less
2 (Good)	daily to weekly
1 (Satis)	more than weekly

M5_Does sampling target storm events?

Score	Criteria
3 (Excel)	Yes with autosamplers
2 (Good)	Yes but manually
1 (Satis)	No

M6_Have baseline conditions been adequately characterised?

Score	Criteria
3 (Yes)	there is >1yr of baseline monitoring
2	
1 (No)	there is <1yr of baseline monitoring

M7_Replicated used?

Score	Criteria
3 (Yes)	Yes
2	

1 (No)	No

M8_ Sufficient length of intervention period (at least years)?

Score	Criteria
3 (Excel)	More than 2 years
2 (Good)	1-2years
1 (Satis)	Less than 1 yr

M9_Is intervention close to monitoring location?

Score	Criteria
3 (Excel)	Yes lower catchment or catchment wide
2 (Good)	Middle catchment or mixed locations
1 (Satis)	No upper catchment or not specified

M10_What experimental design is adopted?

Score	Criteria
3 (Excel)	BACI1 or BACI2 design
2 (Good)	With and without intervention
1 (Satis)	Paired catchment or trend at impacted site

Results_Re:

Re1_Is intervention design reported?

Score	Criteria
3 (Yes)	Yes
2	
1 (No)	No

Re2_Is intervention management reported?

Score	Criteria
3 (Yes)	Yes
2	
1 (No)	No

Re3_Are confounding factors taken into account?

Score	Criteria
3 (Excel)	Stated that there are none or they are identified and quantified
2 (Good)	Identified but only taken into account conceptually
1 (Satis)	Not mentioned or mentioned and not taken into account

Re4_Is the significance of the results reported?

Score	Criteria
3 (Yes)	Yes
2	
1 (No)	No

Re5_Are results relevant to or related to environmental quality standards?

Score	Criteria
3 (Excel)	Results are related to environmental quality standards
2 (Good)	Results are relevant to EQS
1 (Satis)	Results are not relevant or related to EQS

Re6_Are results given for individual years post intervention?

Score	Criteria
3 (Excel)	Yes
2 (Good)	A multi year trend is identified
1 (Satis)	No – 1 year or less

Appendix D: Scoring robustness of social science

All types

Is the methodology(ies) clearly described and could it/they be repeated?

1. Not described
2. Described but insufficient detail to replicate
3. Described with sufficient detail (can include references to further documents for more information)

Data Collection: Were the data-collection methods appropriate for studying engagement, behavioural change and/or cost-effectiveness of agricultural practices?

Qualitative[4] data collection

Free text descriptive column of what the research method(s) was including whether one-off or timeseries

Is the sample size and structure suitable for the aims and questions?

1. No comment on sample type or size
2. Sample type and size described but no comment on saturation[5]
3. Sample type and size described and saturation claimed

Does the data collection follow good practice[6]?

1. No evidence for how implemented provided
2. Evidence provided but does not follow good practice
3. Data collection follows good practice

Quantitative[7] (not related to cost-benefit/effectiveness analysis) data collection

Free text descriptive column of what the research method(s) was including whether one-off or timeseries

---

[4] Qualitative research covers a family of data collection and analysis approaches that focus on non-numerical data and interprets meaning to patterns of these data.

[5] Saturation means that the researchers feel they have sufficient material to understand and explain the patterns in their data

[6] There might be a wide range of qualitative research methods including interviews, participant observation, focus groups or workshops, participatory mapping, video etc. Good practice varies but the shared characteristics is that the method has been used before or if an innovation, has been piloted first, that the protocol for data collection is available (interview guide, workshop plan), and actual process of data collection is recorded, including where the approach was adapted to respond to specific circumstances.

[7] Quantitative research collects countable (normally numeric) data that is analysed for statistical patterns (including descriptive statistics)

Are the instruments used for data collection appropriately tested?

1. No piloting or testing
2. Limited piloting and testing
3. Clear description of piloting and testing
4. Piloting / testing not relevant to method (e.g. secondary data analysis)

Is the sample size and structure suitable for the aims and questions?

1. No comment on sample type or size
2. Sample type and size described but no comment on suitability for statistical analysis selected
3. Sample type and size described and appropriate for statistical analysis used

Does the instrument implementation follow good practice[8]?

1. No evidence for how implemented provided
2. Evidence on instrument implementation provided but does not follow good practice
3. Instrument implementation follows good practice
   - Not relevant (e.g. secondary data analysis)

Cost-Effectiveness and cost benefits data collection

Free text descriptive column of what the research method(s) was, including whether one-off or timeseries; physical or monetary units used

Choice of costs and benefit flows included / excluded discussed

1. No justification of Costs and benefits included
2. Some justification
3. All potential costs and benefits are mapped, and inclusion / exclusion is fully justified

Costs benefits / effectiveness measures:

1. No source of data mentioned
2. Valuation of costs or benefits using secondary sources or qualitative assessment
3. Primary data collection to provide monetary values of costs and benefits (e.g. stated preferences approaches for non-market costs and benefits)

If primary data collection, need to use evaluation criteria from "quantitative research"

If secondary data used:

---

[8] Good practice would cover the protocol for data collection is available (questionnaire) and recording details of actual process of data collection (e.g. was the questionnaire online or by phone, how many follow-ups were attempted etc)

1. No discussion of the relevance of source of the secondary data to the case studied
2. Source of secondary data is acknowledged but relies on few and or irrelevant other case studies
3. Source of secondary data acknowledged, using relevant data or data adjusted to case study of interest (e.g. benefit transfer relying on meta-analysis and WTP adjusted to local characteristics)

Data Analysis: Were the reported results based on appropriate analytical approaches for studying engagement, behavioural change and/or cost-effectiveness of agricultural practices?

All types

Does the study provide information limitations and confounding factors influencing results?
1. No information provided
2. Limitations and confounding factors noted but not connected to findings
3. Limitations and confounding factors impact on findings acknowledged (and quantified where appropriate)

Qualitative[9] data analysis

Do the results cover both majority and minority views; and illustrate the range of opinions?
1. Results do not distinguish between participants views
2. Results not differences in passing but do not discuss
3. Results clearly present different views and a full range of opinions

Is theory used to analyse the data?
1. Themes are described but not used to test or build theory
2. Themes are described and generally linked to relevant literature but not theory building or testing
3. Themes are used to build new theory or challenge/confirm existing theory

Quantitative (not CBA or CEA) data analysis

Data analysis assessed by:

Does the analysis clearly relate to the starting theoretical framework/hypothesis?
1. Results are described but not related to any hypothesis
2. Results used to test starting theory/hypothesis

Are the correct statistical tests selected (e.g. non-parametric for non-normal distribution of data)
1. No description of tests provided

[9] Qualitative analysis focusses on making sense of data in terms of patterns or themes and relating them to theoretical frameworks or concepts.

2. Description of tests provided but not explained
3. Selection of tests explained and justified
- 0 for when tests are not appropriate for method

Are confidence levels and strength of association provided?
1. No confidence levels or association provided
2. Confidence levels and association provided but does not meet good practice (e.g. confidence below 90%; or association very weak)
3. Confidence levels and association provided that meet good practice

CBA or CEA data analysis
Has a sensitivity analysis been implemented?
1. No comment on sensitivity of results to assumptions made
2. Some discussion of the sensitivity of results to assumptions made
3. Full sensitivity analysis done for key parameters of the Cost benefit Analysis

Does the CBA/CE include flows of costs and benefits over the lifespan of the intervention?
1. The CBA/CE is based on a single year assessment and does not account for future flows
2. The CBA/CE partially accounts for future flows / discusses potential future flows.
3. The CBA/CE accounts for all expected flows of costs and benefits, present and future

If CBA over multiple years, how are future flows of costs and benefits discounted?
1. No discounting
2. A discount rate is used but its choice is not justified
3. Discount rate used, justified and follows good practice (HM Treasury's Green Book guidance is 3.5% in real terms, and then declining discount rate over time after 30 years)

Reporting - All types
Means of publication
1. Unreviewed Report or other
2. Peer reviewed report (went out to external reviewers but not published in journal)
3. Peer reviewed Scientific Paper

Discussion of how relevant findings are to other contexts
1. No discussion of relevance to other places
2. Some general recommendations but unclear why/how relevant to other places
3. Explicit comments on relevance to other places provided

Conclusions are backed up by well presented data and findings
1. Unclear how conclusions were drawn
2. Limited relationship between findings/data and conclusions drawn
3. Clear relationship explicitly stated between data/findings and conclusions drawn

Confidence Score
1. Very difficult document to evaluate quality of science and scores were hard to assign

2. Reasonable level of detail but still found it difficult to decide scores in some cases
3. Easy to assign scores based on clear detail in the document

Appendix E: Social Science Methodology Applied

Overall 'Population' of Documents

Of the 173 documents provided 77 were submitted to the Hutton team for analysis to answer the social science questions (having previously excluded PowerPoint presentations and documents describing the tenders for the research). Documents were given a unique identifier (ID) to distinguish them, given that there were often multiple similar citations (e.g. all the newsletters). Full citations are provided in Appendix B, whilst the report text uses the IDs to reference the documents.

Following the protocol, around 20% of all documents were checked by a second researcher, and where any divergence in classification occurred, a discussion took place and the entries updated if required. Furthermore, the PIs of the DTC projects are being given the opportunity to respond to draft findings and identify missing documents or amend extraction or robustness classifications. Whilst these steps will improve the quality of the review, we cannot claim to have complete coverage of all DTC outputs – we have reviewed all evidence made available to us within the time constraints of the project.

Researchers did not review the DTC Evidence Compendium (Defra, 2020) WT15116 prior to the evidence analysis process to avoid bias. A cross-check of the evidence presented by the DTC consortia with our findings was completed as part of the process of peer review described above.

Table E 1: Overall number of documents broken into categories[10] :

SPICE documents	Relevant Documents	Replicates	Discarded
A35	A45	A55	A1
A9	A26	AC10	A57
A11	A43	AC13	A71
AC15	A44	AC14	A72
AC21	A47	AC16	AC11
AC25	A64	AC19	AC20
AC26	A66	AC28	AC22
AV11	A70	AC29	AC24

[10] Appendix B provides a list of all documents with their unique IDS.

AV8	A74	AC31	AC34
E9	AC1	AC6a	AC36
O13	AC12	O16	W 18
O11	AC5	O12	W19
W21	AC4	O14	W35
W23	AC23	W4	
W42	AC27	W13	
	AC28	W16	
	AC32	W7	
	AC6b	W94	
	AC30		
	AV2		
	AC9		
	AV3		
	AV4		
	AV9		
	O15		
	W11		
	W12		
	W30		
	W32		
	W33		
	W41		
Total = 15	Total = 31	Total = 18	Total = 13

The DTC project was organised in three phases: Phase 1 (December 2009 – January 2015); Phase 2 (January 2015 – March 2018) and Phase 3 (April 2018 to March 2019). One might expect the results of relevance to the SPICE question "How effective were DTC engagement processes in DTC catchments in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?" to be mainly contained in the documents from the later phases. Currently there are only four documents in the dataset dated 2018 or 2019; and 27 from phase 2 (2015-2018). Of these post-2015 documents, only eight are SPICE documents (see next page). This means that there is low coverage of results coming out of phase three; and given the lag in academic publication, many of the later journal publications refer to evidence collected in phase 1 and phase 2 only. The cut -off for our evidence was all documents held by Defra in July 2019.

Documents remaining after screening for SPICE relevance

We assessed documents for SPICE relevance (providing evidence allowing us to answer the question "How effective were DTC engagement processes in DTC catchments in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?") - 29 documents were deemed SPICE compliant, but of these, 14 contained replicate information from other reports, leaving 15 documents to be reviewed for relevant evidence and robustness of their science. Many documents have multiple entries for in the spreadsheet, to allow differentiation between catchment, Diffuse Water Pollution (DWP) mitigation measures or other factors, as shown below.

Table E 2: Number of rows in extraction table per document

Row Labels	Count of Article reference
A35	5
A9	36
A11	1
AC15	17
AC21	14
AC25	2
AC26	4
AV11	1
AV8	1
E9	3
O13	7
O11	34
W21	4
W23	1
W42	1
Grand Total	131

Documents with robustness analysis

Robustness could only have been performed on those 15 documents that were SPICE compliant.

Robustness was evaluated based on good practice for data collection, data analysis and data reporting. However, different criteria were used dependent on whether the methodology was qualitative, quantitative or economic as it is important to compare the robustness of the science by appropriate criteria – for example the behavioural component does not have a quantitative design (AC5) so quantitative criteria would be inappropriate. Documents often used a mixture of approaches in which case a combination of criteria were applied.

However, only eight documents were assessed for robustness as shown below. The other documents did not have sufficient methodological information to permit robustness analysis. Those without robustness analysis were: AC25 (newsletter), AC26 (newsletter), AV11, AV8, W23, E9 (short 1 or 2 page summaries) and W42 (conference proceedings).

Table E 3: Results of Robustness Analysis

ID	Type	Types of Science	Overall Robustness	Confidence levels
W21	Non-reviewed Report	Economic	Satisfactory	Low
O11	Non-reviewed Report	Qualitative, Quantitative & Economic	Satisfactory	Medium
O13	Non-reviewed Report	Qualitative and Quantitative	Satisfactory	Medium
AC21	Non-reviewed Report	Quantitative and Economic	Good	High
AC15	Non-reviewed Report	Qualitative and Quantitative	Good	Medium
A35	Peer Reviewed Paper	Qualitative and Quantitative	Good	Medium
A9	Peer Reviewed Paper	Quantitative and Economic	Good	High
A11	Peer Reviewed Paper	Economic	Excellent	High

Average data collection Robustness score: Good. Weaker scores were for not providing information on the questionnaire piloting or testing; failing to give information on sampling strategy, size or representativeness, and limited detail on how data collection was carried out during the discussion groups etc. In some cases, e.g. O11 or O13, there was very little methodological information provided in the reports so low scores were assigned to most criteria, even though the content was extremely useful.

Average data analysis robustness score: Satisfactory. Weaker scores were for not providing information on limitations or confounding factors, insufficient use of theory or statistics to generalise results, lack of information on statistical or economic tests and inappropriate timescales for economic assessments. However, as with data collection, the insights from the analysis were often useful and interesting, but the document lacked evidence that good practice had been implemented.

Average data reporting robustness score: Good. Weaker scores were for not publishing the report in a peer-reviewed publication, not discussing the relevance of the material to other contexts and failing to clearly relate the findings to the overall conclusions or recommendations of the document. Of course, non-reviewed short reports with short or no discussion sections will have scored poorly in this category, but maybe very useful documents for non-scientific audiences.

Overall Average robustness score: Good. The overall confidence score was 1.89 – in most cases the reviewer was fairly confident in their ability to assess the document evidence (score two), and in a couple of cases, very confident but equally there were a couple of cases were the reviewer had low confidence in the ability to judge the robustness due to lack of information. The protocol for robustness scoring has been provided as a technical annex 2.

We also considered whether the documents recorded any 'confounding factors' influencing the implementation of the DTC research activities and none of the documents with robustness analysis actually reflected on this. AC26 did highlight the extension of phase 1 and change to a single DTC project in phase 2 during 2014-15, but this was not presented as influencing the implementation of the research. Indeed, only A9, AC21 and AC26 explicitly listed limitations to their study in their conclusions (limitations were also noted by A66, AC5 and AV4 relevant documents).

These comments do raise a question about robustness scoring for documents that were not intended for scientific audiences. For this reason, robustness scores were not allocated to documents with **no** methodological material, in order to distinguish between lower scores where the information suggested limitations and not absence of any information. The quest for conciseness and relevance may mean that technical methodological details are omitted in the reporting but could have been undertaken in practice. Furthermore, journal paper word restrictions often mean there is limited methodological detail provided in papers as well. Future projects may wish to provide a technical annex or separate report with more information to allow robustness evaluation to be carried out on the full suite of information.

Documents not SPICE but relevant

33 documents were deemed SPICE relevant, of which two are duplicates (AC6a and O14), leaving 31. This means that whilst the documents did not specifically provide information on the SPICE question, they had relevant insights for the relevance and robustness of the science undertaken within the three phases of the DTC project. These documents are therefore considered in sections Eight and Nine regarding general insights on farmer engagement and uptake of DWP measures.

Documents discarded and Replicate documents

Steering Group agreed to discard all PowerPoint presentations. Another 13 papers were considered neither SPICE compliant nor SPICE relevant for the following reasons:

- Conference Abstract – A1 (no further paper could be found – no results reported in abstract)
- 3 newsletters (AC 24, 22 and W18) that contained no social science engagement, uptake or cost-effectiveness results
- 3 scientific publications (A57, A72 and A71) – A57, A71, A72 did not report on specific DTC findings but rather reported on mitigation at the national UK scale.
- 6 reports (W19, AC36, W35, AC4, AC11, AC34) – AC4 is a list of useful resources; AC11 is a list of emerging research questions but no findings; AC34 provides a plan for phase 2 of the DTC but no findings; AC36 sets out the objectives of the DTC but does not provide any empirical findings; W35 sets out list of potential DWPA measures only; W19 did not report on specific DTC findings.

There were 18 replicates in total (15 were SPICE compliant and 3 were SPICE relevant) but they were also discarded as the same information could be found in other, more comprehensive, documents that were reviewed. It is important to discard replicates to avoid double counting.

General Comments on methodology

This analysis is focussed on whether an explicit statement could be found in the documents reviewed. Lack of a reference does not automatically mean that good practice was not followed, it means that it was not explicitly stated in the document.

It is very possible that there is evidence available in the documents reviewed that a reviewer did not recognise. Given the size and heterogeneity of the data set, it is quite possible we have missed something or misunderstood what was presented. Having said that, our protocol requires things to be explicitly presented, such that noting different aspects in different sections within a document, or different documents, without a statement linking them together, would not be considered 'evidence' in our sense.

This strict approach is necessary to enable a rigorous review. It is also important to recognise that non-academic readers are unlikely to have the time to piece together evidence; so the main policy and practice relevant messages need to be clearly stated with an evidence trail to allow an assessment of robustness in order to maximise the utility and impact of the research. For example, very few of the earlier reports had an executive or policy summary. Nine SPICE (A35, A9, A11, AC15, AC25, AV11, E9, O13, and O11) documents had recommendations and 12 other relevant documents contained recommendations; but this is only around half of the documents.

The extraction table was problematic as the questions have different scales for analysis – CEA needs rows for individual measures in individual catchments, or even farms. But often the engagement or confounding factors is reported across all catchments, or generically within the catchments. This makes using the same spreadsheet for very different types of scientific questions, reporting at very different levels of granularity, quite challenging.

The need to be flexible and adaptive was highlighted in A44, whilst the need to be ‘opportunistic’ was noted in A45 and reinforced in AC5. AC5 therefore observed that the adaptive approach meant limited experimental rigour was possible regarding understanding uptake of specific interventions. AC6b noted that empirical social science research has been modest in scope during the first phase, and it is possible to conclude the more resources to collect, analyse and report on the realities of how farmers were willing and able to uptake measures beyond the end of phase two would have been useful.

Appendix F: Physical science evidence tables

This appendix contains evidence tables for the primary (Table F1) and secondary (Tables F2 to F6) physical science questions.

Table F 1: Primary Physical: How effective were DTC agri-environment interventions in DTC catchments for improving and maintaining water quality?

DTC Intervention	Water quality parameter	Effect of intervention?	Modelling or monitoring study?	Coverage of catchment	Pathway assessed for change	Robustness of results (mean score: Re1 to Re6)	Robustness of reporting (R1)	Document ID
Farmyard management (Biobed and drainage field inline)	Pesticides	Reduced total pesticide concentration by >90% at the plot scale	Monitoring	100%	Surface and soil water	S	S	W4
Crop type (winter oilseed raddish)	Nitrate, TP	Nitrate >75% reduction at field scale, N_2O no impact	Monitoring	100%	Soil water	G	E	A15, A27
		Nitrate no change at catchment scale		20%	River			
		TP no impact at field scale		100%	Soil water			
		TP reduced at catchment scale (~33%)		20%	River			
Tillage (direct drilling and shallow non inversion tillage)		Nitrate and N_2O no impact at field scale		100%	Soil water			
		Nitrate no impact at catchment scale		20%	River			
		TP no change at field scale		100%	Soil water			
		TP reduced at the catchment scale (~33%)		20%	River			
Livestock-Animals (stopping riparian poaching)	Suspended sediment	Reduced sediment load by 2.3% at subcatchment scale	Monitoring	98m of river fenced	River water	G	E	A63
Field drainage (track runoff interception_135m2)	TN, TP and sediment	Annual amounts retained at the intervention scale are reported. Absolute reductions in specific yields (kg/ha/yr) of TP (0.06), TN (0.16) and sediment (42) given for the catchment outlet (1.6	Monitoring	~25%	Surface runoff	S	S	A5
Field drainage (modified ditch system)								
Field drainage (track								

runoff interception_55m2)		km²).						
Field drainage (attenuating ditch and overland flow)								
Field drainage (track and runoff interception_80m2)								
Combination (tracks, ponds, fences)	TN, TP, Suspended sediment	Mean sediment concentration increased by 8% (50-54mg/l), total nitrate decreased by 26% (6.1 – 4.5mg/l) and total P decreased by 50% (0.77 – 0.39mg/l). At subcatchment scale.	Monitoring	Not extracted	River	S	S	AC21
Combination (Tracks, ponds, fences and roofing)		Mean sediment concentration increased by 42% (50-71mg/l), total nitrate decreased by 5% (6.1 – 5.8mg/l) and total P decreased by 34% (0.77 – 0.51mg/l). At subcatchment scale.		Not extracted	River			
Wetlands and ponds (roadside wetlands)	Turbidity, sediment concentration, sediment load, total P and nitrate	Decreases reported in turbidity (14%), suspended sediment concentration (14%), sediment load (82%). No change reported in total P. Increase in nitrate (15%) At subcatchment scale	Monitoring	2.7%	River	G	E	A11
Wetlands and ponds (instream pond and wetland)	TP	Decrease in concentration	Monitoring	Not extracted	River	G	Ex	A6
Fertiliser application	NO3, TP, SS, NH3	Percentage decreases reported: 22 (NO3), 47 (TP), 66 (SS), 30 (NH3)	Modelling	Not extracted	Loading to river	S	E	A69

Fertiliser application	SRP and TP	Percentage decreases reported: 37 (SRP), 40 (TP).	Modelling	100%	River	S	E	A67
Buffer strips and field corner management (2m, 6m)	NO3 and TP	Percent decreases reported: 2m: 2.3 (N), 12.2 (P); 6m: 4.6 (N), 16.9 (P)	Modelling	Not extracted	River	G	E	A62
Cultivation (conservation tillage)	NO3 and TP	Percent increases reported: 4.7 (N), 3.8 (P);			River			
Cultivation (no tillage)		Percent increases reported: 6.3 (N), 7.2 (P)			River			
Field drainage (no tile drain)	NO3 and TP	Percent decreases reported: 58.9 (N), 31.6 (P)			River			
Cover cropping (red clover)	NO3 and TP	Percent decreases reported: 19.6 (N), 1.6 (P)			River			
Combination of above measures	NO3 and TP	Percent decreases reported: 24.1 (N), 17.9 (P)			River			
Livestock grazing/stocking (reduced stock density and avoiding poaching)	SS and TP	Percent decreases reported: 6.5 (SS), 4.7 (P)	Modelling	100%	River	G	E	A2
Buffer strips and field corner management	SS and TP	Percent decreases reported: 3.0 (SS), 2.0 (P)		10%	River			
Other: Additional water storage	SS and TP	Percent decreases reported: 4.5 (SS), 4.0 (P)		5%	River			

Note: S=Satisfactory, G=Good, E=Excellent

Table F 2: SP1: Have DTC monitoring methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment

DTC Intervention	Scale	1. Are hydrological observations considered as part of conclusions (M1)?	2. Automated monitoring and ground truthing (M2)?	3. Field and/or laboratory analytical methods (M3)?	4. Temporal resolution of monitoring (M4)?	5. Storm event sampling (M5)?	6. Are baseline conditions adequately characterised (M6)?	7. Are replicates used (M7)?	8. Is length of post intervention period sufficient (M8)?	9. Relative location of intervention and monitoring (M9)?	10. Experimental design (M10)?	Average score for monitoring method (M1 to M10)	Robustness of reporting (R1)	REFERENCE ID
Farmyard Management	Plot	No	S	G	S	No	Yes	No	G	E	E	G	S	W4
Cover crops	Plot	No	S	E	S	No	No	Yes	G	E(E)	G (G)	G (G)	E(E)	A15 (A27)
	Field	No (No)		E(E)	G(G)	No(No)	No (No)	Yes (Yes)	G (G)					
	Sub-catchment (7.14km2)	Yes & not co-located		S	E	Manual	No	No	G					
Tillage	Plot	No		E	S	No	Yes	Yes	S					
	Field	No (No)		E(E)	G(G)	No (No)	Yes (No)	Yes (Yes)	S (G)					
	Subcatchment (7.14km2)	Yes & not co-located		S	E	Manual	Yes	No	S					
Livestock-animals	Subcatchment (0.7km2)	Yes & not co-located	E	E	E	Auto	No	No	S	E	G	G	E	A63
Field drainage and	Intervention, inference	No	S	S	S	No	No	No	S	S	G	S	S	A5

farmyard management	subcatchment (1.6km2)													
Field drainage and farmyard management	Subcatchment (4.6km2)	Yes & co-located	S	S	S	No	Yes	No	E	S	G	G	S	A21
Wetlands and ponds	Subcatchment (19.7km2)	Yes & co-located	G	E	E	No	Yes	No	E	G	G	G	E	A11
Wetlands and ponds	Subcatchment (4.6 and 8.5 km2)	Yes & co-located	S	E	S	No	No	No	S	E	E	G	E	A6

Note: S=Satisfactory, G=Good, E=Excellent

Table F 3: SP2: Have DTC modelling methodologies resulted in robust evidence that enables the effectiveness of a variety of agri-environment interventions (in mitigating rural diffuse pollution) to be assessed at a range of scales from plot to catchment?

Intervention	Scale	Has uncertainty Been reported (Mo1)?	Have authors undertaken calibration and/or validation (Mo2)	Is there peer reviewed evidence for the validity of the model Mo3)?	Model time step sufficient to represent key processes (Mo4)?	Duration of model testing period (Mo5)?	Duration of scenario simulation (Mo6)?	Average score for modelling method (M1 to M6)	Robustness of reporting (R1)	REFERENCE ID
Fertiliser application	Catchment (1700km2)	No	No validation	Yes	S	S	S	S	E	A69
Buffer strips and field corner management In field manure slurry management Field drainage Cover cropping Combination	Sub catchment (19.6km2)	Yes	Yes both	Yes	G	S	E	G	E	A62
Livestock grazing/stocking Buffer strips and field corner management Additional water storage	Sub-catchment (2km2)	No	No validation	Yes	E	S	S	G	E	A2
Fertiliser application	Catchment (>37km2)	No	Yes both	Yes	G	G	E	G	E	A67

Note: S=Satisfactory, G=Good, E=Excellent

Table F 4: SP3: Based on DTC evidence how effective are specific agri-environment interventions and combinations of measures in mitigating diffuse pollution and improving or maintaining ecology? Are these relevant to or assessed against regulatory standards? Does effectiveness change over time?

DTC Intervention	Water quality parameter	Magnitude of effect of intervention?	Effect and its significance (Re4)	Relevant or related to standards (Re5)?	Is effect reported to vary over time (Re6)?	Robustness of results (mean score: Re1 to Re6)	Robustness of reporting (R1)	Document ID
Farmyard management (Biobed and drainage field inline)	Pesticides	Reduced total pesticide concentration by >90% at the plot scale	Not specified	No	No but results for 2 yrs post intervention	S	S	W4
Crop type (winter oilseed raddish)	Nitrate, P	Nitrate >75% reduction at field scale Nitrate no change at catchment scale P no impact at field scale. P reduction at catchment scale	Reduction: significant No change: significant No change: significant Reduction: significant	Yes for Nitrate at catchment scale No No	No	G E G G	E	A15
Tillage (direct drilling and shallow non inversion tillage)	Nitrate, P	Nitrate no impact at field and catchment scale P no impact at field scale. P reduction at catchment scale	No change: significant No change: significant Reduction: significant	Yes for Nitrate at catchment scale No No	No	G		

Livestock-Animals (stopping riparian poaching)	Suspended sediment	Reduced sediment load by 2.3% at subcatchment scale	Not extracted	No	No	G	E	A63
Field drainage (track runoff interception_135m2)	N,P and sediment	Annual amounts retained at the intervention scale are reported. Absolute reductions in specific yields of P, N and sediment given for the catchment outlet.	Not extracted	No	No	S	S	A5
Field drainage (modified ditch system)								
Field drainage (track runoff interception_55m2)								
Field drainage (attenuating ditch and overland flow)								
Field drainage (track and runoff interception_80m2)								
Combination (tracks, ponds, fences)	N,P,Sediment	Mean sediment concentration increased by 8% (50-54), total nitrate decreased by 26% (6.1 – 4.5) and total P decreased by 50% (0.77 – 0.39). At subcatchment scale.	Not extracted	No	No	S	S	AC21
Comination (Tracks, ponds, fences and roofing)	N, P, Sediment	Mean sediment concentration increased by 42% (50-71), total nitrate decreased by 5% (6.1 – 5.8) and total P decreased by 34% (0.77 – 0.51). At subcatchment scale.	Not extracted	No	No but two year average given.			
Wetlands and ponds (roadside wetlands)	Turbidity, sediment concentration,	Decreases reported in turbidity (14%), suspended sediment concentration (14%), sediment load	Significant changes apart from TP	Yes for suspended sediment	No	G	E	A11

	sediment load, total P and nitrate	(82%). No change reported in total P. Increase in nitrate (15%) At subcatchment scale						
Wetlands and ponds (instream pond and wetland)	TP	Decrease in concentration	Decrease, significant	No	No	G	E	A6
Fertiliser application	NO3, P, SS, NH3	Percentage decreases reported: 22 (NO3), 53 (P), 66 (SS), 34 (NH3)	Decrease, not specified	No	No	S	E	A69
Fertiliser application	SRP and TP	Percentage decreases reported: 37 (SRP), 40 (TP),	Decrease, not specified	No for TP. Yes for SRP	No	S	E	A67
Buffer strips and field corner management (2m, 6m)	NO3 and TP	Percent decreases reported: 2m: 2.3 (N), 12.2 (P); 6m: 4.6 (N), 16.9 (P)	Decrease, not specified	No	No	G	E	A62
Cultivation (conservation tillage)	NO3 and TP	Percent increases reported 4.7 (N), 3.8 (P);	Increase, not specified	No	No			
Cultivation (no tillage)	NO3 and TP	Percent increases reported: 6.3 (N), 7.2 (P)	Increase, not specified	No	No			
Field drainage (no tile drain)	NO3 and TP	Percent decreases reported: 58.9 (N), 31.6 (P)	Decrease, not specified	No	No			
Cover cropping (red clover)	NO3 and TP	Percent decreases reported: 19.6 (N), 1.6 (P)	Decrease, not specified	No	No			
combination	NO3 and TP	Percent decreases reported: 24.1 (N), 17.9 (P)	Decrease, not specified	No	No			
Livestock grazing/stocking (reduced stock density and avoiding poaching)	SS and TP	Percent decreases reported: 6.5 (SS), 4.7 (P)	Decrease, not specified	No	No	G	E	A2
Buffer strips and field corner management	SS and TP	Percent decreases reported: 3.0 (SS), 2.0 (P)	Decrease, not specified (1)	No				
Other: Additional water storage	SS and TP	Percent decreases reported: 4.5 (SS), 4.0 (P)	Decrease, not specified	No				

			(1)					

Note: S=Satisfactory, G=Good, E=Excellent

Table F 5: SP6: What evidence is there from the DTC programme that the effectiveness of agri-environment interventions varied between DTC catchments and was this related to differences in the design and/or management of the interventions?

DTC Intervention	Water quality parameter	Magnitude of effect of intervention?	Effect and its significance	DTC catchment	Intervention design detail	Intervention Management detail	Document ID
Farmyard management (Biobed and drainage field in line)	Pesticides	Reduced total pesticide concentration by >90% at the plot/field scale	Not specified	Wensum	49m2 biobed and 200m2 drainage field	Not specified	W4
Crop type (winter oilseed raddish)	Nitrate, P	Nitrate >75% reduction at field scale Nitrate no change at catchment scale P no impact at field scale. P reduction at catchment scale	Reduction: significant No change: significant No change: significant Reduction: significant	Wensum	143ha trial in catchment of 714ha.	Herbicide and Molluscicide added	A15
Tillage (direct drilling and shallow non inversion tillage)	Nitrate, P	Nitrate no impact at field and catchment scale P no impact at field scale. P reduction at catchment scale	No change: significant No change: significant Reduction: significant		Not Extracted	Not Extracted	
Livestock-Animals (stopping riparian poaching)	Suspended sediment	Reduced sediment load by 2.3% at subcatchment scale	Not extracted	Eden	Impact of a 98m unfenced area assessed	Not managed	A63
Field drainage (track runoff interception_135m2)	N,P and sediment	Annual amounts retained at the intervention scale are reported. Absolute reductions in specific yields of P, N and sediment	Not extracted	Eden	135m2	Not extracted	A5

Field drainage (modified ditch system_20m2)	N,P and sediment	given for the catchment outlet.			20m2	Not extracted	
Farmyard management (track runoff interception_55m2)	N,P and sediment				55m2	Not extracted	
Field drainage (attenuating ditch and overland flow_150m2)	N,P and sediment				150m2	Not extracted	
Field drainage (track and runoff interception_80m2)	N,P and sediment				80m2	Not extracted	
Combination(tracks, ponds, fences)	TN, TP, Suspended sediment	Mean sediment concentration increased by 8% (50-54 mg/l), total nitrate decreased by 26% (6.1 – 4.5 mg/l) and total P decreased by 50% (0.77 – 0.39 mg/l). At subcatchment scale.	Not extracted	Avon	Not extracted	Not extracted	AC21
Combination (Tracks, ponds, fences and roofing)		Mean sediment concentration increased by 42% (50-71 mg/l), total nitrate decreased by 5% (6.1 – 5.8 mg/l) and total P decreased by 34% (0.77 – 0.51 mg/l). At subcatchment scale.					
Wetlands and ponds (roadside wetlands)	Turbidity, sediment concentration, sediment load, total P and nitrate	Decreases reported in turbidity (14%), suspended sediment concentration (14%), sediment load (82%). No change reported in total P. Increase in nitrate (15%) At subcatchment scale	Reductions in turbidity and suspended sediment, significant. Increase in nitrate,	Wensum	Not extracted	Not managed	A11

			significant. No significant change reported for TP.				
Wetlands and ponds (instream pond and wetland)	Nutrients	Decrease in concentration	Decrease, significant	Avon	Not extracted	Not managed	A6
Fertiliser application	NO3, P, SS, NH3	Percentage decreases reported: 22 (NO3), 53 (P), 66 (SS), 34 (NH3)	Decrease, not specified	Avon	No	No	A69
Fertiliser application	SRP and TP	Percentage decreases reported: 37 (SRP), 40 (TP),	Decrease, significance tested	Avon	No	30% reduction in fertiliser applications	A67
Buffer strips and field corner management (2m, 6m)	NO3 and TP	Percent decreases reported: 2m: 2.3 (N), 12.2 (P); 6m: 4.6 (N), 16.9 (P)	Decrease, significance tested	Wensum	Width specified	Not managed	A62
Cultivation (conservation tillage)	NO3 and TP	Percent increases reported: 4.7 (N), 3.8 (P);	Increase, significance tested		No	Not managed	
Cultivation (no tillage)	NO3 and TP	Percent increases reported: 6.3 (N), 7.2 (P)	Increase, significance tested		No	Not managed	
Field drainage (no tile drain)	NO3 and TP	Percent decreases reported: 58.9 (N), 31.6 (P)	Decrease, significance tested		No	Not managed	
Cover cropping (red clover)	NO3 and TP	Percent decreases reported: 19.6 (N), 1.6 (P)	Decrease, significance tested		No	Not managed	
Combination of above measures	NO3 and TP	Percent decreases reported: 24.1 (N), 17.9 (P)	Decrease, significance tested		No	Not managed	
Livestock	SS and TP	Percent decreases reported: 6.5	Decrease, not	Eden	Areal extent reported	No	A2

grazing/stocking (reduced stock density and avoiding poaching)		(SS), 4.7 (P)	specified				
Buffer strips and field corner management	SS and TP	Percent decreases reported: 3.0 (SS), 2.0 (P)					
Other: Additional water storage	SS and TP	Percent decreases reported: 4.5 (SS), 4.0 (P)					

Table F 6: **SP7: What evidence is there from DTC data that confounding factors (e.g. climate, non agricultural pollution) may be important in the interpretation of the results?**

DTC Intervention	Water quality parameter	Magnitude of effect of intervention?	Account of confounding factors (Re3)?	What confounding factors were considered?	Were confounding factors quantified?	How did confounding factors affect the results?	Document ID
Farmyard management (Biobed and drainage field)	Pesticides	Reduced total pesticide concentration by >90% at the plot/field scale	S	No specified	No	N/A	W4
Crop type (winter oilseed raddish)	Nitrate, P	Nitrate >75% reduction at field scale Nitrate no change at catchment scale P no impact at field scale. P reduction at catchment scale	G	Legacy pollution from past fertiliser additions. Rainfall	No but conceptualised	Legacy stores of nutrients in soils and sediment will buffer against changes Peaks in TP related to rainfall	A15
Tillage (direct drilling and shallow non inversion tillage)	Nitrate, P	Nitrate no impact at field and catchment scale P no impact at field scale. P reduction at	G	Legacy pollution from past fertiliser additions	No but conceptualised	Legacy stores of nutrients in soils and sediment will buffer against changes	A15

		catchment scale		Rainfall		Peaks in TP related to rainfall.	
Livestock-Animals (stopping riparian poaching)	Suspended sediment	Reduced sediment load by 2.3% at subcatchment scale	E	Rainfall	Yes	The importance of river flow affecting sediment concentration was taken into account.	A63
Field drainage (track runoff interception_135m2)	N,P and sediment	Annual amounts retained at the intervention scale are reported. Absolute reductions in specific yields of P, N and sediment given for the catchment outlet.	S	Not specified	No	N/A	A5
Field drainage (modified ditch system_20m2)	N,P and sediment						
Farmyard management (track runoff interception_55m2)	N,P and sediment						
Field drainage (attenuating ditch and overland flow_150m2)	N,P and sediment						
Field drainage (track and runoff interception_80m2)	N,P and sediment						
Field drainage (tracks, ponds, fences)	N,P,Sediment	Mean sediment concentration increased by 8% (50-54), total nitrate decreased	Sediment (G) Nutrients (S)	For sediment_Prior agri-environment schemes	Conceptualised	Concentrations have increased over time as a result of works that have	AC21

		by 26% (6.1 – 4.5) and total P decreased by 50% (0.77 – 0.39). At subcatchment scale.				taken place.	
Farmyard management (Tracks, ponds, fences and roofing)	N, P, Sediment	Mean sediment concentration increased by 42% (50-71), total nitrate decreased by 5% (6.1 – 5.8) and total P decreased by 34% (0.77 – 0.51). At subcatchment scale.					
Wetlands and ponds (roadside wetlands)	Turbidity, sediment concentration, sediment load, total P and nitrate	Decreases reported in turbidity (14%), suspended sediment concentration (14%), sediment load (82%). No change reported in total P. Increase in nitrate (15%) At subcatchment scale	E	Climate	Yes	Low flows of spring/summer 2017 taken into account.	A11
Wetlands and ponds	TP	Decrease in	S	Not specified	No	N/A	A6

(instream pond and wetland)		concentration					
Fertiliser application	NO3, P, SS, NH3	Percentage decreases reported: 22 (NO3), 53 (P), 66 (SS), 34 (NH3)	S	None	Not specified	Not specified	A69
Fertiliser application	SRP and TP	Percentage decreases reported: 37 (SRP), 40 (TP),	G	Not specified	Not specified	Not specified	A67
Buffer strips and field corner management (2m, 6m)	NO3 and TP	Percent decreases reported: 2m: 2.3 (N), 12.2 (P); 6m: 4.6 (N), 16.9 (P)	G	Multiple	Not specified	rating curve uncertainty under high-flow conditions, difficulties in modelling responses to extreme conditions, difficulties in modelling antecedent conditions, incorrect timing of management practices	A62
In field manure slurry management (conservation tillage, no tillage)	NO3 and TP	Percent increases reported: cons: 4.7 (N), 3.8 (P); no: 6.3 (N), 7.2 (P)	S	Multiple	Not specified	As above	

Field drainage (no tile drain)	NO3 and TP	Percent decreases reported: 58.9 (N), 31.6 (P)	S	Multiple	Not specified	As above	
Cover cropping (red clover)	NO3 and TP	Percent decreases reported: 19.6 (N), 1.6 (P)	S	Multiple	Not specified	As above	
combination	NO3 and TP	Percent decreases reported: 24.1 (N), 17.9 (P)	S	Multiple	Not specified	As above	
Livestock grazing/stocking (reduced stock density and avoiding poaching)	SS and TP	Percent decreases reported: 6.5 (SS), 4.7 (P)	G	Rainfall	Quantified or stated none occur	To investigate the temporal scaling of the results, three large events during a wet 5-month period in 2012 were analysed.	A2
Buffer strips and field corner management	SS and TP	Percent decreases reported: 3.0 (SS), 2.0 (P)	E	Rainfall	Quantified or stated none occur		A2
Other: Additional water storage	SS and TP	Percent decreases reported: 4.5 (SS), 4.0 (P)	E	Rainfall	Quantified or stated none occur		A2

Note: S=Satisfactory, G=Good, E=Excellent

Appendix G: Social Science tables of robustness and evidence

The following tables provide more detail to explain the results summarised in section 5 (Rapid Evidence Assessment: Farmer engagement and implementation)

Table G 1: Robustness Results for Primary SS Question "How effective were DTC engagement processes in fostering and retaining uptake of DTC agri-environmental interventions for improving and maintaining water quality?"

Which interventions taken up?	Number of sources linking engagement to uptake	Robustness
Buffer Strips and field corner management	O11	Satisfactory (S)
Farmyard management	AC26, O11	Not Recorded (NR), S
Traffic on fields	O11	S
Other (nutrient management advice)	AC26	NR
Other (mitigation measures for tackling Phosphorous losses)	AC26	NR
Livestock- grazing and stocking	O11	S
Other (reversion maize to grass)	O11	S
Other (nutrient management advice)	O11	S
Wetlands and ponds	O11	S
Other (integrated manure and fertiliser advice)	O11	S
Cultivation	O11	S
Cover Cropping	O11	S
Pesticide Use	O11	S

Other (woodland planting)	O11	S

Table G 2: Robustness Results for SS1 "What evidence is there that the DTC engagement methodologies appropriately informed, consulted and actively involved farmers and other stakeholders to maximise uptake of interventions?"

Types of stakeholders Engaged	No of docs	Robustness (Satisfactory, Good, Excellent or no record)
Farmers	12	2 Satisfactory (O13, O11), 4 Good (A35, A9, AC15, AC21) & 6 Not Recorded
Advisors	6	2 Satisfactory (O11, O13), 3 Good (AC15, AC21, A9) and 1 Not Recorded
CSFOs	2	1 Good (AC21)
EA	2	1 Satisfactory (O13), 1 Good (AC21)
NGOs	4	1 Satisfactory (O13), 2 Good (AC15, A9)
Defra	2	1 Satisfactory (O13), 1 Not Recorded
Contractors	1	1 Satisfactory (O13)
Utilities	2	1 Satisfactory (O13), 1 Good (AC15)
Reach of engagement	5	1 Satisfactory (O13), 1 Good (AC21) and 3 Not Recorded
Depth of engagement Informed	6	2 Satisfactory (O11, O13), 2 Good (A35, AC21) and 2 Not Recorded
Consulted	11	2 Satisfactory (O11, O13), 4 Good (A35, A9, AC21, AC15) and 5 Not Recorded
Actively Involved	5	1 Satisfactory (O13), 2 Good (A35, AC21) and 2 Not Recorded

Table G 3: Robustness Scores for SS2 "What evidence does the DTC data provide about non-participants and why they did not engage in the process?"

Type of Non-Engagers reported	**Reasons for Non-Engagement reported**	**Effect on uptake of intervention reported**	**Evidence robustness (no of sources**
Farmer (4 documents)	Tenancy and succession issues (1 document) No details (3 documents)	No information provided	2 document Good (A9, AC21) 1 documents Satisfactory (O11) 1 document Not Reviewed

Table G 4: Robustness for SS3 "What evidence is there that the DTC considered behavioural factors when engaging farmers in implementing interventions?"

Types of factors reported	Effect of factors on uptake of interventions reported	Evidence robustness (no of sources)
Level of awareness of diffuse pollution	2 documents report that the level of awareness of diffuse pollution from agriculture was recorded	2 Good (AC15, A35)
Understanding of diffuse pollution	The understanding of diffuse pollution did influence uptake of interventions, for a wide range of interventions	2 Good (AC15, A35) 1 Not Recorded
Farmers' perception of efficacy of intervention	Farmers report mixed perceptions for most interventions	3 Good (A35, AC15, AC21) 1 Satisfactory (O11) 4 Not Recorded
Farmers' attitudes to the environment	2 documents report that attitudes to the environment have been recorded	1 Good (A35) 1 Not Recorded

Farmers' agri-environment experience	3 documents report that farmers' agri-environment experience was recorded.	2 Good (A9, AC15) 1 Not Recorded
Farmers' attitudes to change and innovation	6 documents report that farmers' attitudes to change and innovation were recorded	2 Good (A35, AC21) 4 Not Recorded

Table G 5: Interventions across Catchments

The social science documents discussed these measures the various catchments. Note this may differ from Physical science due to different SPICE documents being used. The plain text in table G5 refers to specific DTC measures implemented as part of the DTC programme and the italics refer to measures already in place or potential measures that farmers might implement them in the future.

Catchment	Intervention Type Discussed	Document
Avon	Buffer strips and field corner management	AC21, O11
	Farmyard management	AC21, AC26, O11
	Traffic on fields	AC21, O11
	Wetlands and ponds	AC21, O13, O11
	Other (mitigation measures for tackling Phosphorous losses)	AC26
	Livestock- grazing and stocking	O11
	Other (reversion maize to grass)	O11
	Other (nutrient management advice)	O11
	Cover Cropping	*O13*
	Field drainage	*O13*
	In-field manure slurry management	*O13*

	Livestock-feeding and diets	*O13*
Total: 8 DTC measures and 4 other measures reported in 4 documents		
Eden	Buffer strips and field corner management	AC21
	Farmyard management	AC21, O11
	Other (aeration of grass to improve drainage)	AC21, O11
	Other (Instream ditch barriers)	AC21, O11
	Wetlands and ponds	AC21, O11
	Traffic on fields	AC21
	Other (integrated manure and fertiliser advice)	AC21, O11
	Cultivation	O11
	In-field Slurry Management	E9
	Livestock- grazing and stocking	O11
	Other (woodland creation)	O11, AC21
Total: 10 DTC intervention types reported in 3 documents		
Tamar	Buffer strips and field corner management	AC21
	Farmyard management	O11
	Traffic on fields	O11
	Livestock- grazing and stocking	O11
Total: Four intervention types reported in 2 documents		
Wensum	Buffer strips and field corner management	AC21
	Cover Cropping	AC21, O11
	Cultivation	AC21, O11

	Pesticide Use	O11
	Other (woodland planting)	O11, AC21
	Fertiliser Application	W21
	Plants	W21
	Traffic on Fields	W23
Total: 5 DTC + 3 other intervention types reported in 4 documents		
Avon & Tamar	*Cultivation*	*AV11*
Avon, Eden and Wensum	*Cover Cropping*	*A35, A9, AC15*
	Cultivation	*A35, A9, AC15*
	Livestock -grazing and stocking	*A35, A9, AC15*
	Buffer strips and field corner management	*A9*
	Farmyard management	*A9, AC15*
	Fertiliser Application	*A9, AC15*
	Field Drainage	*A9, AC15*
	In-field Slurry Management	*A9, AC15*
	Livestock -Animals	*A9*
	Livestock – feeding and diets	*A9, AC15*
	Plants	*A9*
	Traffic on Fields	*A9*
	Cropping	*AC15*
	Other (substitute metaldehyde with ferric phosphate)	*AC15*
Total = 12 proposed or non-DTC interventions types in 5 documents		

Table G 6: Robustness Results for SS4 "What evidence is there of engagement and uptake of DTC interventions varying between catchments?"

	Number of documents with comparative data	Evidence robustness
Engagement	8	2 Satisfactory (O11, O13), 3 Good (A9, AC15, AC21) & 3 Not Recorded
Behaviour	4	1 Satisfactory (O11), 2 Good (AC15, AC21) & 1 Not Recorded
Intervention	3	1 Satisfactory (O11), 1 Good (AC21) & 1 Not Recorded

Table G 7: List of types of intervention assessed for cost-effectiveness in the DTC projects

Type of Intervention	**Number of Cost-effectiveness assessments found in SPICE documents**
Buffer strips and field corner management	4
Combination	14
Cover cropping	7
Cropping	1
Cultivation	13
Farmyard management	19
Fertiliser application	7
Field drainage	3
In-field manure slurry management	4
Livestock -animals	1
Livestock- grazing and stocking	11
Livestock-feeding and diets	4
Other	14
Pesticide use	3
Plants	2
Traffic on fields	3

Wetlands and ponds	9
Grand Total	119

Table G 8: Robustness Results for SS5 "What evidence is presented on the cost effectiveness and benefits of DTC interventions, during the initiative and for the five year period beyond the end of the initiative?"

Number of Sources	Range of costs and benefits considered	Evidence robustness (no of sources E, G, S, NR)
6 documents	Direct costs to farmers	1 Excellent(A11), 1 Good (A9), 2 Satisfactory (W21, O11) and 2 Not Recorded
1 document	Costs to other stakeholders	1 Good (AC21)
4 documents	Private benefits to farmers	3 Good (AC14, AC21, A35), 1 Not Recorded
5 documents	Benefits to society	1 Excellent (A11), 2 Good (A35, AC21), 1 Satisfactory (W21) and 1 Not Recorded

Table G 9: Robustness Results for SS6 "What evidence is provided that confounding factors (e.g. existing non DTC activities) were accounted for when reporting engagement and uptake of DTC interventions?"

	Confounding factors noted (no of sources)	Evidence robustness
Engagement	2	1 Good (AC15), 1 Not Recorded
Behaviour	7	2 Satisfactory (O11, O13), 4 Good (A35, AC15, A9, AC21), 1 Not Recorded
Uptake	6	1 Satisfactory (O11), 3 Good (A35, AC15, AC21), 2 Not Recorded
Additionality	1	1 Good (A9)

Appendix H: Report on Demonstration Test Catchment (DTC) programme science evaluation dissemination event

29 June 2020 13:00 – 14:30

Location

The dissemination event was held online using Webex. It was originally planned to occur face-to-face, but was changed to online due to COVID-19 restrictions.

Attendees

Invitations to the event were sent out to government agencies and people who worked on the DTC programme. Eighty people registered to attend the event and approximately 50 attended. Owing to technical difficulties a proportion of those registered could not join. To include those persons a recording of the event along with a transcript of the written chat was shared by email.

Agenda

1. Introduction to the purpose of the project: (J.Phoenix, Defra).
2. Rapid Evidence Assessment Methodology (G.Old, UKCEH).
3. Physical science findings (including questions of clarification) (G.Old, UKCEH).
4. Social science findings (including questions of clarification) (K.Blackstock, JHI).
5. Additional resources that could be utilised to extend knowledge (participants).
6. Closing comments and recommendations for future (from participants).
7. Closing remarks and thanks (J.Phoenix, Defra).

Introduction to the purpose of the project: Defra

Jess Phoenix began by welcoming participants, introducing the DTC programme and stating the objectives of the evaluation project. Gareth Old reminded participants of the five discrete outputs from the projects: 1) full evidence review (report) 2) evidence compendium (available on Defra ScienceSearch: WT15115), 3) physical evidence table (excel spreadsheet), 4) social evidence tables (excel spreadsheet), and (5) inventory of datasets (excel spreadsheet).

Rapid Evidence Assessment (REA) Methodology (Compendium slides 4 -7)

Gareth Old introduced the REA methodology and the primary aim of the project which was to evaluate the knowledge gained on intervention effectiveness for water quality and farmer engagement. The REA was focused on one physical and one social primary question which were supported with seven secondary physical and six secondary social questions. Secondary project aims were included to enable an evaluation of the broader physical and social science methodologies that were adopted and the approach for generalising the findings. It was acknowledged that the evaluation is focused on specific questions and is not therefore a full evaluation of the outcomes of the DTC programme.

The primary and secondary physical and social science questions were introduced.

The methodologies used to screen the physical and social evidence sources were described. For the physical science inclusion and exclusion criteria were defined using the PICO (Population, Intervention, Comparison, Outcome) framework whereas for the social science the SPICE (Setting, Perspective, Intervention, Comparison and Evaluation) framework was used. After screening the 173 eligible evidence sources provided 12 PICO and 15 SPICE compliant documents remained. However, all evidence sources were included in the evidence extraction tables.

Physical science findings (compendium slides 8-18)

For the primary and each of the secondary questions the following information was summarised: 1) available evidence sources, 2) key findings from the available evidence including a consideration of robustness of the findings where appropriate, 3) views from principal investigators, and 4) recommendations from the evidence assessment.

Findings of the broader evaluation were then presented. This included an evaluation of the monitoring and modelling methodologies and the approaches to generalising the science.

Social science findings (compendium slides 19-28)

For the primary and each of the secondary questions the following information was summarised: 1) available evidence sources, 2) key findings from the available evidence, 3) views from principal investigators, and 4) recommendations from the evidence assessment.

Further insights from Non-SPICE compliant documents on engagement, behaviour and uptake were also presented.

Summary of discussion points

This section provides a summary of discussion points, both questions and answers, raised in the webinar through oral questions and written chat.

Please further explain the REA methodology

Rapid evidence assessments are useful to get specific answers to specific questions. This approach was chosen due to the specific questions wanted to be answered by policy colleagues for future policy development, such as Environmental Land Management.

Strict inclusion/exclusion criteria were agreed and a trial was undertaken by the social and physical science review team to ensure consistency in the screening. All evidence provided by Defra was considered. It is acknowledged that some documents, or details within them, may have been missed owing to the rapid review of such a large volume of evidence. To mitigate against this PI's were consulted to cross-check the evidence that had been included.

What is the applicability of some of the improvements in water quality?

The applicability of the reported changes in water quality concentrations (i.e. chemicals and sediments) to specific flow pathways (e.g. soil water, overland flow, ditch, stream) was questioned. This was an important question that reflects the high level questions and answers that are being presented here in this project. Some reductions related to small scale in-field changes whilst other reductions related to changes at the larger scale in stream/river water. Participants were encouraged to look at the evidence tables included in the full report to gain a detailed understanding of the results.

Could the financial cost and time of collecting long term robust data be disproportionate to the benefits gained?

Sustainable cost effective environmental management requires a robust understanding of the functioning of catchments and interventions. Important considerations include:

- Targeting monitoring to collect the right data at optimal locations;
- Adopting new technologies to optimise data collection;
- Collaboration of interested parties to share data and enable joint monitoring;
- Maximising the transferability of results; and
- Acknowledging the multiple benefits of interventions and including these in monitoring plans.

The longer term costs of not monitoring should also be taken into account and these may include:

- Potential for environmental damage costs of poor water quality resulting from ineffective mitigation;
- Investments in sub-optimal interventions or interventions sited in sub-optimal locations (with associated maintenance costs); and
- Higher water treatment costs downstream.

Were data on non-agricultural pollutant inputs used in the evaluation?

Understanding the significance of non-agricultural pollutant contributions is important in understanding the effectiveness of interventions. Nitrogen deposition was likely to have been included in catchment matcher tool and these data may have also been used in models. Furthermore, load apportionment modelling has illustrated the importance of point source (sewage) contributions (considered in Section 6 of the report). Although these data and methodologies may complement the work done within the DTC project, the inclusion of these data was beyond the scope of this evaluation project.

How was uncertainty accounted for in the REA?

Undertaking a quantitative assessment of the uncertainty in assessments of intervention effectiveness was beyond the scope of this evaluation project. It is complex to quantitatively compare uncertainties between studies and this is beyond the scope of an REA; although it could be appropriate for a full systematic review (which needs considerably more resources).

To account for uncertainty, we included in our assessments of robustness whether or not uncertainty was quantified/considered either in terms of replication in monitoring or uncertainty analysis in modelling (see Section 3 and Appendix C of report).

In utilising model results in policy the value of considering ensembles of models to account for uncertainty was emphasised.

How are the findings relevant for catchment management?

This project was aimed at assisting policymakers and government advisors in making use of the evidence from the DTC programme. Alongside this project, Defra will publish an evidence compendium (WT15116) which compiled findings from the DTC programme to make the findings more usable and understandable for those who engage in catchment and farm management.

It was emphasised that farm advisers from different organisations have different priorities (e.g. economic vs environmental) and the importance of trusted advisers was noted. Given the heterogeneity of farmers (e.g. their economic position, preference) bespoke advice and targeting of interventions is likely to be important (speculation). Further, information on barriers to entry and the role of advisers in supporting farmer decision making may be found in this compendium (WT15115).

Additional resources that could be utilised to extend knowledge

Participants were invited to highlight additional datasets or studies that may contribute to the body of evidence. The following resources were raised:

1. Attendees suggested that the findings from this project could be transferred to similar catchment types and farm systems through the NERC funded Natural Flood Management programme[11].
2. Complimentary data are likely to be available for parts of the catchments from the catchment sensitive farming initiative.
3. Natural England has synthesised evidence along similar lines and highlighted a number of studies. Although these studies are unlikely to provide additional evidence on the effectiveness of interventions, they are relevant to assessing the effectiveness of

[11] https://nerc.ukri.org/research/funded/programmes/nfm/

agricultural pollution mitigation measures. For example, Catchment Risk Assessment modelling has been undertaken in the Wensum and Avon catchments using SAGIS and FARMSCOPER. Although applications of FARMSCOPER were included in this review (sources W20, W21 and A69), studies applying SAGIS did not feature. Applications of SAGIS may be useful in understanding agricultural pollution in the context of pollution from other sectors. In particular, SAGIS may help quantify confounding factors at larger scales (~>30km^2).

4. Natural England project NECR222 (May et al., 2016) was also mentioned. It developed a risk assessment tool to evaluate the significance of septic tanks around freshwater SSSIs. This methodology may also be useful in accounting for non-agricultural pollutants that may mask signals from agricultural activities.

5. Many studies have also brought together useful knowledge on diffuse pollution in these catchments. For example, Diffuse Water Plans have been developed and bring together catchment characteristics, evidence gaps and actions needed for the Avon, Wensum and Eden catchments[12]. Attention was also drawn to the sediment pathways project on the Somerset Frome where sediment pathways were identified using field surveys[13]. Surveys of this types could be combined with the outputs of the evaluation to identify pollution sources and contribute to the targeting of mitigation measures.

Reference

MAY, L., DUDLEY, B.J., WOODS, H. & MILES, S. 2016. Development of a Risk Assessment Tool to Evaluate the Significance of Septic Tanks Around Freshwater SSSIs. Natural England Commissioned Reports, Number222.

[12] https://www.gov.uk/government/publications/nutrient-management-plan-hampshire-avon
[13] https://bristolavonriverstrust.org/somerset-frome-sediment-pathways-project/