The University Of Sheffield.

Department of Civil and Structural Engineering

Development of Real-Time Surface Water Abstraction Management Tools

By

Alemayehu Shitaye Asfaw

Submitted in part fulfilment of the requirements for the degree of Engineering Doctorate in the Faculty of Engineering, University of Sheffield

October 2018

"The LORD is my strength and my shield; in him my heart trusts, and I am helped; my heart exults, and with my song I give thanks to him."

Psalms 28:7

Abstract

Efficient use of available water resources to meet demand, whilst maintaining the quality of the aquatic environment has become increasingly important. Water quality challenges associated with diffuse agricultural pollutions have also become widely recognized problems globally. This thesis presents the development of new approaches to improve surface water abstraction management with a view to mitigate the challenges associated with increasing pressures on availability of water resources for public water supply and diffuse agricultural pollution. The first part of the thesis presents the development of a real-time surface water abstraction management scheme that integrates a conceptual rainfall-runoff model, a Bayesian inference based uncertainty analysis tool and a water resources management model that incorporates various operating rules to represent real-world operational constraints. The developed approach enables efficient utilization of available water resources and thus provides improved capability to deal with emerging issues of increasing demand, climate adaptation planning and associated policy reforms.

The second part of the thesis describes the development of a new travel time based physically distributed metaldehyde prediction model, which enables water infrastructure operators to consider informed surface water abstraction decisions. Metaldehyde is a soluble synthetic aldehyde pesticide used globally in agriculture and has caused recent concerns due to high observed levels in surface waters utilized for potable water supply. The model provides new approach to represent spatially and temporally disaggregated runoff generation, routing and build-up/wash-off processes using a grid based structure in a GIS environment. Furthermore, a state-of-the-art Monte Carlo based spatial uncertainty analysis tool is employed to assess uncertainties in the metaldehyde prediction model. The structure of the metaldehyde model combined with the availability of high spatiotemporal resolution data has enabled the application of spatial uncertainty analysis of the catchment scale metaldehyde model, which is currently lacking in water quality modelling studies.

Keywords: Surface water abstraction, Diffuse pollution modelling, Metaldehyde, Rainfall-runoff, Hydrological forecasting, Water resources, Uncertainty analysis

Acknowledgements

I feel pleasure in expressing my deep sense of gratitude to my supervisor, Dr. James Shucksmith, for his esteemed guidance, invaluable suggestions and continuous support in the development and delivery of this work. Special thanks to my industrial supervisor, Keiron Maher, for his commitment to this project, unreserved support, positive attitude and encouragement.

I would like to acknowledge EPSRC and Severn Trent Water for funding this project. I'm also thankful to the STREAM IDC program for creating an enjoyable experience during the course of my EngD study. My special thanks go to the Innovation Team at Severn Trent, for enabling such fruitful research collaboration and providing me with essential advice and support during my EngD research. My sincere thanks extend to Bob Stear, Katherine Filby and Julie Hart, who have all been truly supportive of this project.

I wish to avail this opportunity to express my love and appreciation to my wife Tsegaab. It was the power of her prayers and constant support that gave me strength, courage and confidence to pursue my dreams. Above all, I thank God, without whom none of this would have been possible.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AMC	Antecedent Moisture Condition
BADC	British Atmospheric Data Centre
CAPEX	Capital Expenditure
CEH	Centre for Ecology and Hydrology
CV	Coefficient of Variation
DEFRA	Department for Environment, Food and Rural Affairs
DE-MC	Differential Evolution-Markov Chain
DREAM	Differential Evolution Adaptive Metropolis
DWD	Drinking Water Directive
DWI	Drinking Water Inspectorate
EA	Environment Agency
EC	European Commission
EPSRC	Engineering and Physical Sciences Research Council
EU	European Union
FIO	Fecal Indicator Organisms
FOCUS	Forum of the Co-ordination of Pesticide Fate Models and their Use
GIS	Geographical Information System
HSG	Hydrologic Soil Group
IQR	Inter Quartile Range
LandIS	Land Information System
MAC	Maximum Admissible Concentration
MC	Monte Carlo
MCA	Multi Criteria Analysis
MCMC	Markov Chain Monte Carlo
MORECS	Meteorological Office Rainfall and Evaporation Calculation System
NAO	National Audit Office
NSRI	National Soil Resources Institute
NSE	Nash Sutcliffe Efficiency
ODI	Outcome Delivery Incentive
Ofwat	Office for Water

Opex	Operational Expenditure
PCV	Prescribed Concentration or Value
PPDB	Pesticide Properties Database
RTAM	Real-Time Abstraction Management
RWM	Random Walk Metropolis
SCS	Soil Conservation Service
SD	Standard Deviation
STW	Severn Trent Water
TOTEX	Total Expenditure
UKWIR	UK Water Industry Research
USDA	United States Department of Agriculture
WFD	Water Framework Directive
WINEP	Water Industry National Environment Program

Chapter 1

Introduction

This research has been motivated by water quantity and quality challenges associated with operational surface water abstraction management. The focus of this thesis is to develop new approaches to help in informing surface water abstraction management with a view to tackle the challenges associated with increasing pressures on availability of water resources for public water supply and diffuse pollution problems. Thus, the study is compartmentalized into two major parts with distinct objectives. The first part focuses on the development of real-time water resources management model with the aim of improving resilience to existing water resources infrastructure. The second part focuses on understanding the dynamics of diffuse pollutant transport processes in catchments and developing a pollutant prediction model with a view to inform surface water abstraction management. Thus, the following two sections are used to individually discuss the background information and problem statements associated with each part.

1.1 Abstraction management – Water resources

Global availability of freshwater is more than adequate to meet all current and foreseeable water demands. However, its spatial and temporal distributions pose challenges in providing adequate supply to meet demands while minimizing environmental consequences. These challenges are raising concerns, especially with the changing and uncertain future climate, increasing demand from a rapidly growing population and associated impacts on the environment (Cosgrove & Loucks 2015). Surface water and groundwater are the principal sources of drinking water in the UK. The variability in the occurrence of rainfall events at different places and periods raises the need for careful management of water resources to ensure sustainable supply. In England and Wales, two-thirds of drinking water comes from surface water, including reservoirs, lakes and rivers (Gray 2008). Water abstraction for public water supply or other uses are currently putting water resources in parts of England and Wales under pressure. Twenty-five percent of rivers and groundwater in

England and seven percent in Wales are currently unable to provide water for new consumptive abstractions reliably (EA 2011). Environment regulators in the UK are working with the water industry to implement a list of environmental improvement schemes through the Water Industry National Environment Program (WINEP) and the Restoring Sustainable Abstraction (RSA) program, which are both aimed at ensuring water courses in the UK meet European and national targets related to water. Through these programs water companies are required to undertake environmental improvement schemes which will bring about investments to resolve historical abstraction sustainability issues and investigations to prevent future environmental impacts from abstraction. Abstraction licences are likely to be revoked or reduced in cases where abstractions are identified to cause environmental problems, in which case water companies are required to find alternative source of supplies to meet future demands. Cost effective and sustainable solution that enable efficient utilization of existing water resources play an important role in offsetting the loss of available water and minimizing the amount of investment needed to build alternative source of supplies.

Two main factors are known to control surface water abstractions from rivers depending on flow levels in the river. During low flows, abstraction licence conditions are likely to limit the amount of surface water abstraction volumes; whereas hydraulic capacities and/or storage levels in reservoirs determine abstraction levels during high flow periods. Abstraction licences often include constraints with the aim of maintaining a minimum daily flow in the river. Thus, during low flow periods the amount of water available for abstraction on a particular day varies with flow availability in the river. There are flow rate measurements at the abstraction site, which Environment Agency use to control abstraction volumes. However, flow measurement data isn't available to abstraction decision makers on a real-time basis. Hence, abstraction decisions are not currently supported by information on the daily availability of water in the river and abstraction decisions are often made focusing only on avoiding a breach of licence conditions. As a result, significant amount of opportunities to sustainably abstract more water are missed. Enabling to use

opportunities to abstract more water, especially during low reservoir storage periods can have significant impacts in terms of raising storage levels and avoiding the need to trigger drought management actions. Drought permits, which reduce the level of constraints imposed by abstraction licences, are triggered during dry periods when reservoir storages are low. Thus, effective implementation of real-time abstraction management scheme to abstract more water provides a potential to avoid or delay triggering drought permits, hence providing a double benefits of reducing impacts of drought on the water resources production system and the environment.

The use of flow forecasting models for abstraction management application has specific modelling requirements such as incorporating uncertainty methods and integration with the available water resources infrastructure. Hence, the development of water resources management model to enable real-time abstraction management focuses on two parts. The first part investigates the development of a suitable real-time stochastic flow forecast model by combining a conceptual rainfall-runoff model with Bayesian based uncertainty analysis method. The second part focuses on devising a water resource management model and integrating it with the flow forecast model to enable efficient implementation of real-time abstraction management scheme and understanding implications of the scheme on water resources system.

1.2 Abstraction management – diffuse pollution

Historically, water quality issues have received less attention than water quantity in general. However, in recent decades policy makers, scientists and the public have begun to recognize the significant role water quality plays in economic, social and environmental developments and has consequently attracted a lot of attention (du Plessis 2017). Over the past few decades, improvements have been observed in the quality of river waters in UK. Nonetheless, diffuse pollutants such as fecal indicator organisms (FIOs) and pesticides originating from farmlands remain a problem for the management of river systems and can have serious financial consequences for water companies due to increased cost of drinking water treatment and compliance to

drinking water quality standards. The National Audit Office (NAO) has estimated the cumulative cost of water pollution to be between £700million and £1.3 billion a year (OFWAT 2011). Intensified pesticide application rates and new emerging products on the market have increased pesticide levels found in raw drinking water (Carvalho 2017). Weather impacts can also have a dramatic impact on pesticide runoff or application e.g. heavy rain at key planting period increases application of the slug control chemical metaldehyde. Most of this pollution is in the form of diffuse pollution from farmlands with insensitive agricultural practices, which contaminates surface and ground water supplies of drinking water. Treatment of contaminated water to comply with drinking water standards imposes considerable capital and operating costs on water companies and contributes significantly to the industry's carbon footprint. In the European Union (EU), the Drinking Water Directive (DWD; 98/83/EC) sets out the maximum allowed concentration (MAC) in treated drinking water for an individual pesticide as 0.1 µg/l (EC 1998). The DWD also states that total concentration of all pesticides should not exceed 0.5 µg/l.

Diffuse agricultural pollution is widely recognized as a significant threat to the quality of water resources, including in catchments used for drinking water supply (Castle et al. 2017). Metaldehyde is a soluble synthetic aldehyde pesticide used globally in agriculture which has caused recent concern due to high observed levels (exceeding the European and UK standards for pesticides in drinking water value of 0.1µg/l) in surface waters utilized for potable water supply (Kay & Grayson 2014). This study describes the development of a new travel time based physically distributed metaldehyde prediction model which aims to describe the short term fluctuations of metaldehyde concentrations in surface waters caused by rainfall runoff events. This will enable water infrastructure operators to consider informed control decisions in order to improve the quality of abstracted surface water. Several studies in the literature have revealed the significant role runoff plays in transporting pesticides and the occurrences of short lived peak pesticide concentrations following rainfall events (Huber et al. 1998; Huber et al. 2000; Bach et al. 2001; Wu et al. 2004; Heathwaite et al. 2005). Development of metaldehyde prediction model that

aims to predict short lived concentrations with a view to inform surface water abstraction management requires detailed understanding of the short-term dynamics involved in runoff based pesticide generation and transport processes, which include transport of event water through surface runoff and drains. Thus, the development of the metaldehyde prediction model in this study mainly focuses on the following points.

- Enabling physically distributed representation of runoff and metaldehyde generation thoughtout the catchment
- Improved representation of spatiotemporal variability of pollutant transport
- Understanding temporal dynamics of short lived peak pollutant levels at catchment scale
- Investigate spatial representations of diffuse pollutant generation areas and associated risks to water supplies at catchment scale
- Collection of new high resolution water quality dataset in study catchment

1.3 Aims and objectives

The overall aim of this study is to develop an advanced surface water management scheme based on real-time information and modelling approaches. This enables surface water abstraction volumes to be adjusted based on predicted availability of water and metaldehyde concentrations in the river at abstraction sites. Additionally, uncertainties associated with the developed real-time models will be analysed and presented to help in making risk aware decisions. Sub-aims and corresponding objectives of this study are presented in Table 1.1.

Table 1.1. Sub-aims and corresponding objectives

Sub-Aims	Objectives
a) Investigate the use of hydrological forecasting in order to maximize the amount of water	I. Identify a rainfall-runoff modelling approach, data assimilation and uncertainty

abstracted in surface water catchments and develop smarter abstraction management scheme that enables to vary abstraction volumes based on availability of water in the river.		analysis methods suitable for real-time abstraction management and assess and enable operational suitability.
	II.	Develop a real-time stochastic flow forecast model by combining a conceptual rainfall-runoff model with Bayesian based uncertainty analysis method.
	III.	Develop a water resources management model and integrate it with flow forecast model with a view to assess the potential benefits of the real-time abstraction management scheme and investigate implications on water resources.
b) Develop a new travel time based physically distributed catchment scale metaldehyde prediction model and improve understandings of short term fluctuations in metaldehyde concentrations at surface water catchment outlets caused by rainfall runoff events.	I.	Assess currently available water quality modelling practices and water quality datasets with regards to their applicability for describing short term fluctuations in pollutant concentrations at catchment scale.
	II.	Enable physically distributed representation of runoff and metaldehyde generation throughout the catchment and develop improved representation of spatiotemporal variability of pollutant transport.

	III.	Calibrate and validate the metaldehyde model using newly collected high resolution water quality data in the study catchment and assess operational suitability of the model to enable smarter abstraction management.
c) Investigate propagation of catchment scale spatially distributed input and parameter uncertainty in the metaldhyde prediction model and enable risk aware abstraction management decision making.	I.	Define and parametrize probability distribution functions of identified model input and parameter uncertainties, and generate realizations from predefined probability distributions to represent uncertainties in inputs and parameters.
	II.	Analyse propagation of spatially distributed input and parameter uncertainties in the metaldehyde prediction model.
	III.	Summarize results from uncertainty analysis using various statistical variables and asses relative uncertainty contributions from each input and parameter.

1.4 Thesis structure

The structure of the thesis is illustrated in Table 1.2. The thesis begins by providing an introduction to the research in this chapter. This describes mainly the problem statements and the motivations driving the research work. It also details the aims and objectives that this study addresses and presents the objectives in parallel with the chapters where they are addressed in the thesis.

Chapter 2 provides a review of a literature and discusses identified knowledge gaps and research questions. Chapters 3 - 5 are the three core chapters of the thesis and describe the development of real-time abstraction management schemes to enable smarter surface water abstraction management. Chapters 2 - 5 are briefly summarised and their association with the research objectives are shown in Table 1.2. Chapter 6 is the business case and implementation chapter which aims to set out a business case for implementation and discusses potential exploitation of the developed real-time abstraction management scheme in catchments throughout Severn Trent Water Ltd (STW) region. Chapter 7 is the final chapter which provides an over-arching summary and conclusion of the work described in the thesis. Limitations of the study and recommended future works in the area of surface water abstraction management are described in this chapter.

1.5 Publication and research dissemination

Findings from this study have been disseminated in several academic and industrial platforms. The publications from this work include:

- Asfaw, A., Maher, K. & Shucksmith, J.D., 2018. Modelling of metaldehyde concentrations in surface waters: A travel time based approach. Journal of Hydrology, 562, pp.397–410. doi: 10.1016/j.jhydrol.2018.04.074

- Asfaw, A., Shucksmith, J. & Macdonald, K., 2016. Parameter Uncertainties in a Conceptual Rainfall-runoff Model and Implications on Surface Water Management

and Planning Decisions. Procedia Engineering, 154, pp. 299–307. doi: 10.1016/j.proeng.2016.07.479

Findings from the study have also been presented in several academic and industrial events including the European Geosciences Union General Assembly 2015 (April 2015, Vienna, Austria), International Conference on Hydroinformatics 2016 (August 2016, Incheon, South Korea), Institute of Water annual conference (July 2016, Birmingham), TWENTY65 Annual Conference (April 2018, Manchester). Findings of the study have also been presented on numerous occasions within Sheffield University and at Severn Trent Water Ltd.

Authorship

I hereby confirm that I am the primary contributor in the writing of each of the above listed papers including the design and conduct of the reported research in each paper.

Copyright

I hereby confirm that all the necessary permissions have been obtained from relevant publishers of the above listed published materials for use in this thesis.

Table 1.2. Descriptions of Chapters 2- 5 and objectives addressed under each chapter

	Objectives								
	a (I)	**a (II)**	**a (III)**	**b (I)**	**b (II)**	**b (III)**	**c (I)**	**c (II)**	**c (III)**
Chapter 2 - Literature Review									
This chapter presents a review of the literature on a range of areas associated with real-time rainfall-runoff modelling, uncertainty analysis methods and modelling of pollutants to assess existing works and set out research gaps that need to be addressed in this study.	✓			✓					
Chapter 3 - Flow Prediction to Inform Surface Water Abstraction									
This chapter presents investigations on the use of hydrological forecasting in order to maximize the amount of water abstracted in surface water catchments and development of smarter abstraction management scheme that enables to vary abstraction volumes based on availability of water in the river.	✓	✓	✓						
Chapter 4 - Modelling of Metaldehyde Concentrations in Surface Waters									
This chapter presents the development of a new travel time based physically distributed catchment scale metaldehyde prediction model and improvements in our understandings of short term fluctuations in metaldehyde concentrations at surface water catchment outlets caused by rainfall runoff events.				✓	✓	✓			
Chapter 5 - Uncertainty in Metaldehyde Prediction Model									
This chapter presents investigations of propagation of catchment scale spatially distributed input and parameter uncertainty in the metaldhyde prediction model. This enables risk aware abstraction management decisions.							✓	✓	✓

Chapter 2

Literature Review

This chapter presents a review of the literature on a range of areas associated with real-time rainfall-runoff modelling, uncertainty analysis methods and catchment scale modelling of diffuse pollutants. It starts with a brief overview of the regulatory changes associated with the changing water environment and implications on water resources management in the UK. The chapter then presents the various methodologies used in the literature to represent hydrological processes in rainfall-runoff models and their suitability for real-time modelling applications. Furthermore, various applications of real-time models in the literature and associated challenges are discussed. More focus has been given to identifying specific requirements of real-time models for abstraction management applications, and assessing the performance and suitability of available rainfall-runoff flow forecasting modelling approaches for use in real-time abstraction management (RTAM) applications. The water quality section in this chapter presents a review of the literature on the occurrence and monitoring of emerging pollutants, catchment scale modelling of diffuse pollutants and assessment of uncertainty associated with catchment scale diffuse pollution models. The main focus is assessing existing catchment scale modelling approaches and identifying their suitability and capabilities to model short term dynamics associated with runoff driven pollutant generation and transport suitable to inform a real-time abstraction scheme. Assessment of uncertainty associated with catchment scale water quality models is also discussed.

2.1 Challenges of the changing water environment and regulatory reforms

Effective management and utilization of available water resources is fundamental to society in terms of public health, commercial activity and environmental protection. In recent years, population growth, climate change and increasing pressures from emerging pollutants have become key water resource issues facing the water sector. Moderate drought events have been observed in the UK with an unusual frequency

over the last twenty years (Kendon et al. 2013). The 2010 – 2012 periods have been remarkable in climatic terms with the occurrences of unusual rainfall events, which have characteristically departed from the typical seasonal rainfall patterns in most parts of the UK. Particularly in 2012, a dry period followed by one of the wettest spring and summer on record have caused widespread water quality and quantity challenges for water supply utilities (CEH 2012). Due to the impacts of climate change, more severe and frequent droughts and changes to rainfall patterns are expected in the future. By the 2050s, it is anticipated that under the medium emissions scenario of UK Climate Projections (UKCP09) summer temperatures will increase and summer rainfall will decrease in most parts of the UK with more frequent occurrences of short duration drought conditions (OFWAT & EA 2011). Increased occurrences of short term dry periods cause challenges in infilling water resource reservoirs and are likely to trigger drought management actions more frequently. This indicates the need to improve the existing management of water resources to make operations more resilient to these challenges. In 1999 the Restoring Sustainable Abstraction (RSA) program was set out to investigate the extent of environmental damage on rivers and wetland sites suspected of being affected by over-abstraction in England and Wales, with the aim of finding and implementing remedial solutions. In recent years, this has caused sustainability reductions to abstraction licences with a view to protect international/national designated conservation sites (Habitats Directive, Sites of Special Scientific Interest or Biodiversity 2020 sites), and deliver Water Framework Directive (WFD) objectives. Furthermore, changes in demand due to population growth and impacts of climate change are likely to deteriorate the status of licences currently deemed sustainable. Environment regulators in the UK are working with the water industry through WINEP to implement schemes that are aimed at ensuring water courses in the UK meet European and national targets related to water (section 1.1).

Furthermore, the UK Department for Environment, Food and Rural Affairs (DEFRA) recently published the Natural Environment White Paper that set out an evidence-based case for change in water abstraction management in the context of a future in

which there will be less water available for people, businesses and the environment (DEFRA 2011a). This document made clear the UK Government's intention to reform the current water abstraction management system with a view to enabling a more responsive and flexible approach to abstract water.

The Water White Paper (DEFRA 2011b) set out a proposed direction, principles and processes for water abstraction management reform aiming to promote resilient and sustainable use of water while protecting the environment. This paper identified that the way we manage water resources needs to be reconsidered to ensure continued protection of future water ecosystems and make water available for abstraction. While demand management was stated to have an important role to play, tackling unsustainable abstraction and development of new water resources have been identified as key actions to tackle these water resource challenges (DEFRA 2011b). As a significant stakeholder in the water environment, it was recommended that water companies adopt more dynamic, flexible and innovative approach to carry out water abstraction.

The UK Environment Agency investigated current and future water availability and supported a proposal to reform abstraction licensing (EA 2011). As a result, DEFRA worked with key stakeholders to identify two options of abstraction reform which are named 'Current System Plus' and 'Water Share'. Different approaches and methods were used to link abstraction to water availability in each of these reform options. The Current System Plus option is similar to the current system in using daily and annual volumetric constraints (i.e. periodic daily constraints and annual volumetric constraints in current licences), but it is modified to further restrict abstraction during low flows and allow more water to be abstracted when more is available in the source. The Water Share option aims to achieve shared responsibility for water resources in the catchment. This option gives abstractors a share of the available water in the catchment rather than an absolute amount (DEFRA 2014). Both options aim to increase the amount of water available for use by linking abstraction to water availability in surface waters whilst protecting the environment during periods of low flow. The Water Share option allows more active water trading as compared to

Current System Plus, though both options are believed to make water trading between abstractors quicker and easier. Following the dissemination of the proposed options for water abstraction licensing reform, the UK Government has conducted a series of open consultation and workshops throughout the country. Responses from a range of abstractors and other stakeholders supported most of the proposals in principle (DEFRA 2014). The Water companies' main concern has focused on the resulting uncertainty of the reforms on their deployable output, defined as the overall output of a water source or a group of sources as constrained by environment, abstraction licence, water quality and infrastructure capacities. The complexities of operation and increased implementation costs mainly associated with the Water Share option have also been raised as concerns. As major water abstractors, water companies are highly vulnerable to the challenges of climate change, increased demand and associated abstraction reforms, so there is a need to look for innovative ways to improve their abstraction management. Increasing availability of data capturing technologies and computational capabilities provide potential to devise catchment level models with a view to help prepare abstraction management systems for these challenges.

In addition to abstraction reform, a proposal to increase competition across wholesalers, part of water companies that governs water production infrastructures, has also been introduced in the Water White Paper to help in stimulating a more dynamic wholesale market for alternative water resources (DEFRA 2011b). This reform is also believed to incentivise water companies to look for innovative options to secure future water resources rather than adopting capital intensive projects (such as developing alternative water resources) to meet future demands. Moreover, upstream competition in the water sector helps to maximize the economic value of available water resources by giving more value to water, whereas values in the current system only reflect works done to get the water to the tap. The draft water bill was published in 2012 with the main objective of delivering the key commitments set out in the Water White Paper and is likely to facilitate the transition to a resilient and sustainable water sector (DEFRA 2012). Technological advancements in capturing

hydro-meteorological data, computational capabilities and improved theoretical understandings have boosted the development of real-time monitoring and modelling methods, which the water industry can exploit to meet the demands imposed by these reforms and efficiently manage water resources. Suitable flow forecasting models that are capable of informing surface water abstractions can be developed by employing uncertainty analysis and data assimilation techniques on rainfall runoff-models. Operationally suitable real-time abstraction management scheme can then be enabled by coupling these models with water resource management models that represent licence conditions, infrastructure capacities and reservoir operation rules.

2.2 Rainfall-runoff modelling

Rainfall-runoff models are widely used in water resources management for a wide range of applications and play a significant role in informing water resources management and planning decisions in catchments (Beven 2012). Complex non-linear hydrological processes, which exhibit high spatial and temporal variations, are commonly represented using relatively simple structures in rainfall-runoff models. The purpose of model application, available data and the effort required for model development and application determine the level of hydrological system to be described in the model (e.g. spatiotemporally distributed detail or in a more aggregated way) (Beven 2012). For this reason, there is no a unique and seamless rainfall-runoff model fit for all purposes, but there are several descriptions of the hydrological system (rainfall-runoff models) each of which may be suitable for addressing different set of scientific questions. Different rainfall-runoff models differ in their aggregation level of the structure and processes of the hydrologic system, process representations and mathematical structures and formulations.

2.2.1 Types of rainfall-runoff models

Given the wide range of rainfall-runoff models detailed in the literature and their increasing availability as software modelling packages, the key questions relates to deciding which are the most suitable models and what are the limitations of these models. To answer these questions, it is primarily important to look into the generic

classifications of available rainfall-runoff models. Many different ways of categorizing rainfall-runoff models is available in the literature, but this section focusses on the two major classification methods, assessing suitability and limitations of model types in these categories. Rainfall-runoff models can be categorised into three main groups, namely: distributed physical based models, conceptual models and input-output or black-box models (Beven 2012). The methodologies used to represent hydrological processes in each of these model types and their suitability and limitations for use in a range of applications are discussed below.

Distributed physical based models are based on our current best understanding of hydrological processes and try to represent real world physical processes in catchments. These models are also called mechanistic models as they incorporate principles that govern physical processes in the catchment. System states in the catchment are represented using state variables, which are measurable and are spatially and temporally distributed (Devia et al. 2015). The spatiotemporal dynamics of water movement in the catchment are often represented using Saint-Venant equations in models used for theoretical and applied studies. Finite difference method is commonly used to numerically solve these governing equations in models (Bell & Moore 1998a). Large numbers of parameters employed in physically based models are derived from various temporally and spatially distributed measured variables such as soil moisture content, initial water depth, topography, soil type and dimensions of river networks in the catchment. As a result, evaluation of the parameters that are used to describe catchment characteristics in the model is often an extensive task (Feyen et al. 2000). Parameters used to describe catchment characteristics in physically based models have direct physical interpretations, which provide useful information for modellers and enable direct measurements of most of the parameters.

In contrast, conceptual models are developed based on data and use mathematical concepts to represent hydrological processes in a suitable structure and parameters set. In these models, hydrological process components are represented using a

number of interconnected reservoirs (Lan Anh et al. 2008). The conceptual reservoirs are recharged using rainfall, infiltration and percolation processes and are emptied by evaporation, drainage and surface and sub-surface runoff processes. Semi-empirical equations are mainly used to describe the interaction between the conceptual reservoirs (Siebert 1999). Calibration procedures are used in models to adjust model parameter values by forcing them with in the required margin of uncertainty to obtain representation of modelled process that satisfy pre-agreed goodness-of-fit criteria (Gan & Biftu 1996). Unlike physically based distributed models, most parameters in general conceptual models lack physical basis and thus cannot be inferred from direct measurements (Beven 2012). Hence, calibration procedures that require extensive hydrological and meteorological data are often used to estimate these parameters. The calibration process involves solving an inverse problem of estimating parameter values using recorded data on system responses such as discharge and water level data (Zhang et al. 2015). Different set of discharge or water level data is required to validate calibrated models, which is a process performed to verify that models are performing as expected in line with their design, objective and intended application (Gan & Biftu 1996; Zhang et al. 2015). Conceptual rainfall-runoff models are developed at different degrees of complexity with varying number of parameters depending on the intended application of the model. Over the past few decades, a wide range of applications have used conceptual rainfall-runoff models as it provides modellers a potential to achieve a reasonable accuracy in modelling hydrological processes with relatively simple computational complexity (Zhang et al. 2015; Devia et al. 2015). However, the amount of effort in developing conceptual rainfall-runoff models and the complexity of these models seem to have increased continually with the expanding availability of high computational power and advanced computational methodologies (Beven 2012),

In contrast, black-box models only consider hydrologic input and output data from catchments and do not attempt to represent internal hydrologic processes. These are highly empirical observation based models that use information derived from existing system input and output data to estimate the behaviour of the system. As a

result, catchment characteristics and hydrological processes are not considered at all and hence these models are also called data driven models (Sitterson et al. 2017; Xu 2002). These models use information derived from system input and output time-series data to derive a set of mathematical equations that are used to emulate the physical processes involved in the system. Black-box models are not suitable for application outside the calibrated boundary as they lack hydrological processes representations (Knapp et al. 1991). Artificial neural network and fuzzy logic models developed using machine learning techniques are typical examples of black-box models. Statistical models that derive functional relationship between input data and system responses using regression or correlation techniques can also be categorised as data driven or black-box models (Rajurkar et al. 2002).

A second consideration is whether to use these models in deterministic or stochastic setup, which is a model choice based on the output types delivered by the models. Deterministic models use one set of inputs and parameters and produce a single output for each time-step, whereas stochastic models produce a set of output values at each time step, which can be presented using various summary statistics (Shaw et al. 2017). Variations in stochastic model outputs are fundamentally generated due to model input, parameters and structural uncertainties. The vast majority of rainfall-runoff modelling applications have commonly used deterministic models due to its relatively simple setup and lesser computational demand. However, the growing interest in identifying the impact of model prediction uncertainties on water resource management decisions has led to an increase in use of stochastic models in recent years (Voge 2016). Moreover, rapidly advancing computational capability has enabled users to implement various uncertainty analysis methods (e.g. Bayesian methods) to enable stochastic model predictions. Variances or other measure of model output dispersion resulting from stochastic models provide additional information to water resource managers and enables them to make risk-aware decisions (Jonsdottir et al. 2006).

2.2.2 Real-time applications of rainfall-runoff models

In real-time modelling, models are applied to make predictions about future system states and outputs based not only on historical data but also using newly received and/or forecasted data. Knowledge on future state systems and outputs are valuable for pro-active water resource management and planning. Thus, real-time modelling is employed in a wide range of geo-science and environmental fields including hydrology (Seo et al. 2003; Cloke & Pappenberger 2009) and meteorology (Golding 2000; Thorndahl et al. 2010). Typical application of real-time modelling in these areas include providing warning systems of future events such as flooding (Penning-Rowsell et al. 2000). Amidst an intense period of real-time model development and application, several studies have focused on employing real-time modelling in urban drainage system and distribution networks (Henonin et al. 2013). In addition, most of the published studies on real-time rainfall-runoff modelling have examined its application for flood forecasting (e.g. Seo et al. 2009; Rogelis & Werner 2018). Less attention is given on applying real-time models for managing human interventions to the river systems, which is partially due to old perceptions on the availability of water and limited understanding of the changing water environment and impacts of our interventions (e.g. abstractions) on it. In recent years, there has been a significant improvement in our understanding of the needs of ecosystems in water and an increase in realization of the combined effects of climate change and a growing population (Morrison et al. 2009). As a result, abstraction management systems have been/are being reformed in a number of countries including United States, South Africa, Australia, Russia, England and Wales (Erfani et al. 2014). Fundamentally, the reforms aim to minimize impacts of abstraction on the environment while maximizing the use of available water resources.

A comprehensive study on utilizing available advanced hydro-meteorological data capturing technologies (e.g. radar rainfall and sensors) combined with real-time modelling capability to enable dynamic abstraction management is lacking. The use of flow prediction models for abstraction management application has specific modelling requirements due to a number of operational issues associated with

abstraction management. These challenges can be categorized into two main areas of abstraction management as follow: 1) Real-time models for abstraction management application need to accurately predict a variety of flow conditions, which requires the models to have flexibility to increase the capability of simulating the full range of service flows (flow range between minimum environmental flows and maximum abstraction capacity). As a result, the divergence of the models, which is defined as the relative accuracy of the model during the calibration and validation processes, needs to be measured over the service flow range rather than at high flows as used in flood forecasting applications (Vaze et al. 2011). Moreover, to keep the balance between enabling to abstract more water and avoid breaching abstraction licence conditions, uncertainties associated with forecast service flow range needs to be carefully analysed and presented to enable risk aware abstraction management decisions. 2) Effective implementation of real-time abstraction management scheme requires integrating water resource systems (such as reservoir operation rules, abstraction licence conditions, available storage volumes and pump and water main capacities) and forecasted flow data to determine the amount of water that can be abstracted at any specific period in real-time. This requires development and integration of a water resources management model with flow forecast models.

2.2.3 Comparison of rainfall-runoff models for real-time applications

Extensive research has been conducted in hydrology on the development of various types of hydrological models to represent the rainfall-runoff processes and their wide range of applications. While the main scientific objective of most works have been to obtain better understanding of the complexity of catchment responses and informing strategic water resources management and planning, some of the researches have been motivated by the need to apply these models for short term water resource management and informing operational decisions (Burn et al 1999; Vaze et al. 2011). Given the wide range variability of rainfall-runoff models, a question arises as what type of model is the most suitable for real-time use to inform short term water resource management and operational decisions. The use of any of the rainfall-runoff model types in real-time application requires consideration of factors such as

computational requirement, implementation of real-time updating and forecasting, and operational suitability (Burn et al 1999). Due to the high number of parameters involved and the associated level of complexity, the use of physics-based models such as Systeme Hydrologique Europeen (SHE) (Abbott et al. 1986) and HYSIM (Manley 1978) are computationally intensive and require large datasets. Furthermore, these models do not readily lend themselves to having their forecasts updated in real-time in a computationally efficient manner. Consequently, the choice of a rainfall-runoff model for most real-time applications currently rests between conceptual and black box models.

Considerable number of studies to assess the suitability of rainfall-runoff models and compare them for use in real-time forecasting under various conditions have been conducted (Shamseldin & O'connor 2003; Todini 2005; Jorgeson & Julien 2005). Furthermore, a number of studies have compared available rainfall-runoff models for operational flow forecasting purposes (Goswami et al. 2005; Te Linde et al. 2008). However, most of these studies are focused on large catchments (with an area of more than 1000 km^2) and coarse time resolution (6-12 hrs time steps), which makes them less relevant for a typical situation in the UK, where flow estimations are required at the outlet of relatively small catchments and thus finer time resolutions (in the region of 1-2hrs). The studies conducted by the Institute of Hydrology to compare various rainfall-runoff models for flood forecasting purposes across UK catchments (Moore et al. 1992; Moore & Bell 2001; Young 1997) are of much greater relevance to this study. These studies have assessed EA operational models including Thames Catchment Model (TCM) (Moore & Bell 2001), the Midlands Catchment Runoff Model (MCRM) (Moore & Bell 2001), the Probability Distributed Moisture model (PDM), the Isolated Event Model (IEM), the ISO-function model and black-box models (e.g. Transfer Function, TF and Physically Realizable Transfer Function, PRTF) (Beven 2012), and other rainfall-runoff models such as the US National Weather Service Sacramento Model (Burnash 1995), the NAM model (Agrawal & Desmukh 2016), the Grid Model (Bell & Moore 1998b). These studies highlighted that the type of catchment processes explicitly represented (such as

infiltration, percolation, surface runoff, subsurface and groundwater flow processes) or the effectiveness of representing aggregated processes (such as semi distributed, lumped, black-box and various conceptual model representations) are the main factor which determine the appropriate level of complexity in the models. Otherwise, the reviewed conceptual rainfall-runoff models are found to be similar despite the various 'brand names' (Moore & Bell 2001). Consequently, the overall conclusions of the model reviews have been informative rather than judgemental giving further useful information associated with performances of each model rather than providing preferences.

Though forecast accuracy has been used as the main criteria of comparison, model configuration, initialization and calibration are also considered in the reviews. Overall, no one model is found to consistently outperform all the other models across all catchments in these reviews. Peak flows in extreme flood events such as the Easter 1998 flood in the UK Midlands are underestimated by most of the models with NWS underestimating more than most and Grid Model performing best. TCM is found to perform better when the R^2 statistic is used to compare models in forecast mode. Although and PDM is in the second tier of models, it performed very well when threshold critical success index (CSI) statistic is used for comparison, indicating its suitability in forecasting flows at different peak levels. TCM is suggested to be more applicable to more complex catchments where flow responses come from heterogeneous areal response zones. The reviews identified PDM and MRCM as models which provide good balance between simplicity and complexity. PDM's capability to represent a large range of catchment behaviour through the wide selection of model structures it accommodates has helped its wide range of application as a flow forecasting model. The R^2 performance of black-box models TF and PRTF are found to be poor, but improved significantly when error-prediction updating is used. IEM has performed particularly well as compared to other simple models mainly in terms of threshold CSI. Grid Model was found to be slow to run and calibrate when applied to large catchments and is suggested that its primarily use is as a research tool (Bell et al. 2001). Following this review of the models and

discussions with staffs at Severn Trent Water Ltd, we have decided to use the PDM model in this study.

2.3 Probability distributed model

The Probability Distributed Model (PDM) is a conceptual rainfall-runoff model principally based on probability distributed moisture capacities and routing of runoff and drainage through routing stores. The design of the model algorithm enables the representation of a range of catchment responses in different ways. A series of publications (Moore 1985; Moore & Bell 2002; Moore 2007) provide detail of the theoretical background of the model.

The search for models that are parsimonious of parameters and are capable of accurately predicting various range of flows has been the source of PDM rainfall-runoff model. With this aim, Moore & Clarke (1981) have conceptualized the representation of runoff production from catchments using a probability distribution of moisture capacity rather than the usual catchment averaged single soil moisture store representation. This led to a more realistic representation of runoff production over the catchment and enabled optimization of the variability in soil moisture storage distribution across the catchment. Moore (1985) has further simplified the theory with a view to improve the representation of runoff generation and routing processes. The major improvement being the development of a procedure to enable interaction between the soil moisture storages to equalize water depth across the distributed storages throughout the catchment (Moore 1985), which has led to the formulation employed in the current PDM model. In the previous setup of the model, probability distributed time of travel (considered as instantaneous unit hydrograph) has been used to route direct runoff to catchment outlet using convolution operation. The parallel routing formulation, which uses soil moisture feedback to separate direct runoff and groundwater flow of total runoff and enable non-linear runoff routing procedure has been used in the current model. This method has initially been developed based on concepts from Dooge (1973) and Dooge & O'Kane (2003). The use of non-linear routing procedure instead of convolution to route runoff has

enabled real-time updating much easier. Moreover, improvements made by Moore (1985) to PDM have also included addition of parameter interdependence to the model, introduction of parallel configuring of routing stores and inclusion of more complex recharge functions, which are incorporated to represent a wide range of catchment processes. This raised the need to rely on manual calibration with the use of visual support and only use automatic optimization for refinement of selected parameters. In recent years, PDM has evolved as a fairly general conceptual rainfall-runoff model well suited for real-time flow forecasting with relatively few parameters. Currently, PDM has widespread application throughout the world (Teuling et al. 2004; McIntyre et al. 2005; Cabus 2008) and earned a solid reputation as a real-time flow forecasting model in the UK.

2.4 Data assimilation

Real-time modelling provides an opportunity to incorporate new observed system response data thereby increasing the calibration dataset with a view to improve model predictions, which is known as data assimilation. It mainly functions by comparing model outputs with new independent observations in order to modify certain states/parameters of the model and by considering the discrepancy between the latest observed data and model outputs as feedback to mitigate real-time model uncertainties (Moradkhani & Sorooshian 2009). Data assimilation is a widely used technique in different areas of geosciences with most applications involved in meteorological and hydrological forecasts. Advanced telemetry systems, increased availability of hydrological data observations (e.g. from radar and satellite) and the need for better forecast accuracy have contributed to the increased use of data assimilation as a powerful tool in flow forecasting techniques (Brocca et al. 2013). Over the past twenty years, considerable research has been carried out on the use of data assimilation techniques to improve hydrologic model predictions. Lee et al. (2011) and Lee et al. (2012) reviewed the main developments in the area including assimilation of a range of in-situ and remotely sensed observations into hydrological models. Flow, water level or soil moisture data can be used in data assimilation techniques implemented in rainfall-runoff models depending on data availability and

model structure and output types. The most relevant studies in the context of the hydrologic application in this study are assimilation of observed flow to improve rainfall-runoff forecasts (e.g. Seo et al. 2003; Weerts & El Serafy 2006; Clark et al. 2008; Seo et al. 2009; Lee et al. 2012; McMillan et al. 2013). Different types of data assimilation techniques are available in the literature that can be used to assimilate observed flow data to improve flow predictions. In this section, the state correction and error forecasting methods that are employed in the PDM rainfall-runoff model are discussed (Adediran 2015).

2.4.1 State correction method

The state correction method focusses on updating state variables in the model that define the initial state of the model before making forecasts. The term state in the state correction model refers to these variables in the model which define the state of the system and mediate between model input and outputs(Szollosi-Nagy 1976). The type of state variables and their physical interpretations vary depending on the structural formulation of models. For example, typical state variables in conceptual rainfall-runoff models such as PDM are surface water and groundwater stores, which represent storage components in the hydrologic process (Moore 2007). In the state correction method, mismatches between model predicted and observed values are assumed to have been caused by incorrect state variables. Hence, the feedback derived from the error between observed and predicted values is used to correct the state variables (Adediran 2015). State correction method is fundamentally applied based on the Kalman filter algorithm as detailed in a number of literatures (Jazwinski 1970; Gleb 1974; Wang et al. 2009). An extended form of Kalman filter based on a linearization approximation is often implemented in non-linear dynamic models such as PDM. This enables adjustment of state variables using a simpler and more intuitive scheme than the more complex and formal extensions of the Kalman filter implementation in non-linear dynamic models (Clark et al. 2008). The PDM rainfall-runoff model provides an option to use the linearization approximation scheme, which is also called empirical state adjustment scheme that enables to make sensible physical interpretations of state variable adjustments (Moore et al. 2007). A typical

simple implementation of Kalman filter using a linearization approximation is used in the PDM rainfall-runoff model. In this method, proportions of the error between observed and simulated flows are used to adjust surface and groundwater stores based on their contribution to the total flow (Moore et al. 2007; Adediran 2015). This is mathematically expressed as

$$q_b^* = q_b + \alpha g_b \varepsilon \tag{2.1}$$

$$q_s^* = q_s + (1 - \alpha) g_s \varepsilon \tag{2.2}$$

where

$$\alpha = q_b/(q_s + q_b) \tag{2.3}$$

q_b – groundwater contribution to total flow, q_s – surface water contribution to total flow, q_b^* – adjusted groundwater contribution to total flow, q_s^* – adjusted surface water contribution to total flow, g_b - gain coefficient to groundwater contribution and g_s - gain coefficient to surface water contribution, ε – error in predicted flow ($Q - q$) where Q – observed flow and $q = q_s + q_b$.

When values of gain coefficients (g_b and g_s) are equal to one, the combination of adjusted surface and groundwater contributions matches observed flow values. Different adjustments to the gain coefficients allow varying proportions of errors attributed to surface water and groundwater contributions, thus g_b and g_s can be used as forecast model calibration parameters that can be used to achieve best-fit between observed flows and forecasts. This can be implemented by using a more general expression for α as

$$\alpha = \frac{q_b}{\beta_1 q_s + \beta_2 q_b} \tag{2.4}$$

where, β_1 and β_2 are weight parameters that can be used to allocate the error distribution towards or away from one of the flow contributions. Computation of parameter α and the flow contribution adjustments are carried out at each new data point. Based on the method used to compute the value of α, three different flow adjustment schemes can then be defined as the proportional adjustment scheme (equation (2.3)), super-proportional adjustment scheme (equation (2.4)) and the simplest non-proportional adjustment scheme, which involves replacing α and $(1-\alpha)$ in equation (2.3) and equation (2.4) by one. The fundamental Kalman filter adjustment technique, which is based on updating system states based on new information derived from new observations (model error combined with gain coefficients), is used in all of these adjustment schemes.

2.4.2 Error prediction method

The error prediction method focuses on analysing error structures with a view to enable future error predictions, which can then be used to improve forecasts. As a result, model errors are not attributed to any of the model components in this method and no adjustments are made to model variables. The development of error prediction models is based on the persistence of error characteristics from conceptual rainfall-runoff models, which results in sequences of positive or negative errors and identifiable trends (Casale & Margottini 2012). Information on structural dependences of these error sequences are used to develop error prediction models. The error prediction method is developed totally external to models and thus is suitable for use in conjunction with any types of models. The PDM rainfall-runoff model has incorporated the error prediction method as an option for updating forecast simulations (Moore et al. 2007).

Moore et al. (2007), Casale & Margottini (2012) and Adediran (2015) have described the details of the procedures involved in the operation of error prediction method to update flow forecasts based on predicted errors, which are summarized below. Observed flow ($Q_{t_o|t}$) and forecasted flow (q_{t_o+t}) made using the rainfall-runoff

model in simulation mode with lead time *t* can be related using the following expression

$$Q_{t_o+t} = q_{t_o+t} + \varepsilon_{t_o+t} \tag{2.5}$$

A better forecast accuracy can be achieved if the error (ε_{t_o+t}) can be predicted using a forecast model and used to adjust predicted flow (q_{t_o+t}). Considering $\varepsilon_{t_o+t|t}$ as the predicted error at lead time *t* made using error prediction model, then the forecasted flow using a forecast-mode rainfall-runoff model at lead time *t* can be expressed as

$$q_{t_o+t|t} = q_{t_o+t} + \varepsilon_{t_o+t|t} \tag{2.6}$$

and the forecast-mode model error is expressed as

$$e_{t_o+t|t} = Q_{t_o+t} - q_{t_o+t|t} \tag{2.7}$$

Depending on the performance of the error prediction model used, error from forecast-mode model at lead time *t* $(e_{t_o+t|t})$ should be less than the error from the simulation-mode model at lead time *t*

$$\varepsilon_{t_o+t} = Q_{t_o+t} - q_{t_o+t} \tag{2.8}$$

thus improving flow forecast performances.

The autoregressive (AR) and autoregressive moving average (ARMA) models are typically used error prediction models used to improve model performances as they provide a suitable structure to incorporate dependences on historical errors from simulation models. In general, ARMA model is considered as a more parsimonious choice as it provides the same level of approximation as AR but with less number of

parameters. The error prediction method in PDM is implemented using ARMA. ARMA parameters that are used to adjust and improve flow forecasts are estimated using automatic optimisation in PDM (Moore 2007).

2.5 Uncertainty analysis in rainfall-runoff modelling

Accurate and reliable flow forecasting is critical as it governs the reliability of water management decisions, and hence impacting on the associated benefits or losses which accrue from them. Calibrated parameter values are often used in hydrological models without considering parameter uncertainties and their implications in decisions based on uncertain model outputs are not assessed. Flow prediction uncertainties originating from inaccurate structural and parametric representation of catchments in models have a major influence on the reliability of forecasts (Beven & Freer 2001). Other factors that bring uncertainties in real-time forecasting include inputs, initial conditions (e.g., soil moisture states) and observed flow data used for model calibration (Hutton et al. 2014). Several studies have analysed these sources of uncertainties and developed various methods of dealing with them. These include first-order approximations and multi-normal distributions (Kuczera & Mroczkowski 1998), simple uniform random sampling over feasible parameter space (Uhlenbrook et al. 1999), parameter sampling using Markov chain Monte Carlo methods (Campbell 2001; Vrugt 2003; Feyen et al. 2007).

The choice of an uncertainty analysis tool to be used in rainfall-runoff models depends on the following four main factors: i) the type of rainfall-runoff model used; ii) uncertainty source to be dealt with; iii) uncertainty representation (e.g. probability theory); iv) the purpose of uncertainty analysis; v) data and resource availability (e.g. computational power) (Reichert 2014). Pappenberger et al. (2006) has described the processes involved in selecting a suitable uncertainty analysis tool using a decision tree. First-order approximations and multi-normal distributions, which are uncertainty analysis methods based on traditional statistical theory (X. Zhao et al. 2011), are not effective in dealing with non-linear and complex rainfall-runoff models. Explorations of parameter space using simple uniform random sampling

methods are relatively easy but are not efficient in arriving at satisfactory solutions that are able to identify the shape of the resulting uncertainty distribution (Tomassini et al. 2007). In contrast, Markov chain Monte Carlo (MCMC) methods use adaptive Markov chains to efficiently sample from the target posterior distribution and thus enable application of inferences and optimisations that require highly efficient sampling methods (Reichert 2014). The posterior parameter distribution is used to quantify model parameter uncertainty and its formulation varies depending on the type of uncertainty representation or inference used. Bayesian inference, which is a statistical inference method based on Bayes' theorem that is used to update probability of prior beliefs/hypothesis as more data is available, provides an ideal framework to analyse parameter uncertainties of conceptual rainfall-runoff models by combing prior knowledge of model parameter values with model outputs and observed flow data (Feyen et al. 2007). This is done by conditioning the joint distribution of model parameters (prior) and model outputs (likelihood) by observed data.

$$p(\theta|D) = \frac{p(D|\theta)p(\theta)}{p(D)} \propto p(D|\theta)p(\theta) \tag{2.9}$$

Where, $p(\Theta|D)$ is posterior probability, $p(D|\Theta)$ is likelihood probability, $p(\Theta)$ is prior probability, D is observed data and Θ is parameter.

Thus, the posterior distribution of the parameter is formulated as the direct proportion of the product of prior and likelihood function (equation (2.9)) with observed data used in the likelihood function as shown in Figure 2.1. Difficulties involved in efficient sampling of the resulting posterior distribution or estimating its parameters have hindered the widespread application of Bayesian inference in hydrology. However, recent advances in the development of MCMC methods that enable efficient sampling from the posterior distribution have enabled to overcome these difficulties. Application of such MCMC techniques requires tuning algorithmic parameters to improve the convergence rate of the chain and careful diagnosis of

convergence (Marshall et al. 2004). Detailed description of a state-of-the-art uncertainty analysis tool based on Bayesian inference is given in section 2.5.2.

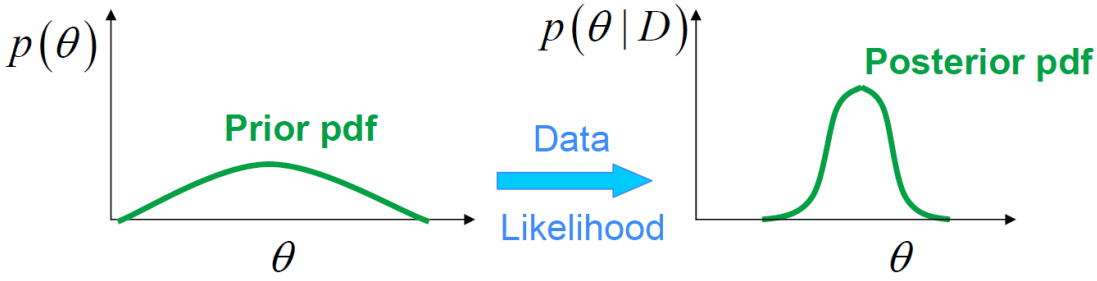


Figure 2.1. Formulation of posterior distribution using Bayes' theorem

2.5.1 Markov chain Monte Carlo method

There are only a small number of posterior distributions and models for which Bayesian inference can be conducted analytically. For this reason, numerical techniques are very important for implementation of Bayesian inference and most applications are based on Monte Carlo simulation (Reichert 2014). In Monte Carlo based methods, a sample is drawn from the posterior distribution and properties of the population distribution are approximated based on properties of the sample. Of particular importance to application of Bayesian inference are Markov Chain Monte Carlo (MCMC) techniques, which are a special type of random process, that construct a Markov chain which asymptotically converges to the posterior distribution. MCMC methods use Markov chain, which contains a sequence of random variables that is characterised by the probability density of each random variable in each chain conditional on the immediate predecessor and independent of any other random variable of the chain (Campbell 2001). A Markov chain can be characterised by its transition probability density function, $f_{trans}(\theta^{new}, \theta^{old})$, which is the conditional probability density for reaching a new state θ^{new} given a previous state of θ^{old}. Homogeneous Markov chains, which have similar transition probability density for the whole chain, result in a stationary distribution f_{stat}, which is invariant for each step (Reichert 2014):

$$\int f_{stat}(\theta') f_{trans}(\theta, \theta') d\theta' = f_{stat}(\theta) \tag{2.10}$$

Equation (2.10) is used to show that if a Markov chain has a stationary distribution and it is aperiodic and recurrent, then the Markov chain will converge to its unique stationary distribution. This stationary distribution is also often referred as equilibrium or invariant distribution. The condition of detailed balance can be used as a sufficient but not a necessary condition to check if a Markov chain has a stationary distribution and is expressed as

$$f_{stat}(\theta) f_{trans}(\theta', \theta) = f_{stat}(\theta') f_{trans}(\theta, \theta') \tag{2.11}$$

In equation (2.11), the expression $f_{trans}(\theta', \theta)$ represents the probability of Markov chain, which has a stationary distribution of $f_{stat}(\theta)$, moving from $\theta \; to \; \theta'$. Hemce, the equation of detailed balance shown in equation (2.11) will be met when the probabilities of Markov chain moving from $\theta \; to \; \theta'$ and $\theta' \; to \; \theta$ is the same. As a result, equation (2.11) represents the reversibility that for every two parameter states $\theta \; and \; \theta'$, the probabilities of transitioning to one of the parameter states or the other are the same. A Markov chain will converge to a unique stationary distribution given that it is periodic, recurrent and has a stationary distribution. Tierney (1994) and Gamerman (1997) have showed that most of chains in MCMC simulations fulfill these characteristics and thus can be used as a representative samples from a stationary distribution. The main objective in the use of MCMC simulations in Bayesian inference is to generate a Markov chain that has a posterior distribution as its stationary distribution as shown in Figure 2.2. A sample of this Markov chain can thus be used to estimate characteristics of the posterior distribution.

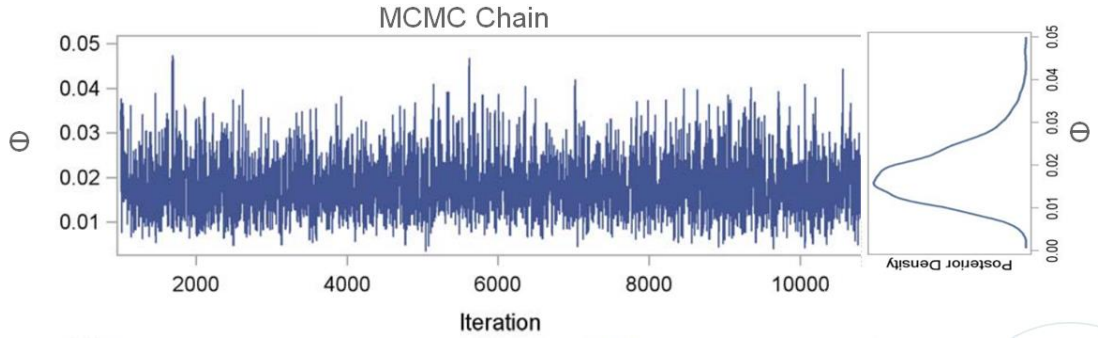


Figure 2.2. Markov chain and its convergence into stationary distribution

Some MCMC methods split the transition density function, $f_{trans}(\theta', \theta)$ into proposal density function, $f_{prop}(\theta', \theta)$, and an acceptance probability, $P_{accept}(\theta', \theta)$ (Tierney 1994; Gelman et al. 2014). Hence, the probability of staying at the same state θ is given by

$$P_{stay}(\theta) = 1 - \int f_{prop}(\theta', \theta)\, P_{accept}(\theta', \theta) d\theta' \tag{2.12}$$

Denoting the transition density function as

$$f_{trans}(\theta', \theta) = f_{prop}(\theta', \theta) P_{accept}(\theta', \theta) + P_{stay}(\theta)\delta(\theta - \theta') \tag{2.13}$$

For a given proposal density function, application of the condition of detailed balance (equation (2.11)) mean that the following condition has to be fulfilled by the acceptance probability

$$P_{accept}(\theta', \theta) = \frac{f_{stat}(\theta')}{f_{stat}(\theta)} \frac{f_{prop}(\theta, \theta')}{f_{prop}(\theta', \theta)} P_{accept}(\theta, \theta') \tag{2.14}$$

If this condition is fulfilled by the acceptance probability, the transition density (equation (2.13)) fulfils the stationary condition (equation (2.10)). Hence, the use of any proposal distribution combined with an acceptance probability, which fulfils the

condition expressed in equation (2.14), results in a Markov chain that can be used as a sample from a given stationary distribution (Gelman et al. 2014). Only the ratio of the stationary density is needed in equation (2.14) and thus a numerical scheme can be applied even if all other parameters used for normalization are not known (Reichert 2014). This characteristic is very important in application of Bayesian inference as most of posterior distribution functions normalization factors are not known. Two different implementations of this general method are Metropolis sampling and Metropolis-Hastings sampling which are discussed in detail by Hasting (1970), Gamerman (1997) and Gelman et al. (2014).

Slow convergence is a main issue in the application of MCMC methods to sample from posterior distributions, particularly when implementing Bayesian inference. Optimization of the proposal distribution in several initial runs does not often alleviate this problem (Reichert 2014). A number of adaptive MCMC methods are developed over the past few decades to avoid slow convergence, which are all based on continuously adapting the proposal distribution during simulation runs (Tierney & Mira 1999; Haario et al. 2001; Green & Mira 2001; Haario et al. 2006; Vihola 2012). In recent years, adaptive MCMC methods based on the use of several Markov chains run in parallel and that couple them adaptively have received a lot of attention (Ter Braak 2006; Vrugt et al. 2009; Laloy & Vrugt 2012). The procedure discussed in the next subsection is a particular implementation of these adaptive methods.

2.5.2 Differential evolution adaptive Metropolis method

Recognizing the limitations associated with existing MCMC methods, Vrugt et al. (2009) presented the Differential Evolution Adaptive Metropolis (DREAM) method. The DREAM technique is fundamentally based on the Differential Evolution-Markov Chain (DE-MC) method developed by Ter Braak (2006). The main extensions incorporated in DREAM to improve the search efficiency of DE-MC include: i) generate proposals by using higher order pairs to improve diversity in the searching process; ii) subspace sampling that helps to modify each dimension with a crossover probability; iii) detect an outlier chain during burn-in and remove them

using Inter Quartile Range; iv) estimate crossover probability distributions during burn-in and allow larger jumps, which helps to speed-up convergence. These extensions are observed to improve the overall performance of the MCMC algorithm, sometimes dramatically (Vrugt et al. 2009).

In DREAM, N different Markov chains $\{x_i, i = 1, \dots, N\}$ are run simultaneously in parallel. At initial stage a population of Nxd is formed, where *N* denotes number of chains and *d* is dimension of parameter space. The state of the i^{th} chain can be represented using a *d* -dimensional vector x_i *(i=1,…,N)* and the j^{th} element of x_i can be referred using x_i^j *(j=1,…,d)*. The procedures involved in generating parallel Markov chains using the DREAM algorithm are detailed below (Vrugt et al. 2009):

i) Initiate the parallel chains using a population $\{x_i, i = 1, \dots, N\}$ sampled from the prior distribution.

ii) Compute the density $\pi(x_i)$ for $i=1,\dots,N$, where $\pi(.)$ denotes probability distribution function of the target distribution.

For $i = 1,\dots,N$, do chain evolution

iii) Generate a candidate point, z_i in chain *i.*

$$z_i = x_i + (1 + e)\gamma(\delta, d_{eff})\left[\sum_{j=1}^{\delta} x_{r_1(j)} - \sum_{n=1}^{\delta} x_{r_2(n)}\right] + \varepsilon \qquad (2.15)$$

where δ signifies the number of pairs of chains used to generate a proposal, γ is jump rate and $r_{1(j)}$ & $r_{2(n)} \in \{1,\dots,N\}$; $r_{1(j)} \neq r_{2(n)} \neq i$. The value of γ depends on the number of pairs used to create the proposal. Random Walk Metropolis (RWM) guidelines suggest that, a good choice of $\gamma = 2.38/\sqrt{2\delta d_{eff}}$, where d_{eff} denotes the number of dimensions that will be updated. The value of *e* is drawn from *Ud(-b, b)* where $|b| < 1$ and white noise ε is drawn from $N_d(0,b^*)$ where b^* is small compared to the width of the target distribution.

iv) Replace each element of the proposal, z_i^j *(j=1,…,d)* with x_i^j using a binomial function using crossover (CR) probability of *1-CR*. When *CR=1* all dimensions are updated together *(d'=d)*.

$$z_i^j = \begin{cases} x_i^j & if\ U \leq 1 - CR, \quad d' = d' - 1 \\ z_i^j & otherwise \end{cases} \tag{2.16}$$

v) Compute $\pi(z_i)$ and $\alpha(x_i,z_i)$ of the candidate point and apply Metropolis selection rules to decide whether these proposals should be rejected or not, where $\alpha(x_i,z_i)$ denotes acceptance probability of the candidate point.

vi) If the candidate point is accepted $x_i=z_i$, otherwise remain at x_i.

End of chain revolution

vii) Use Inter Quartile Range (IQR) statistic, which is a measure of statistical dispersion that covers the range between the 75^{th} and 25^{th} percentiles, to remove outlier chains during burn-in. In MCMC procedures, burn-in is a practice of throwing away some samples at the beginning of MCMC run with a view to give Markov Chain time to reach its equilibrium distribution, which is particularly important if the Markov Chain started from a point far from the equilibrium distribution.

viii) Diagnose the convergence of the chains using Gelamn-Rubin, $\hat{R}_j$, for each dimension $j = 1, \dots, d$ using the last 50% of the sample in each chain. Gelman–Rubin is used to evaluate MCMC convergences by comparing the estimated between-chains and within-chain variances and analysing the difference between multiple Markov chains for each model parameter. Large differences between the variances indicate non-convergence.

ix) Stop if $\hat{R}_j < 1.2$ for all *j*, otherwise go to **chain evolution**.

Application of DREAM in different case studies, which involved a wide range of complex problems in terms of nonlinearity, dimensionality and multimodality, have showed that it generally outperforms existing MCMC methods (Vrugt et al. 2009; Laloy & Vrugt 2012; Shi et al. 2012; Mara et al. 2016a). Its efficiency in dealing

with complex problems, ergodicity and detailed balance make it a preferable tool to handle uncertainties in rainfall-runoff models.

2.6 Diffuse pollution

The aim of this section is to provide a review of the literature on the occurrence and monitoring of pesticides, existing catchment scale water quality models and uncertainty analysis associated with water quality models. The main focus is on the occurrence and modelling of diffuse pollutions in storm water runoff at catchment scale and thus the section is not intended to be a full review of pollutant fate and transport models at various scales.

2.6.1 Occurrence and monitoring of pesticides

Pesticides are widely used in modern agricultural practices with a view to increase production quantity and quality. As a result, water travelling through surface and groundwater pathways in agricultural catchment may be exposed to these pesticides and leaves the catchment mostly with its quality deteriorated. Most of this pollution is in the form of diffuse pollution from farmlands with insensitive agricultural practices (Bach et al. 2001), which contaminates surface and ground water bodies posing significant water quality issues for drinking water sources. Treatment of contaminated water to comply with drinking water standards imposes considerable capital and operating costs on water companies and contributes significantly to the industry's carbon footprint. In the European Union (EU), the maximum allowed concentration (MAC) in treated drinking water set out by the Drinking Water Directive (DWD; 98/83/EC) for an individual pesticide is 0.1µg/l (EC 1998). The DWD also states that total concentration of all pesticides should not exceed 0.5µg/l. EU member states are required to implement these standard limits as minimum requirements in their national legislations and can adopt even tighter water quality standards of choice. The value of 0.1µg/l MAC set by DWD is based on the typical limit of detection for most pesticides when the DWD was set out in 1980. Thus, the limits are not based on health risk considerations (i.e. combination of toxicological exposure and effects on human health) but represent a precautionary surrogate zero

to reflect the principle that drinking water should be free from pesticides. Consequently, the standard limit is considerably lower than the concentration level at which it is considered to cause any health impact. For example, the World Health Organization has set the concentration limit of 0.9mg/l for AMPA (Alpha-Amino-3-Hydroxy-5-Methyl-4-Isoxazole Propionic Acid) alone or in combination with glyphosate based on health impacts (Breach 2010). In addition to concerns associated with drinking water quality standards, high pesticide concentrations in water bodies can cause impacts to the water ecology (Warren et al. 2003). However, only concentrations significantly higher than 0.1µg/l are likely to cause any considerable eco-toxicological damage (Crommentuijn et al. 2000).

Observed levels of emerging pollutants in raw drinking water have increased as agricultural pesticide application rates have intensified, detection methods have improved and new products emerge onto the market (Geissen et al. 2015). The characteristic behaviour of these pollutants such as pesticides (e.g. metaldehyde and quinmerac) mean that existing drinking water treatment processes are inadequate to reduce levels to drinking water regulation limits and thus have become of rising concern to UK water companies (Kay & Grayson 2014). These pollutants can enter river systems via diffuse sources through a number of pathways including surface runoff, tile drains, drift, and groundwater flow. A number of studies have shown that rapid runoff in the form of overland and drain flows are the pathways considered to be the major non-point source of pollutants in most catchments (Huber et al. 1998; Dils & Heathwaite 1999; Huber et al. 2000; Bach et al. 2001; Harris & Catt 2006; Moore 2016). The proportion of diffuse pollutants transported through overland flow and drain flow varies in different catchments depending on meteorological conditions, soil type, land slope, agricultural practices and network of drains in the area (Bach et al. 2001). In catchments that consist network of drains, preferential water flow through macropores to tile drains plays an important role in the rapid transport of pesticides to water bodies (Tang et al. 2012). The importance of drain flows caused by a rainfall event water in transporting diffuse pollution to water courses is highlighted in a number of literatures (Harris & Catt 2006; Granger et al.

2010; Tiktak et al. 2012). The transport of pollutants in both surface runoff and drain flows, which form the dominant pathway for diffuse pollution transport into water bodies, is storm-driven and the initial storm after pesticide application produces higher flux of pollutants than subsequent storms (Stone & Wilson 2006). Thus, the combined event water contribution to water bodies transported via overland and drain flows can be referred as rapid runoff/ runoff. Catchment characteristics combined with hydro-meteorological parameters determine the amount and rate of runoff generated following a particular rainfall event. During rainfall events, the impact of raindrops and overland flow detach soil particles and cause the transport of pesticides both found in solution and adsorbed by sediment particles. Transport of highly soluble pesticides such as metaldehyde through runoff is considered far more important than transport via soil erosion (Holvoet et al. 2007). Migration of pollutants through erosion is considered significant only for highly adsorbing substances with a sorption coefficient of active ingredient to organic carbon (K_{OC}) value greater than 1000 L kg^{-1} (Kenaga 1980). The time interval between pesticide application and a rainfall event combined with application doses over this period determine the accumulation of pesticide in the active zone at the soil surface (Müller et al. 2003). This directly affects the amount of pesticide transported to water bodies through runoff.

UK water companies currently use grab sampling methods to monitor pesticide pollutant levels in catchments to help in the planning and development of catchment management strategies. The use of appropriate sampling methods in water quality monitoring studies determine the ability of collected data to reflect the actual conditions at the study sites and the validity of conclusions made based on the data (Ort et al. 2010). Sampling regimes can fail to provide good set of data which adequately represent catchment processes and can be a significant source of error misleading management decisions. In a research recently conducted in the Cherwell catchment, UK, the Metaldehyde Stewardship Group has indicated that a significantly large amount of metaldehyde losses occur during the rainfall event following metaldehyde applications, thereby highlighting the importance of

capturing the impacts of individual rainfall events (Kilburn 2010). Currently, grab samples are collected days or weeks apart by water companies and the pesticide concentrations in between sample collection times are unknown. Consequently, there is a concern regarding the suitability of water quality data collected through the existing grab sampling methods to inform the development of catchment management strategies. Moreover, model development activities to enable prediction of potential exposure to pesticide concentrations can be hampered by the inability to capture short term fluctuations in diffuse pollution concentrations caused by rainfall driven runoff. Recent studies have indicated the use of automatic water samplers as a step forward towards addressing this problem (Petersen et al. 2005; Rabiet et al. 2010). High resolution data generated by auto-samplers can be used to better inform the planning and development of catchment management strategies. Furthermore, these datasets also provide a potential to establish relationships between concentrations of pollutants and catchment hydro-meteorological parameters which can be measured more easily in real-time. Studies which exploit this potential to formulate catchment dynamics that relates catchment attributes with diffuse pollutant fluxes to water bodies are lacking. This is mainly due to lack of water quality datasets adequately describing the levels, fate and transport of such pollutants through runoff driven by rainfall events in the catchment. However, this is now becoming more accessible due to recent advances in the sampling and spatial data collection technologies.

2.6.2 Water quality modelling approaches

Water quality models that predict dynamic behaviour of pollutant sources, fate and transport have widely been used to inform water quality management decisions (Wang et al. 2013). A study of the literature reveals that considerable amount of work has been conducted on the development of water quality models that are capable of operating at different scales (Quilbé et al. 2006; Köhne et al. 2009). A choice of suitable water quality model for prediction of selected pollutants involves assessing the models based on the following criteria i) ability to accurately represent the fate and transport process of selected substances, ii) appropriateness to the

complexity of the catchment system and available data, iii) ability to incorporate methods to estimate prediction errors and report prediction uncertainty, iv) sufficient flexibility to allow updates and improvements based on new information and monitoring data (Loucks & Beek 2005). Additionally, practical suitability of the models for the intended specific application needs to be considered.

There is a wide variety of water quality models in the literature that represent processes involved in pollutant fluxes from land to water bodies and pollutant transport in channel networks at various levels of detail and structure. Quilbé et al. (2006) and Wang et al. (2013) have provided wide ranging reviews of these pollutant fate and transport models. Detailed water quality models such as MACRO (Larsbo & Jarvis 2003) have large data requirements and are known to be computationally intensive. As a result, these models are typically applied at small spatial scale and are more suitable to run with larger time-steps. In contrast, large scale models such as the European scenario models FOCUS 2000 and FOCUS 2001 do not provide the level of spatial details required to apply at catchment scale. Catchment scale water quality models, which are able to predict pollutant concentrations at outlets of catchments, are of particular interest for applications aimed at informing short term water quality management decisions such as surface water abstraction management. Catchment scale water quality models widely applied in the literature to predict diffuse pollution exposure in surface water bodies include the Soil and Water Assessment Tool - SWAT (Neitsch et al. 2002), Hydrological Simulation Program Fortran - HSPF (Donigian, A. S. et al. 1995) and AnnAGNPS (Bingner & Theurer 2005). Pullan et al. (2016) has also introduced a new integrated model for pesticide transport (IMPT) focused on operational application, which enabled to predict pesticide concentrations and hydrograph responses at different catchment scales. However, most of these models predict diffuse pollution exposure in surface water bodies and are developed with a view to analysing long-term effects of catchment management practices such as impacts of land use and agricultural activities on river water qualities (Nguyen et al. 2017) and thus predict pollutant concentrations with large time scales (daily, weekly, fortnightly, and monthly). As a result, these models are not considered to be

suitable for applications mainly used to enable operationally suitable short-term surface water abstraction management decisions. SWAT provides an option to run with small time-steps and a number of studies have applied it in a range of catchments to predict daily pesticide concentrations (e.g. Kannan et al. 2006; Holvoet et al. 2008; Luo & Zhang 2009). However, application of SWAT at catchment scale requires large amount of data and contains large number of parameters that need to be calibrated (Benaman et al. 2005).

Models capable of predicting pollutant loadings and transport from single rainfall events and run at small time-steps are of particular interest to model applications aimed at informing short-term water quality management (e.g. surface water abstraction management). The Dynamic Watershed Simulation Model - DWSM (Borah et al. 2002), the Agricultural Nonpoint-Source Pollution Model – the AGNPS (Young et al. 1989) and MIKE SHE (Refshaard et al. 1995) are some of the typical rainfall event based models available in the literature, which are widely used to simulate pollutant fate and transport throughout the catchment following rainfall events. A number of studies have shown that peak concentrations of pesticides from agricultural lands are observed following rainfall events, which emphasizes an important role runoff plays in generating and transporting pesticides from application areas to the outlet of the catchment (Borah et al. 2003; Sangchan et al. 2012; Bundschuh et al. 2014). Hence, rainfall event based occurrences of peak concentrations of some pesticides, which are of particular concern to water quality managers, are major factors when making short-term water quality management decisions. However, existing catchment scale models lack detailed representation of spatiotemporally distributed runoff based pesticide generation and transport processes throughout the catchment, which is required for accurate prediction of the arrival of short-term peak pesticide concentrations at catchment outlets following rainfall events. This is mainly due to the scarcity of high resolution model input and validation data that supports development of models that represent short-term dynamics in runoff based pesticide generation and transport throughout the catchment (Bach et al. 2001). The numerical structure of existing models combined

with their lack of spatially distributed representation of model inputs/parameters also limit the capability of analysing uncertainties associated with spatial variabilities of model inputs/parameters such as pollutant build-up and rainfall. However, detailed assessment of uncertainties in water quality models is required when using these models to inform water quality management decisions, particularly to inform decisions associated with short-term fluctuations of pollutant concentrations. In recent decades, technological advancements in data collection techniques such as satellites, rainfall radars and auto-samplers have enabled to collect good quality data with high temporal and spatial resolution. This provides a potential to develop water quality models capable of representing short-term dynamics in runoff based pesticide generation and transport and use these models to predict the arrival of peak pesticide concentrations at catchment outlets following rainfall events.

2.6.3 Uncertainty analysis in water quality modelling

Surface water quality models are developed with the ultimate purpose of making reliable predictions to assist in water quality management and decision making. However, there are a number of uncertainty sources in catchment scale water quality models and sometimes the model output uncertainty can be too high to draw meaningful conclusions (Lindblom et al. 2007; Vandenberghe et al. 2007). Uncertainties associated with water quality models or other hydrological models are typically introduced through the assumptions associated with: 1) the ability of the model to adequately represent the governing physical processes and states (model structural error), 2) the model parameters (calibration and identification error), 3) the input data (Vandenberghe et al. 1987; Beven 2001).

Most model formulations represent components of physical processes using parameters. The uses of parameters in models enable to define relations between system components in the model and help to clearly structure these relations. But, the values of some model parameters have to be estimated using calibration to make the model a good representation of the system (Beven 2001). The process of estimating these parameters should not only provide best estimates of model parameters but also

estimates of their uncertainty (Briggs et al. 2012). Similarly, uncertainties associated with model inputs obtained from measurements or estimated based on collected data and information needs to be analysed. Quantified uncertainties can then be propagated to water quality prediction results, which can be used by water quality managers to make well informed and risk-aware decisions (Hall & Borgomeo 2013).

Environmental systems, which exhibit variable natural processes, are often poorly monitored due to lack of resources and the expensive procedures involved in collecting and analysing datasets, which also often require careful handling and analysis in laboratories (McIntyre et al. 2002). For example, collection of hourly metaldehyde data that can be used to capture short term fluctuations of metaldehyde concentration in rivers requires implementation and continuous operation of automatic samplers during runoff periods and laboratory analysis of sampled water, which is quite an expensive and resource intensive process. As a result, high resolution data to support development of surface water quality model that are capable of predicting short term fluctuations of pollutant concentrations in rivers are generally scarce (Bach et al. 2001). In addition, a significant proportion of these data are prone to high noise levels and bias due to the various procedures involved in collecting, handling and measurement of samples (Keith 1990). The resulting lack of water quality data thus limits the ability to adequately formulate model structures, accurately assign model inputs and precisely identify model parameters often causing large uncertainty in water quality model predictions (Hankin et al. 2016). The consequences of water quality management decisions based on these model results thus become much more difficult to predict (Briggs et al. 2012). In decision making associated with water quality problems, understanding average system responses from deterministic models are often not sufficient and uncertainties incorporated in models are becoming increasingly important (O'Hagan 2012). Water quality management decisions aim to keep the right balance between socio-economic development and socio-environmental protection, which often involves a trade-off between maximizing the human benefit while minimizing the risk of harm caused to the environment and society (Dietz 2003; Reichert & Borsuk 2005). Evaluating the

nature and extent of model output uncertainties can help to explicitly define these risks associated with management decisions and provides a well-informed picture of the possible outcomes, hence enabling to make risk-aware decisions (Burgman 2005). Thus, information on uncertainties of water quality model predictions that enable probabilistic modelling and decision analysis are indispensable for water quality management and decision making.

Several studies have analysed the different sources of uncertainties in water quality models mainly focusing on quantifying uncertainties associated with the values of model inputs and parameters (Reckhow 1994; Sangchan et al. 2012; Rangel-Peraza et al. 2016). A number of studies have also focused on developing new methods of handling model uncertainties. These methods include first order error analysis, which is based on first order terms in the Taylor series expansion of variable dependency relations (X Zhao et al. 2011), Monte Carlo simulation, which draws large samples from distributions of uncertain variables and use them for stochastic simulations to establish model uncertainty (Tomassini et al. 2007), MCMC, which is a special type of random process discussed in detailed in section 2.5.1, generalized likelihood uncertainty estimation (GLUE), which assess global uncertainty based on Monte Carlo and likelihood measures (Beven & Binley 1992). A number of factors need to be considered when choosing an uncertainty analysis tool as discussed in section 2.5. Monte Carlo simulation based on selection of several representative samples of uncertain variables is a preferred method to provide a better estimate of uncertainties in water quality models that have high computational demand and when implementation of other methods is not feasible (Reichert 2014). Most of uncertainty analysis approaches in the literature focus on quantifying uncertainties associated with the value of model inputs, parameters and structures. Studies that assess uncertainties due to spatial aggregation of model input and parameters in catchment scale water quality models are lacking in the literature. This is mainly due to the scarcity of catchment scale high spatiotemporally distributed data and lack of suitable modelling structure in catchment scale models. The absence of structurally suitable physically distributed catchment scale water quality models, which are

required for implementation of spatial uncertainty analysis techniques, has also restricted the progress of spatial uncertainty analysis studies of catchment scale water quality models. These analyses also require computationally efficient spatial uncertainty analysis tools that can be used to deal with uncertainties associated with spatially variable model inputs and parameters. Integration of structurally suitable physically distributed model with a spatial uncertainty analysis tool can help to quantify the significant levels of uncertainties in catchment scale runoff based models, which are caused by representing spatially variable model inputs and parameters using spatially aggregated values (Andréassian et al. 2004). Particularly, water quality models that aim to predict short-term fluctuations in pesticide concentrations following rainfall events need to consider uncertainties associated with parameters that represent pesticide application locations in the catchment.

2.7 Identified knowledge gaps

Following review of the literature, scientific knowledge gaps have been identified in the area of real-time abstraction management and its use to address increasing pressures on the environment. The use of real-time flow prediction models to inform surface water abstraction decisions can help to effectively utilize available water resources and adapt to water scarcity challenges. In a wide range of existing flow forecasting studies, models have been developed in isolation to serve the needs of flood warning and protection systems focused on prediction of exceedance of certain flow thresholds. However, the use of real-time modelling methods for surface water abstraction management purpose needs to focus on service flow ranges, which is the variety of flow conditions ranging between the extreme dry and flood conditions. Effective application of flow forecast models in a real-time surface water management scheme thus raises specific modelling requirements and provides unique challenges. Firstly, the flow forecast model needs to accurately forecast a variety of flow conditions and provide uncertainties associated with forecasted flow values to avoid the risk of breaching abstraction licence conditions due to model errors and thus enable making risk-aware surface water abstraction decisions. Secondly, a water resource management model needs to be developed and coupled

with the flow forecast model to enable to inform surface water abstraction management decisions on the daily volume of water that can be abstracted sustainably. The coupling of water resources management model and a real-time rainfall-runoff model enables to combine forecasted flow data with reservoir operation rules, abstraction licence conditions, available storage volumes and pump and water main capacities to determine the amount of water that can be abstracted at any specific period in real-time. A study that utilises real-time modelling capabilities combined with these modelling requirements to enable dynamic surface water abstraction management is lacking.

To enable operationally suitable real-time abstraction management scheme, the probabilistic forecast model enabled through the uncertainty analysis method needs to be integrated with water resources management model that incorporates various operating rules to represent real-world operational constraints such as reservoir control curves, abstraction licence conditions and minimum flow requirements. Reservoir operational rules are particularly required to be encompassed in water resource models to determine the amount of water available for use to meet demands. These rules define the percentages of reservoir capacity that are reserved for flood control, drought monitoring and dead storage buffer zones (Ajami et al. 2008). Control curves are subjectively developed by water resource planners to enable effective management of impounding reservoir storages and are used to identify these zones. The drought monitoring curves define the minimum storage level required to be maintained in the reservoir at the beginning of each month in order to ensure continuous and reliable supply of water is provided to meet full demand (Ajami et al. 2008). Studies that integrate these kinds of water resources management models with suitable probabilistic flow forecast techniques to develop advanced surface water abstraction management scheme are currently lacking.

In recent years, the presence of emerging pollutants such as metaldehyde in raw and treated drinking water has become a concern for the water industry. Several studies in the literature have shown the occurrences of peak pesticide concentrations following rainfall events and revealed the significant role runoff plays in transporting

pesticides. However, appropriate data which can be used to sufficiently describe pollutant presence in the aquatic environment is notably lacking. Non representative water quality data, which do not reflect dramatic short duration variations in pollutant concentrations, can lead to mis-interpretation of river water quality status and may result in undermining associated risk to the environment and public water supplies. Most existing water quality datasets come from sampling campaigns set out to collect samples with frequency and spatial density suitable for regulatory and long-term catchment management purposes (Berthouex & Brown 2002; Ward et al. 1986). These datasets fail to capture dynamic responses of catchments as required for successful development of catchment scale water quality models that are capable of predicting pollutant concentrations with high resolution and are of high importance to inform short-term water quality management decisions. Consequently, most existing catchment scale models do not sufficiently represent the short-term dynamics involved in runoff based pesticide generation and transport processes, which is mainly responsible for the occurrences of short-term peak pesticide concentrations at catchment outlets following rainfall events. Furthermore, the numerical structure and the processes represented by existing rainfall event based models such as MIKE SHE make these models computationally expensive and their structure does not allow for implementation of suitable uncertainty analysis methods. Concentrations of pollutants such as metaldehyde, which are mainly transported through runoff following rainfall events, at catchment outlets are heavily dependent on the rate of runoff generation from high risk areas throughout the catchment. Thus, accurate representation of spatial variability of rainfall is essential when using catchment scale models to predict concentrations of pollutants at catchment outlets following rainfall events. Furthermore, a suitable modelling structure that combines identified high risk areas in a catchment with spatially variable rainfall data is required to accurately represent spatiotemporally variable generation and transport of pollutants across the catchment. The increasing usages of automatic samplers in recent years combined with high resolution catchment data capturing technologies provide a potential to improve the descriptions of short-term pollutant dynamics in models based on high spatiotemporal representation of pollutant loadings and

transport throughout the catchment. However, a study that exploits this potential to develop runoff based water quality model that enables to predict the arrival of short-term peak pesticide concentrations at catchment outlets following rainfall events with enough level of complexity and practical model structure is missing. With the growing concern associated with emerging pollutants such as metaldehyde in recent years, the use of these models is becoming increasingly important to inform short-term water quality management decisions.

Spatially and temporally distributed catchment scale model inputs and model parameters such as the amount of metaldehyde applied on high risk areas are often difficult to acquire. So, it is often a challenge to accurately represent these spatially distributed data in catchment scale models. Catchment average data are commonly estimated and used in models, which may cause significant uncertainties. Spatially distributed catchment scale models provide the potential to assess and quantify these uncertainties and compare them with uncertainties associated with inputs and parameters values. However, a study on the assessment and quantification of uncertainties in water quality models caused due to spatial mis-distribution of model input and parameters in catchment scale water quality models is lacking in the literature. Uncertainty analysis studies in water quality models in the literature have mainly focused on quantifying uncertainties in model input and parameters values and development of various methods of dealing with these uncertainties as discussed in section 2.5. Quantifying uncertainties associated with spatial distribution of water quality model inputs and parameters such as pesticide applications on farmlands throughout the catchment is particularly important when using models to forecast the arrival of peak pesticide concentrations in runoff following rainfall events. Comparison of uncertainty contributions from different sources can also help to improve future model development and data collection activities.

Chapter 3

Flow Prediction to Inform Surface Water Abstraction

3.1 Introduction

Resilience in public water supply has improved since privatization. However, population growth and the impacts of climate change such as frequent occurrences of drought are likely to stretch this capacity. Increased periods of low flow within surface waters mean that abstraction restriction such as 'hands off' flow conditions could be activated more frequently, increasing the pressure on water supply-demand management systems. A recent study on predicted availability of water and demand growth in the Severn Trent Water (STW) region indicated that if nothing is done to address the problem, there will be significant deficits of water by 2040 due to changes to abstraction licensing and climate change impacts on water resources (Water Forum 2015). Thus, there is a need to consider a sustainable solution to tackle the problems associated with keeping the increasingly delicate balance between public water use and the environment.

Surface water abstraction licensing system in the UK was initially designed in the early 1960s when there was believed to be surplus of water in most areas. However, our understanding of the impact of abstraction on the environment has grown over time and the perceived surplus water has significantly decreased. Consequently, the licensing system has gone through updates mainly during the Water Resources Act 1991 and the Water Act 2003, particularly to better protect the environment (OFWAT & EA 2011). However, the basic principles which form the licensing system remain unchanged. The current licences allow abstractors to take fixed volumes of water and occasionally include further conditions to restrict abstractions with a view to protect the aquatic environment. These conditions restrict daily abstraction volumes based on specified daily flow values at abstraction sites, which are often known as 'hands off' flow conditions referring to a fixed flow in the river below which abstraction is proscribed. Thus, during low flow periods, the amount of available water for abstraction directly varies with river flow. However, current

abstraction decisions are not informed on the amount of predicted daily flows in rivers and usually a conservative approach is taken when making daily abstraction decisions to avoid the risk of breaching licence conditions. This results in missing a certain amount of water daily, which could be abstracted without breaching the licence conditions and potentially help avoid the need to trigger drought management actions during dry periods. In particular, during dry periods reservoir levels drop below target storage alert curves and thus as much water as possible is needed to be pumped into reservoirs. Initial assessment of historical and simulated abstraction data has shown that a significant amount of opportunities to sustainably abstract more water without breaching abstraction licence conditions have been missed at abstraction sites. Enabling to abstract extra amount of available water can help raise reservoir storage levels and particularly in dry periods it can help to avoid the need to trigger drought management actions. In addition to enabling to minimize future investments (estimated as £1M per Ml/d), increase resilience, and avoid drought actions, this can play an important role in minimizing impacts on the environment. Efficient use of available water resources to meet increasing demand, whilst maintaining the quality of the aquatic environment has become increasingly important (section 2.1). This has raised the need to develop improved techniques to assist in the sustainable use and management of water resources. The increasing presence of hydrological and meteorological data from advanced radar and sensor technologies combined with high computational capabilities provide a potential to use real-time monitoring and modelling to tackle these problems by better informing abstraction decision making. A RTAM scheme that enables surface water abstraction decisions to account for water availability in the source in real-time by employing real-time flow forecasting models combined with water resources management model is proposed in this study. This enables to optimise the use of available surface water sources and provides a sustainable solution to tackle water resources problems associated with emerging issues of increasing demand, impacts of climate change and associated policy reforms.

In real-time modelling, models are applied to make predictions about future system states and outputs based not only on historical data but also using newly received and/or forecasted data. Knowledge on future state systems and outputs are valuable for pro-active management and planning. Thus, real-time modelling is employed in a wide range of geo-science and environmental fields including hydrology (Seo et al. 2003; Cloke & Pappenberger 2009) and meteorology (Golding 2000; Thorndahl et al. 2010). Typical application of real-time modelling in these areas include providing warning systems of future events such as flooding (Penning-Rowsell et al. 2000). Amidst an intense period of real-time model development and application, several studies have focused on employing real-time modelling in urban drainage system and distribution networks. In addition, most of the studies on real-time rainfall-runoff modelling have examined its application for flood forecasting. However, less attention is given on applying real-time models for managing human interventions to the river systems, which is partially due to old perception on the availability of water and limited understanding of the changing water environment and impacts of our interventions (e.g. abstractions) on the environment. In recent years, there has been a significant improvement in our understanding of the needs of ecosystems in water and an increase in realization of the combined effects of climate change and a growing population. Real-time modelling provides a potential to inform water management decisions with a view to minimize impacts of interventions on the environment and optimise the use of water resources infrastructures.

The current study investigates the use of integrated hydrological forecasting with water resources management model in order to maximize the amount of water abstracted in a trial catchment. A model has been developed to enable probabilistic prediction of flows over a prediction period of 24 hours, which can be used to inform daily abstraction decisions. A Bayesian based uncertainty analysis method is implemented in the model to assess hydrological uncertainties and enable probabilistic flow predictions, which will allow risk-aware decision makings and minimize the risk of breaching licence conditions. The RTAM scheme thus aims to increase the resilience of the water supply system in the catchment and minimize

impacts on the aquatic environment by dynamically linking abstraction volumes to actual availability of water in the source. The developed approach provides a novel technique to develop and integrate probabilistic flow forecast models and water resources management models that incorporate various operating rules to represent real-world operational constraints.

3.2 Methodology

This section describes the calibration and validation of the PDM rainfall-runoff model using a Bayesian inference method and the integration of the calibrated model with a water resource management model to forecast flows with a view to inform surface water abstraction management is described. First, the study catchment and the conceptual rainfall-runoff model, PDM, used in this study are presented. The second part of this section details the Bayesian inference method used to calibrate the PDM rainfall-runoff model and techniques used to implement the method. Model input, calibration and validation data as well as model parameterization and model performance evaluation criteria are also detailed in this section. Data assimilation techniques employed to use the PDM in forecast mode for real-time flow forecast are also described in this section. Furthermore, this section details about the development and application of a water resources management model that is used to assess the implications of employing the real-time abstraction management scheme on reservoir levels and associated water resources management decisions in the study catchment.

3.2.1 Study area

The Dove catchment used in this study is a sub catchment of the Trent catchment located in the UK Midlands. The Dove catchment drains an area of approximately 1,020km^2 and includes Churnet, Tean, Manifold and Hamps sub-catchments. The elevations in the catchment range between 550m to 50m above sea level from its source to its confluence to the River Trent. The Dove River is 45 miles (72 km) in length flowing generally south to its confluence with the River Trent and is the major river of the southwestern Peak District, in the Midlands of England. In the downstream part of the catchment, the River Dove flows through a wide floodplain

which contains extensive flood embankments constructed to protect villages and farmland in the area.

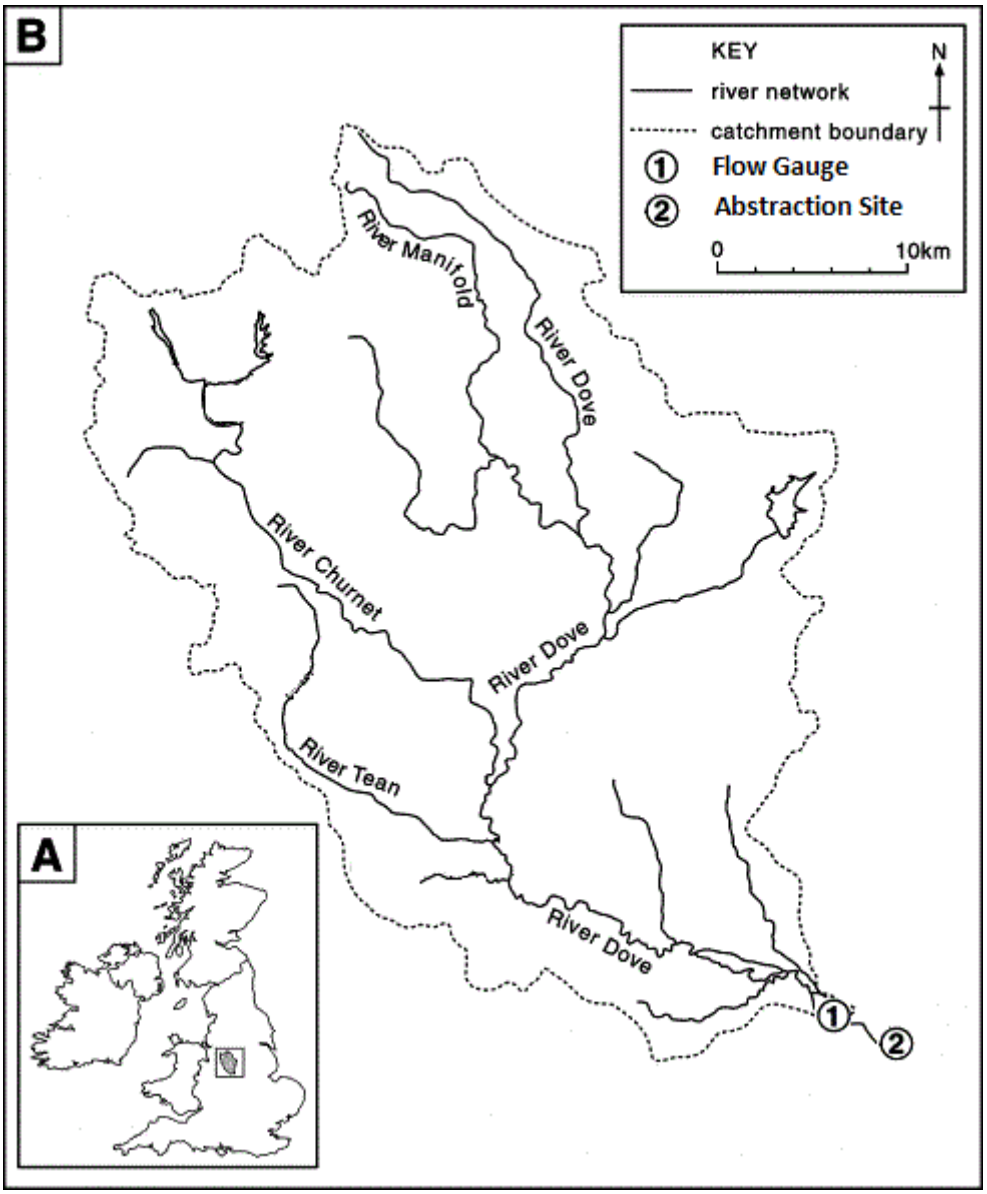


Figure 3.1. (A) Location of the River Dove catchment. (B) The River Dove catchment and location of flow gauge and abstraction site. (Goodson et al. 2002)

The normal flow depth of the River Dove at the gauging station located at the outlet of the catchment is between 0.43m and 0.83m. In extreme weather conditions the water level rises and ranges between 0.49m and 1.75m. A UK Environment Agencies' flow gauging station is situated at the outlet of the catchment to monitor

daily volumes of water abstracted at the public water supply abstraction site downstream (Figure 3.1). The abstraction site is used by a water utility operator to pump water to impounding reservoirs for water supply purposes. Data from the flow gauging station (at 15 min temporal resolution) have been used to calibrate and validate the hydrological model used in this study. The data used have been collected continuously over the period 2004 - 2013.

3.2.2 Description of flow prediction model

In this study, we used a conceptual rainfall-runoff probability distributed model (PDM), which has widespread application throughout the world, both for operational and design purposes (Young & Reynard 2004; Cabus 2008). PDM describes runoff production from a catchment mainly as a function of rainfall, evaporation and absorption capacity of soil columns, which are represented as a succession of soil moisture storages. It is widely recognized that the soil moisture storage capacity widely varies throughout the catchment and this variation is represented in PDM using a probability distributed function. This is based on the fact that on discrete basis, there are more soil moisture storages of one capacity than another and the actual runoff produced from a catchment can be obtained by weighting runoff produced by a store, which is a soil moisture storage column, of a given capacity by its frequency of occurrence. The model provides various modelling options including different probability density functions to enable the representation of a range of catchment responses in different ways. A series of publications (Moore 1985; Moore & Bell 2002; Moore 2007) provided details of the theoretical background of the model.The PDM model initially expresses runoff produced by a store of a given capacity mathematically by using the following equation:

$$q' = \begin{cases} P - E - (c' - S_o) & P > c' + E \\ 0 & P \leq c' + E \end{cases} \quad (3.1)$$

where S_o is the initial depth of water in storage, and *P, E, c'* and *q'* represent the depth of rainfall, potential evaporation, storage capacity and direct runoff produced

over the interval being considered respectively. If *c* is a storage capacity at any point in the catchment, then it can be described using a random variant *f(c)* so that the proportion of the catchment storages with capacity between *c* and *dc* will be *f(c)dc*. At the end of a unit duration at which a rain falls at a net rate P-E, shallowest stores in the catchment with available capacity (C' – So) less than net rate P - E start to produce runoff (Figure 3.2). Let *C**denotes this critical store capacity below which all stores produce runoff. Thus, proportion of the catchment containing stores of capacity less than or equal to *C** is

$$\mathrm{Prob}(c \leq C^*) = F(C^*) = \int_0^{C^*} f(c)dc \tag{3.2}$$

This is also the proportion of catchment area generating runoff (runoff contributing area). Thus, the value of F(C*) (cumulative distribution function of store capacities evaluated at C*) in equation (3.2) above varies between 0 and 1 depending on the proportion of store capacities that are above/ below C*. Consequently, at any particular time the instantaneous direct runoff per unit area *q(t)* from a catchment generated by net rainfall *π(t)* can be calculated from

$$q(t) = \pi(t)F(C^*(t)) \tag{3.3}$$

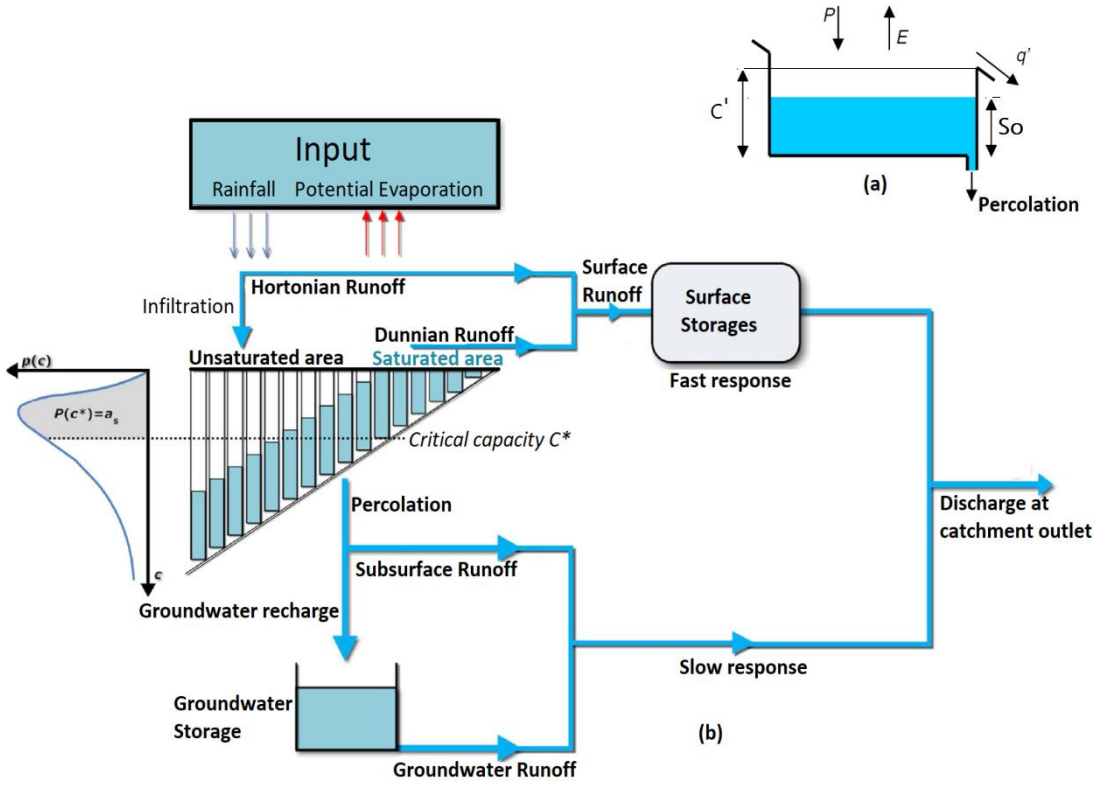


Figure 3.2. Runoff production from a population of stores in the PDM model and the model interacting components (a) Point representation of runoff production by a single store (b) Catchment representation by storage elements of different depth and their representation using probability density function (Noto 2014).

Runoff generation during rainfall events occur as a result of two main processes called Hortonian and Dunnian runoff processes. Hortonian Runoff also known as infiltration excess overland flow, occurs when the rate of rainfall is faster than the rate at which the soil can absorb the water. Saturation excess runoff also often known as Dunnian runoff, occurs when the soil is fully saturated and all the rainfall during this period causes runoff. During rainfall events these runoff generation processes control how much of the rainfall water gets into water bodies and flows downstream to the catchment outlet. During dry periods actual evaporation depletes the water content of soil moisture storage. In PDM, the dependence of actual evaporation loss on soil moisture content is introduced by assuming the following simple function between actual and potential evaporation ratio, catchment level soil moisture deficit

$\left(S_{max} - S(t)\right)$, where S_{max} represents maximum catchment soil moisture storage capacity and $S(t)$ represents catchment soil moisture content at the start of the interval.

$$\frac{E_i'}{E_i} = 1 - \left\{\frac{\left(S_{max} - S(t)\right)}{S_{max}}\right\}^{b_e} \tag{3.4}$$

Usually a linear ($b_e = 1$), thus $E_i' = (S(t)/S_{max})E_i$ or a quadratic form ($b_e = 2$) is used.

Further loss to ground water is introduced by assuming the rate of percolation over an interval (d_i) varies depending on catchment soil moisture content at the strat of the interval as shown below.

$$d_i = \frac{(S(t) - S_t)^{b_g}}{k_g} \tag{3.5}$$

where k_g is drainage time constant and b_g is recharge function exponent and is usually set to 1, S_t is threshold storage below which water gets held by soil tension and no drainage occurs.

Ground water recharge from soil water drainage is routed through subsurface storage that represents subsurface flow paths often referred as slow response system. This system is defined in PDM using a variety of non-linear storage reservoirs. The non-linear storage model is defined by the Horton-Izzard equation (Moore 2007).

$$\frac{dq}{dt} = \mathrm{a}(u - q)q^b, \qquad q > 0, -\infty < b < 1 \tag{3.6}$$

where u is the rate of inflow and q is the rate of out flow per unit area, a and b are parameters that are given by $a = mk^{1/m}$ and $b = (m-1)/m$ where k is the storage

rate coefficient and m is the store exponent. PDM provides recursive solutions of the Horton-Izzard equation for a choice of various non-linear storage forms such as linear, quadratic, cubic, exponential and general non-linear. A cubic for is often considered the most appropriate to represent ground water storage (Moore 2007). When the Horton-Izzard equation is used to represent ground water storage in PDM, the rate of percolation (d_i) will be used as input u and the output q will be the baseflow component of total flow.

In recent years, PDM has evolved as a fairly general conceptual rainfall-runoff model well suited for various flow prediction applications with relatively few amount of parameters. Currently, PDM has widespread application throughout the world (Teuling et al. 2004; McIntyre et al. 2005; Cabus 2008) and has earned a solid reputation as a flow forecasting model in the UK (Bell et al. 2001). PDM's suitability to represent a wide range of flow characteristics, widespread application in the scientific community, suitability for industrial operational application and well-structured model concepts are some of the main factors considered when choosing the modelling tool (Cabus 2008; Bell et al. 2001).

3.2.3 Uncertainty analysis

The use of hydrological models in water resources management introduces parameter uncertainties regardless of the complexity and structure of models used. These uncertainties in the model parameters are likely to impact water resources planning decisions. Calibration or inverse problems use observed system response data to estimate unknown quantities incorporated in mathematical models. In the case of the PDM these unknown quantities are hydrological properties of the catchment (e.g. soil moisture capacities, surface and ground water storages and routing properties), which are represented in the model using parameters (Moore & Clarke 1981; Moore 2007). In non-linear calibration, the Bayesian approach serves the purpose of transferring information from collected data on system responses to the unknown variables, thus updating the posterior probability distributions that describe uncertainty about the unknown variables (Feyen et al. 2007). When more observations become available,

this posterior distribution can be used as a prior to be updated again, which then leads to a logically consistent representation of a sequential learning process. There are only a small number of probability distributions and models for which Bayesian inference can be done analytically. For this reason, numerical techniques for Bayesian inference are very important. The most important of these techniques are based on Monte Carlo simulation. This means that a sample is drawn from the posterior distribution and properties of the distribution are approximated by properties of the sample. The numerical implementation of Bayesian techniques usually requires computationally demanding Monte Carlo techniques. Of particular importance are Markov Chain Monte Carlo (MCMC) techniques that construct a Markov chain which asymptotically covers the posterior distribution. Application of such techniques requires tuning algorithmic parameters to improve the convergence rate of the chain and careful convergence diagnosis (Smith & Marshall 2008). Statistical emulators that interpolate the solutions of computationally demanding simulation models can be useful to shorten computation time.

A state-of-the-art Markov chain Monte Carlo (MCMC) technique entitled Differential Evolution Adaptive Metropolis (DREAM) is used in this chapter to estimate parameter uncertainties in the PDM model (Vrugt et al. 2009). The Differential Evolution Markov Chain (DE-MC) technique initially developed by Ter Braak (2006) forms the building block of DREAM. DREAM is a multi-chain MCMC technique and has excellent performance in sampling complex, multi-modal and high dimensional target distributions. In recent years, DREAM has found widespread application and use for estimation of optimal parameter values and their underlying posterior probability density function on a wide range of model calibration and uncertainty analysis studies. DREAM is basically an adaptation of the Shuffled Complex Evolution Metropolis algorithm (Vrugt 2003). In DREAM, a number of predefined chains are run in parallel to search the parameter space. Latin hypercube or covariance-based sampling methods are used to sample from a prior parameter space with uniform distribution to initialize a specified number of Markov Chains. These parallel chains at the initial stage form an $N \; x \; d$ matrix, where N denotes number of chains and d dimension of parameter space. At each stage,

differential evolution as genetic algorithm creates multivariate proposals to evolve the chains and Metropolis selection rules are applied to decide whether these proposals should be rejected or not (equation (3.7)).

$$z_i = x_i + (1+e)\gamma(\delta, d_{eff})\left[\sum_{j=1}^{\delta} x_{r_{1(j)}} - \sum_{n=1}^{\delta} x_{r_{2(n)}}\right] + \varepsilon \quad (3.7)$$

where δ signifies the number of pairs of chains used to generate a proposal, ϒ is jump rate and $r_{1(j)}$ & $r_{2(n)} \in \{1,\ldots,N\}$; $r_{1(j)} \neq r_{2(n)} \neq i$. The value of ϒ depends on the number of pairs used to create the proposal. Random Walk Metropolis (RWM) guidelines suggest that, a good choice of $\gamma = 2.38/\sqrt{(2\delta deff)}$, where *deff* denotes the number of dimensions that will be updated. The value of *e* is drawn from $U_d(-b, b)$ where $|b| < 1$ and white noise ε is drawn from $N_d(0, b^*)$ where b^* is small compared to the width of the target distribution. The chain moves from X_i to Z_i depending on the Metropolis selection criteria and forms a Markov chain, whose stationary distribution is the posterior distribution of the parameters. Ter Braak & Vrugt (2008) have shown that this Markov Chain converges to the posterior distribution. R statistics of Gelman & Rubin (1992) can be used to monitor convergence of the Markov chain after the Differential Evolution – Markov Chain (DE-MC) becomes independent of its initial values (after burning Period). The generated posterior population can be used to communicate uncertainties in model parameters and model predictions. Vrugt et al. (2009) have provided a detailed description of DREAM.

3.2.4 Model input, calibration and verification data

Rainfall is considered the main driving force in the catchment and river systems. Good information on rainfall intensity is of great importance for hydrological modelling. In recent years, advances in data processing, communications and display technology have enabled the use of meteorological radars to produce rainfall data, which provides better representation of spatial variability of rainfall across catchments (Rabiei & Haberlandt 2015). In this study, composite radar rainfall data with spatial and temporal resolution of 1km^2 and 5 minutes respectively were

acquired from the UK Met-Office and used for the calibration and validation of the flow forecast model (Met Office 2003). A computer program has been developed to process the 1km^2 composite radar rainfall data from the Met-Office website and compute catchment averaged time series rainfall data for the study catchment. Missing values in radar rainfall data are mainly caused by the absence of radar image at a specified time for all of the UK. Thus, infilling methods which use data from nearby locations are not feasible to infill the missing data as these data are not available for all catchments covered by the radar rainfall range. However, the original data is provided as 5 minutes interval data, where consecutive values are highly related to each other. The moving average method was found to be the most suitable technique for infilling the missing values in the rainfall data. Six hours moving average rainfall data have been applied to infill the missing data in the processed time series rainfall data. Daily potential evaporation data are acquired from the UK Met-Office's MORECS system and also used as an input to the PDM model (Hough & Jones 1997). Historical flow data with time intervals of 15 minutes collected over the period of 2004 to 2013, from a flow gauging station situated at the outlet of the catchment and upstream of the surface water abstraction site, were obtained from the UK Environment Agency and are used for calibration and validation of the flow forecast model.

3.2.5 Model Parametrization

Most parameters in general conceptual models such as PDM lack physical basis and mostly cannot be inferred from direct measurements. To reduce the dimensionality of the model calibration in this study some of the PDM parameters are estimated from initial assessments and databases. Prior to calibration, the model was run with default parameter values and water balance outputs are examined to find initial estimates of parameters which control runoff volumes (rainfall factor f and exponent in actual evaporation function b_e). Common and catchment specific values are assigned for some parameters such as a parameter that controls the type and distribution of spatial variability of store capacity (b), soil tension storage capacity (S_t) and constant flow parameter to represent returns or abstractions (q_c), which are all believed to be less

significant in changing runoff outputs in the study catchment. Eight parameters were identified in this study for calibration using measured flow data (Table 3.1). Bayesian based uncertainty analysis methods including DREAM have been used in the literature to solve inverse problems involving similar scale of calibration parameters (Feyen et al. 2007; Bilondi et al. 2013; Muleta et al. 2013; Mara et al. 2016b). Uniformly distributed priors for the identified calibration parameters are provided with upper and lower bounds for use in DREAM (Table 3.1). The PDM model guidelines and previous modelling works are used to define prior parameter ranges to make sure the parameter values remain hydrologically realistic. Minimum depth (C_{min}) and maximum depth (C_{max}) parameters are used in PDM to define the ranges across which storage depth is distributed with a particular distribution. Pareto distribution, which is a skewed, heavy tailed power law probability distribution, is the most commonly used distribution (Moore & Bell 2002). The cumulative distribution and the probability density function for the pareto distribution used in PDM are:

$$\mathrm{F}(c) = 1 - \left(1 - \frac{c}{c_{max}}\right)^{b}, \qquad 0 \le c \le c_{max} \tag{3.8}$$

$$\mathrm{f}(c) = \frac{dF(c)}{dc} = \frac{b}{c_{max}}\left(1 - \frac{c}{c_{max}}\right)^{b-1}, \qquad 0 \le c \le c_{max} \tag{3.9}$$

The pareto distribution can take different forms based on parameter *b*. More deep stores than shallow stores exist when *b* is small (between 0 and 1) whereas greater frequency of shallow stores occurs when *b* is greater than one. Parameters k_1 and k_2 are time constants, which are used to determine the rate of flow from two linear stores that constitute PDM's surface storage and they control the level of peaks in hydrographs. Ground water recharge time constant (k_g) along with soil moisture storage contents control rate of aquifer recharge in a nonlinear function. Base flow time constant $k_{b=}\ k^{-1}$ controls rate of flow from subsurface storage and thus determines length of recession in a hydrograph. Exponent of recharge function (b_g) represents the sensitivity of recharge rate to soil dryness. Horizontal shift of

hydrograph along time axis is adjusted in PDM using a time delay parameter (T_{dly}). Elements of the inverse problem are passed to the DREAM algorithm using the following equation

$$\begin{aligned} &\text{Dream\_Object} \\ &= \text{dreamCalibrate}(\text{FUN}, \text{FUN.type}, \text{pars}, \text{measurement}, \text{control}) \end{aligned} \tag{3.10}$$

where FUN is model function with argument of a vector of parameter values, FUN.type is type of value FUN returns, pars is a list of variable ranges, measurement is required parameter for FUN.type (i.e. observed system response data), control is list of settings for the DREAM algorithm. In this study FUN = call_PDM is used where call_PDM is a function set up to run the PDM model with sampled parameters from the DREAM algorithm and returns the residual vector to be compared with measurement. The sum of squared residuals option from the DREAM algorithm was used as FUN.type to calculate the likelihood. A list of parameter ranges was passed as *pars = list("cmin"=c(0,50),"cmax"=c(100,400),"k1"=c(1,70),"k2"=c(1,70), "kb"=c(1,400),"kg"=c(60000,120000),"bg"=c(1.4,1.8),"tdly"=c(0,0.5))*. Control = *list(nseq=8,ndraw=2500)* is used where nseq = number of of parallel chains to evolve in DREAM and ndraw is the number of function evaluations.

Table 3.1. Prior ranges and description of the PDM parameters used in DREAM

Parameter	**Description**	**Minimum**	**Maximum**
C_{min} **(mm)**	Minimum store capacity	0	50
C_{max} **(mm)**	Maximum store capacity	100	400
k_1 **(h)**	Time constant of surface storage 1	1	70
k_2 **(h)**	Time constant of surface storage 2	1	70
K_b **($h^{1/3}$ $mm^{2/3}$)**	Baseflow time constant	1	400
K_g **(h $mmbg^{-1}$)**	Groundwater recharge time constant	60000	120000
b_g	Exponent of recharge function	1.4	1.8
T_{dly} **(h)**	Time delay	0	0.5

3.2.6 Model evaluation criteria

The performance of the PDM rainfall-runoff model in the stochastic model calibration and validation method employed in this chapter is evaluated using P-factor and R-factor values. P-factor represents the percentage of observations that fall within a given prediction uncertainty bounds, whereas R-factor is used to represent the average width of prediction uncertainty bounds divided by the standard deviation of the observations. These statistical parameters are commonly used to quantify the fit between stochastic simulation outputs, which are commonly expressed using 95 % prediction uncertainty bounds, and observation data expressed in single signals. A P-factor value of 1 and R-factor value of 0 indicate an absolute match between simulated and observed values with model structure and parameter values perfectly representing modelled systems in the catchment under study. Achieving an optimum value for both factors requires keeping the balance between capturing most of observations within a specified uncertainty bound and keeping the width of prediction uncertainty bound as small as possible. In addition, the deterministic model's flow prediction efficiency using optimum parameter values, which are parameter values that are found to give the best fit between simulated and observed flows over the entire hydrograph, was evaluated using the Nash-Sutcliffe model efficiency coefficient (*NS*) given by:

$$NS = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)^2}{\sum_{t=1}^{T}(Q_o^t - \overline{Q_o})^2} \tag{3.11}$$

where $\boldsymbol{Q_m^t}$ is modelled flow at time step t, $\boldsymbol{Q_o^t}$ is observed flow at time step t and $\boldsymbol{\overline{Q_o}}$ is average observed flow.

3.2.7 Model updating and data assimilation

Calibration of the PDM is undertaken in simulation mode where rainfall and potential evaporation form the model inputs and calculated flow forms the model output. Data assimilation techniques can be incorporated into calibrated models to

enable to use new observed flows data to improve model performance. Models that incorporate data assimilation techniques are commonly used in real-time application and are often referred as in forecast mode. Data assimilation techniques employed in forecast mode are mainly required to continuously update the model using current and past flow data to improve the model forecasts for different lead times from the forecast origin. The PDM used in this study provides state correction and error prediction data assimilation techniques, which use a set of rules for adjusting model states and predicting future errors respectively based on new incoming observed flow data (Moore 2007). The term state in the state correction technique is used to describe a variable in the model which mediates between inputs and the model outputs. In the case of PDM the main state variables are the water contents of the surface and ground water stores, and the probability distributed soil moisture storages (Table 3.1). The performance of error prediction model in providing improved forecasts depends on the degree of persistence in model errors, which is discussed in detail in Chapter 2. In the case of flow prediction, errors show a tendency to oscillate rapidly and most widely in the vicinity of the rising limb and peak of predicted flow hydrographs (Casale & Margottini 1999), which are considered important part of the hydrograph in terms of informing surface water abstraction. In contrast, state variables are key hydrological variables in conceptual rainfall-runoff models such as PDM and are known to largely influence the partition of rain between runoff and infiltration, hence control model output flows (Moradkhani et al. 2005). Consequently, the state correction technique, which is discussed in detail in Chapter 2, is found to be more suitable for use in this study and is employed in the PDM forecast model. Figure 3.3 illustrates the process of incorporating new incoming observations using data assimilation technique, where model forecast are corrected based on estimated errors from a set of observed data and forecasted outputs. As shown in Figure 3.3, when observed data is available the difference between the forecast and the observation are used as feedback to correct the model states in earlier time steps and the corrected model is used to make forecast.

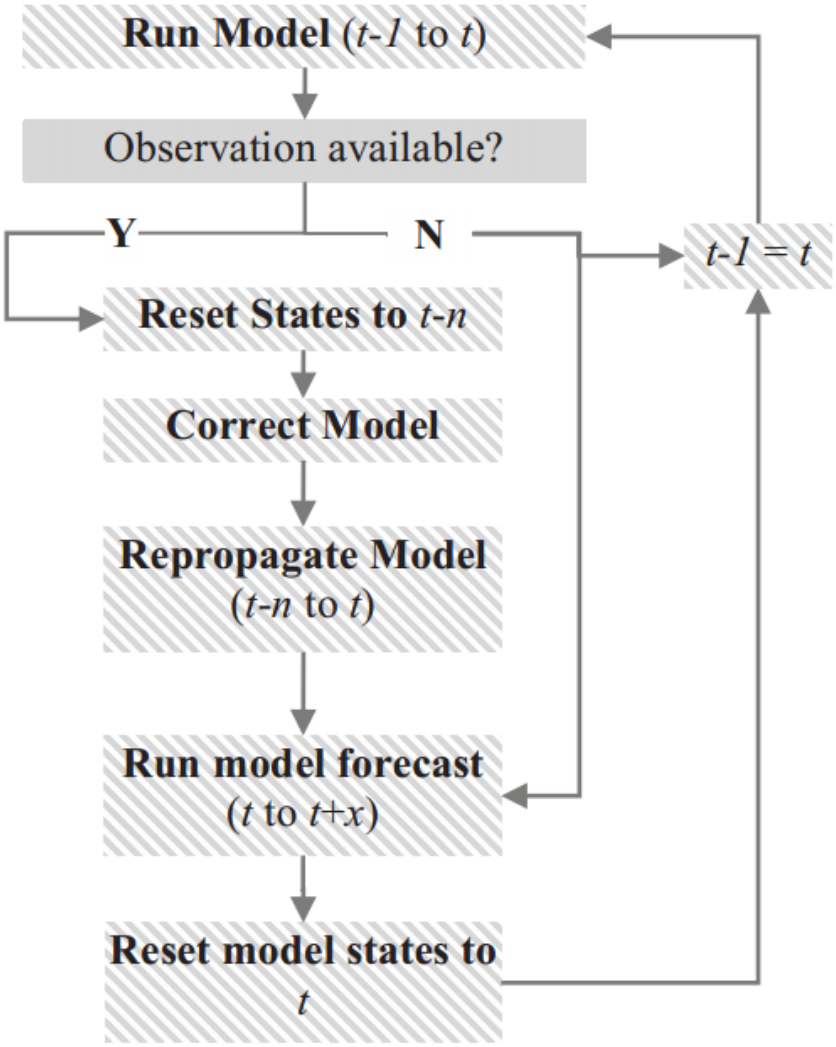


Figure 3.3. Implementation of data assimilation technique in flow forecasting model

3.2.8 Water resource management model

A simple water resource management model is developed and used to propagate simulated surface water abstractions through a water resource system to assess implications of real-time abstraction management scheme on reservoir levels and associated water resources management decisions in the catchment. This involved the simplification of the water resource system as an isolated, self-contained unit served by the storage reservoir which is supplied by the streamflow. A historical analysis is conducted in which actual water supply data from the catchment water

source over the 2011 dry period are used as water demand from the system. This enabled an assessment of the benefit of using real-time abstraction management during a period when abstraction of extra amount of water is highly beneficial. At each time-step, volume in the storage reservoir is calculated based on historical demand, abstraction from the surface water and the operational constraints of the reservoir. Abstraction of water from the river at each time-step is constrained by the intake capacity, the total storage capacity (19845 m^3) and abstraction licence conditions (streamflow should be maintained at prescribed flow of 159Ml/d at the abstraction site in this catchment). In this study, the intake capacity is set based on the average of daily maximum abstraction volumes achieved during wet periods as computed using historical abstraction data. Given these constraints, abstraction is defined within the model as the maximum permissible value at each time-step.

The total reservoir capacity is divided into four operational zones as shown in Figure 3.4 below and different operational policies are applied when available storage is in each of these zones. The flood control zone occupies the top 2% of the total reservoir capacity and is maintained only for use in extreme rainfall events to provide storage for runoff and avoid flooding of nearby areas (i.e. limiting capacity to 98% of the total storage). The dead storage zone occupies the bottom 13% of the total reservoir capacity and is maintained as inactive storage, where the water is not used for operational purposes. A control curve is used to split the remaining capacity into two zones called conservation zone and buffer zone. The control curve defines the storage volume required to be maintained in the reservoir at the beginning of each month in order to ensure continuous and reliable supply of water is provided to meet full demand. Control curves are developed by water planners based on design drought inflows and demands, which can be converted in to required storage volume by the process of balancing inflow and outflow.

At each time-step, if the storage volume is above the control curve, the model allows free release of water from the reservoir to meet demand in full. However, if the storage volume in the reservoir drops below the control curve, a step-wise restriction of water release from the reservoir is also applied depending on percentage drops in

storage volume. These procedures represent drought management actions such as the use of demand restrictions (e.g. hosepipe bans, nonessential use bans and severe water rationing) which are adopted by water resource managers to maintain adequate supply of water during dry periods. In addition, the model allows for relaxation of the abstraction licensing constraint allowing additional water intake from the river despite a drop in river flow beyond the original minimum flow requirement specified in the abstraction licence. This accounts for drought permits, which are implemented by water utility operators in agreement with Environment Agency to reduce abstraction licence restrictions during drought periods. A control curve currently being used for the operation of the reservoir in the study catchment is used in this study as shown in Figure 3.4. If the storage level drops to dead storage zone now water is released from the reservoir. The water resources management model accounts for all these real-world operational constraints and determines daily volume of abstraction with certain level of risk using forecasted flow at the abstraction site and the associated uncertainty information, which are both derived from the probabilistic flow forecast model.

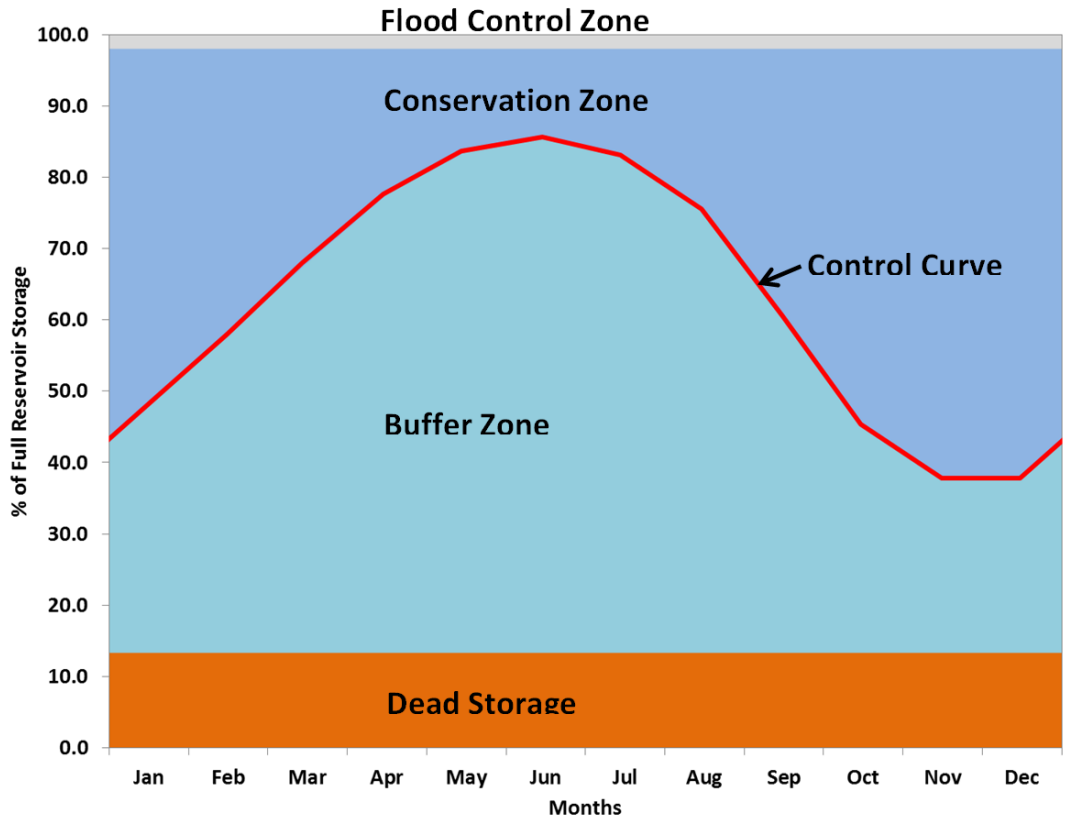


Figure 3.4. Reservoir storage components and control curve for reservoirs in the River Dove catchment

3.3 Results and Discussion

3.3.1 Runoff modelling and uncertainty analysis

The DREAM method was employed to randomly generate 2500 samples from posterior probability distribution function of the model parameters. Initially the parallel MCMC chains in DREAM are initialized using the Latin hypercube sampling, which is used to sample from the uniformly distributed prior parameter ranges specified in Table 3.1. Convergence of the parallel MCMC chains to the posterior distribution is monitored using the R-statistics of Gelman & Rubin (1992). The posterior probability distributions of PDM parameters provided the required information to summarize simulated flow variability caused by parameter uncertainty. Following convergence of the chains to a stationary distribution, PDM is

evaluated for each set of parameter draws derived from DREAM to propagate the parameter uncertainties through the model and obtain simulated flow distribution. These values are summarized using 95 percentiles to reflect the impact of parameter uncertainty on the model prediction (Figure 3.5). Five years of data are used for model calibration using DREAM over the period spanning from April 7, 2004 to April 7, 2009 and the remaining four years of data are used for validation purposes.

Simulated river flows and error statistics over the calibration and validation period are shown in Figure 3.5 and Table 3.2 respectively. The Nash-Sutcliff coefficient computed for the best fitted flow simulation resulted from optimal parameter set for the calibration and validation period are found to be 0.72 and 0.7 respectively. Moriasi et al. (2007) has reviewed performance ratings of catchment scale hydrological models in the literature and in general Nash-Sutcliff coefficient values greater than 0.7 are rated as acceptable performance. However, the intended application of the model and the extent at which model output errors impact on water resource management decisions need to be carefully considered. In this study, the model output is intended to inform abstraction management decisions that focuses on service flow ranges as discussed in section 2.2.2. Thus, impacts of out layers at both high and low flows on abstraction management decisions are considered to be relatively low and the Nash-Sutcliff coefficient values of 0.72 and 0.7 are considered acceptable. The R-factor values in the two periods are found to be 0.87 and 0.84 as shown in Table 3.2, which indicated the average width of the 95% uncertainty band is reasonable as compared to the distribution of observed flow data. The P-factor values in the calibration and validation period are found to be 0.89 and 0.71 respectively, i.e. 89 % of the observed flow values fall within the 95 % prediction uncertainty bounds of the ensemble flow simulation in the calibration period, whereas 71 % fall within the bound in the validation period. It is observed that both in the calibration and validation period significant proportions of low flow observations lie at the tail end of the uncertainty band. This indicates the model is more susceptible to errors when predicting low flows as compared to medium range and high flows, which indicates the models relative limitation in terms of adequately

representing base flow contributions during non-runoff periods. Most of the 11% of observations during calibration period and 29% of observations during validation period that lie outside the 95% prediction uncertainty bound are observed to occur during low flows. Comparison of observed flow data and simulated flow ensembles over the entire calibration and validation period also showed that some level of deviations between the observed and simulated flow ensembles during high flows, which are not of particular importance with respect to the main objective of this work. As discussed above, capability of the model to accurately predict service flow ranges is considered more important for the intended purpose of the model to inform abstraction management, hence the outliers during very low and high flow periods will have less impact on abstraction decisions and the P-factor values of 0.89 and 0.71 computed for calibration and validation periods respectively are considered acceptable. , i.e In general, the simulated and observed flow hydrographs are observed to agree well with acceptable levels of error statistics and the calibrated and validated model can be considered fit for the intended practical purpose of informing surface water abstraction management.

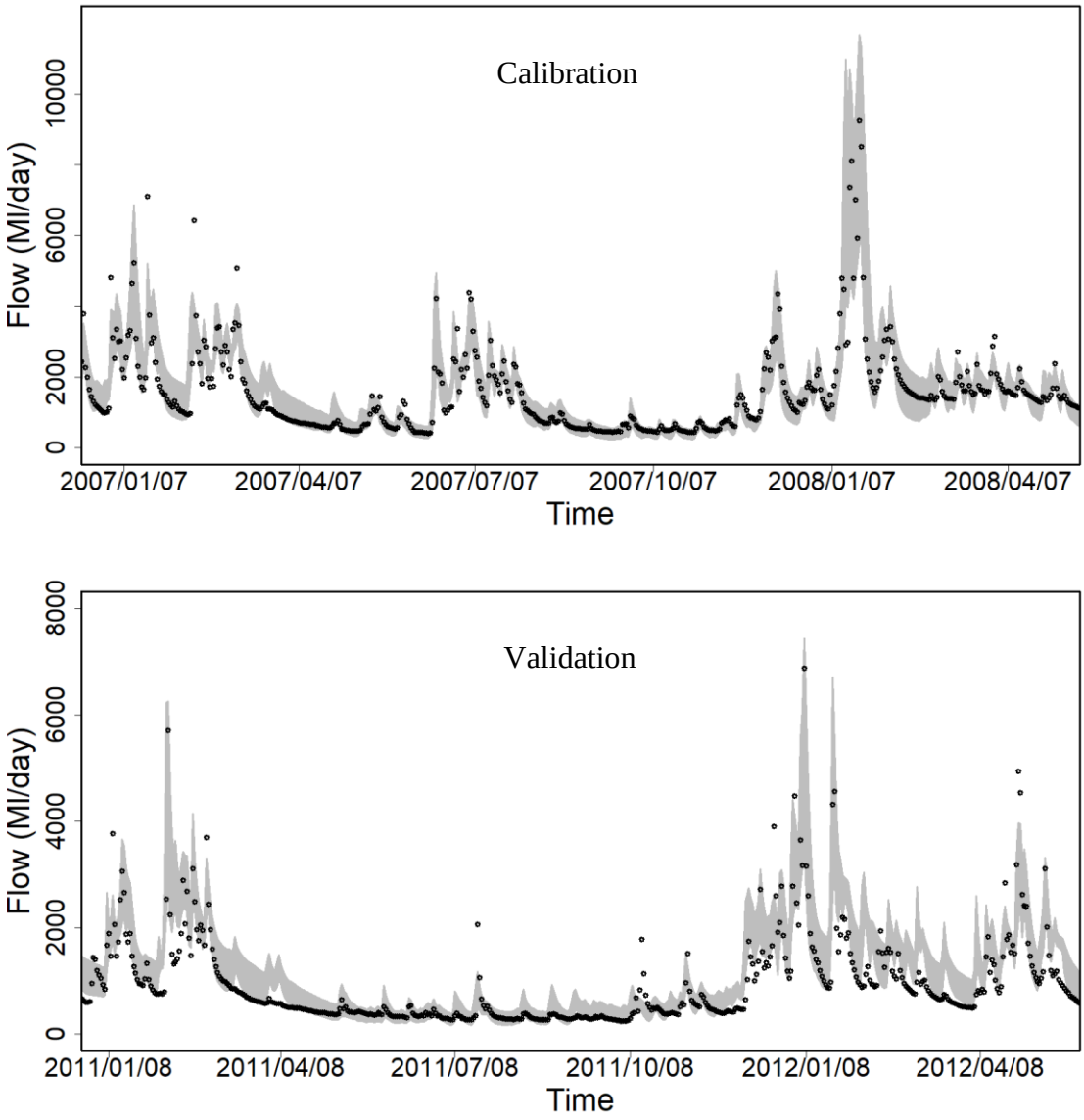


Figure 3.5. Streamflow prediction uncertainty ranges derived with DREAM for a representative portion of the calibration and validation period. The grey shaded area represents 95% prediction interval, whereas the black dots denote recorded streamflow observation data.

Table 3.2. The values of evaluation criteria in calibration and validation period

	NS Coefficient	P-factor	R-factor
Calibration Period	0.72	0.89	0.87
Validation Period	0.7	0.71	0.84

3.3.2 Flow forecasting

Following the calibration and validation of the model, data assimilation techniques are employed into the model and flow forecasts are generated. To assess the flow forecasting capability of the model, a number of fixed origin forecasts are made over the model calibration and validation period, and are compared with observed flows. Fixed origin forecast used here is defined as time series of forecasted flow values computed from a single forecast time origin. Each forecast value is associated with a lead time, which specifies the difference between the times of the forecasted value and forecast origin. The applicability of a flow forecast model to inform surface water abstraction is dependent on the catchment response time, which is influenced by meteorological, hydrological and catchment variables (land use, soil type, storages, catchment and channels geomorphology). Catchment response time is estimated in this study using lag time (LT) parameter, which is defined as the time from the centroid of rainfall to the peak discharge time of total runoff (Gericke & Smithers 2014). This method is employed as shown in Figure 3.6 to estimate lag time in the study catchment using a rainfall event. Relatively long lag time of 26 hours is estimated for the study catchment as shown in Figure 3.6. This indicates that rainfall events that occur after the time of forecast (forecast origin) are less likely to affect flows at the outlet of the catchment within 26 hours of the rainfall occurrence; hence forecasted flows with lead time of up to 24 hours can be effectively used to inform surface water abstraction. However, forecasted rainfall data can be used if prediction time of longer than 24 hours is required.

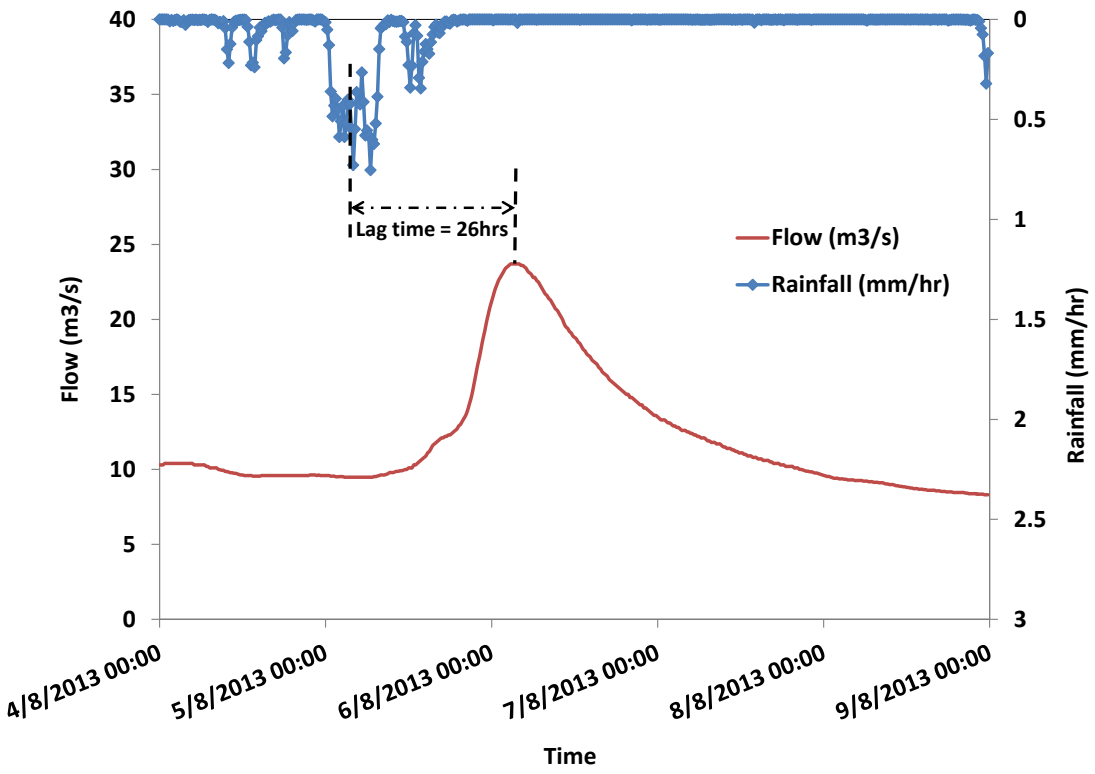
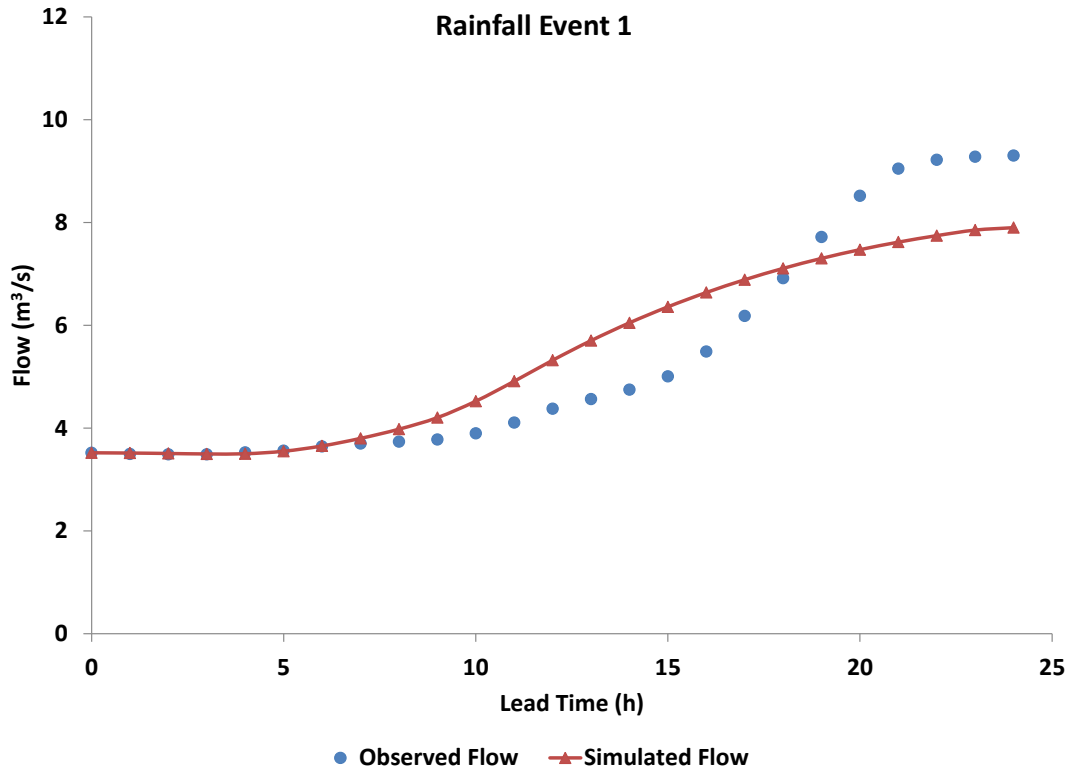


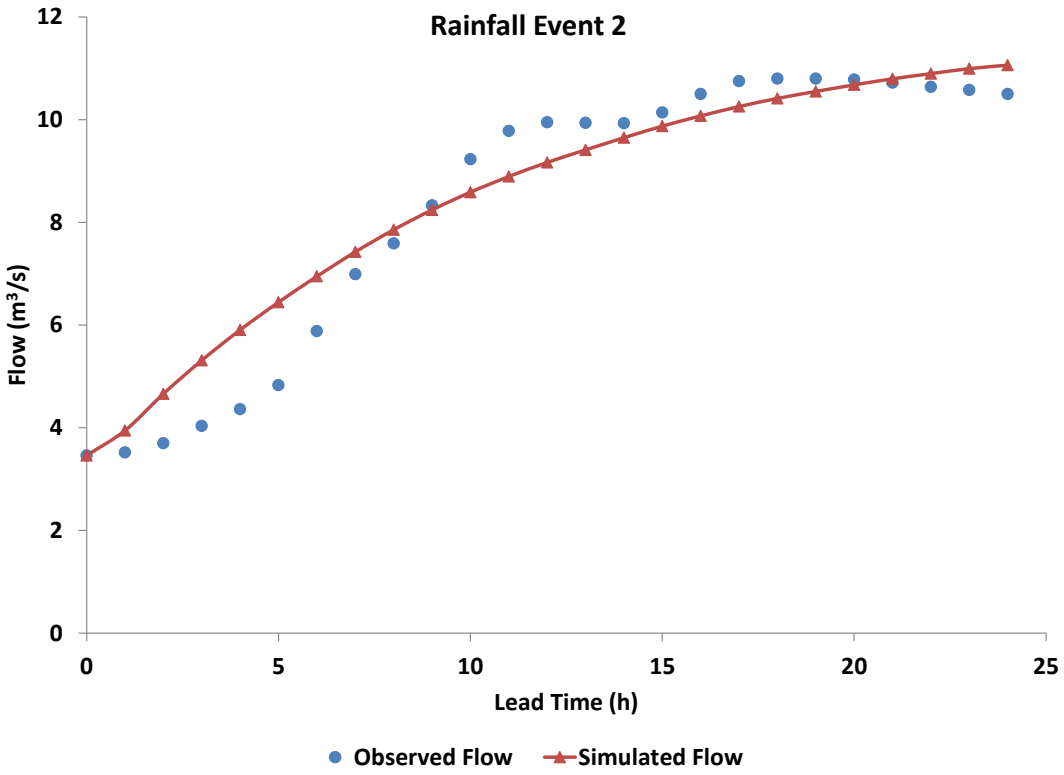
Figure 3.6. Lag time estimated as time between the centroid of rainfall and peak discharge time.

Deterministic and stochastic flow forecasts are performed over four different rainfall events (rainfall events 1-4) and are compared with observed flows to assess the flow forecasting performance and suitability of the calibrated PDM rainfall runoff model to inform surface water abstraction. Optimum parameter values are used to make deterministic flow forecasts over these four rainfall events which produced runoff events of various scales. These forecasts provide information on the availability of water at the abstraction site and enable abstraction operators to make informed decisions. The lead time in Figure 3.7 and Figure 3.8 below specify the time interval between the arrival time of the forecasted flow and time of forecast started (the origin). The performance of the deterministic model in forecasting the four rainfall-runoff events of varying magnitude with up to 24 hour lead time is found to be practically acceptable, considering the level of errors in these forecasts and their impacts on abstraction volumes. However, in real-time it is advisable to use the

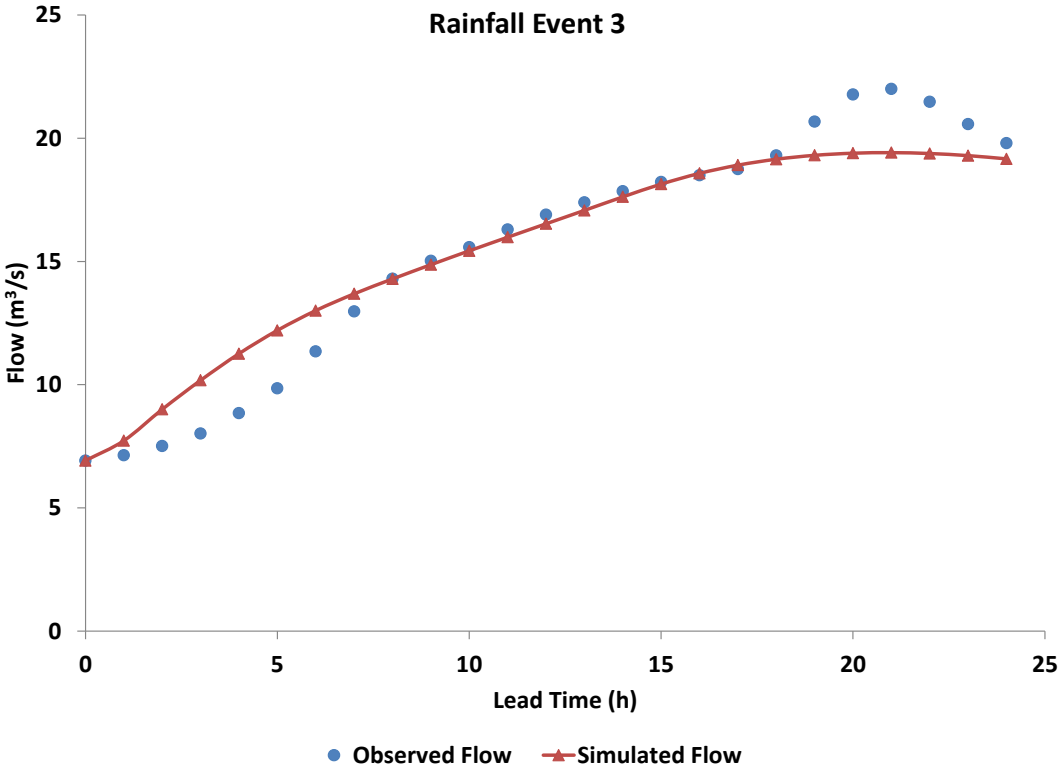
probabilistic forecast model to provide additional information on the level of model forecast errors and enable abstraction operators make risk aware decisions as discussed below.



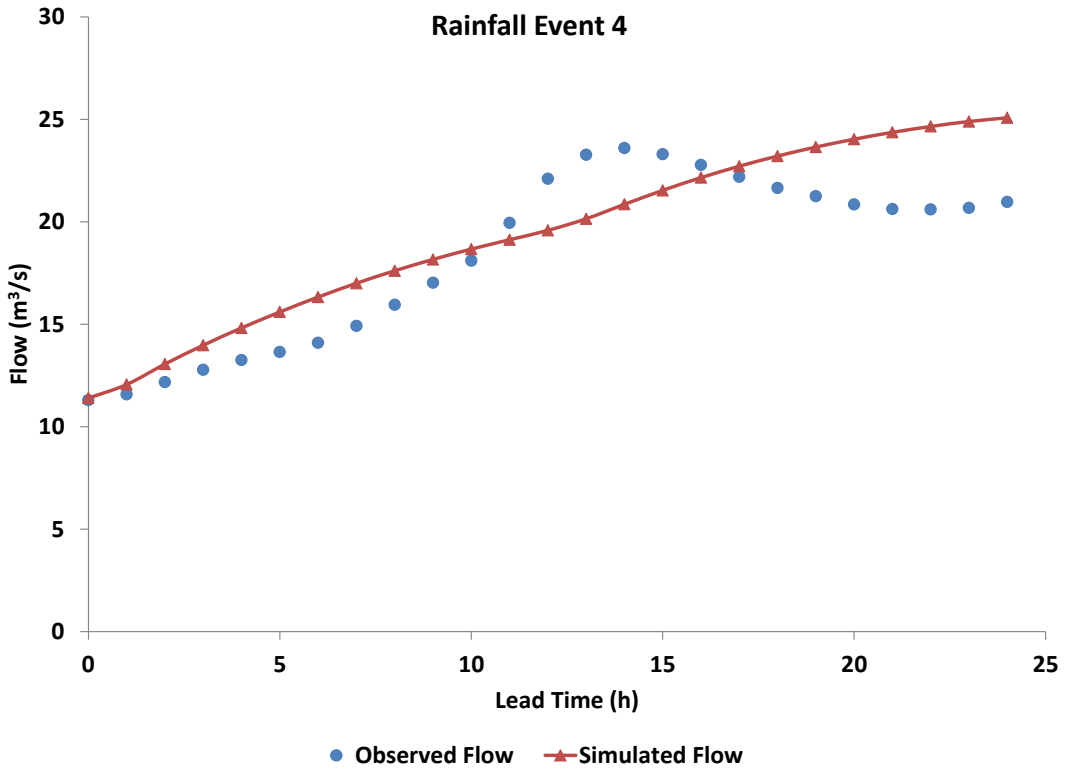
(a) Deterministic flow forecast for rainfall event 1 (on June 12, 2006) with different lead times (h) from forecast origin of 12/06/2006 06:00AM



(b) Deterministic flow forecast for rainfall event 2 (on August 12, 2005) with different lead times (h) from forecast origin of 12/08/2005 05:45PM



(c) Deterministic flow forecast for rainfall event 3 (on May 13, 2007) with different lead times (h) from forecast origin of 13/05/2007 05:15AM

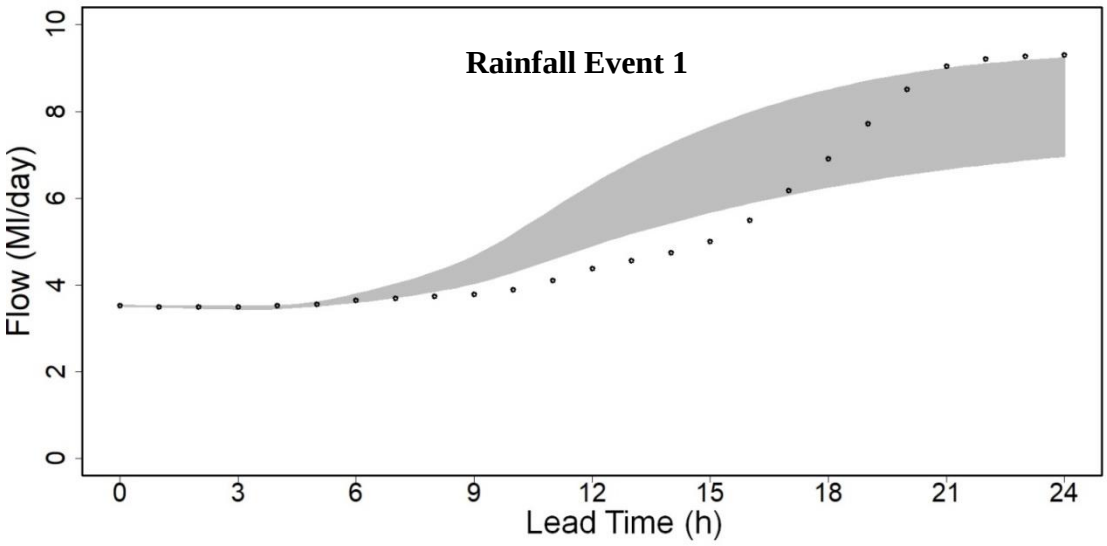


(d) Deterministic flow forecast for rainfall event 4 (on January 10, 2006) with different lead times (h) from forecast origin of 10/01/2006 03:45PM

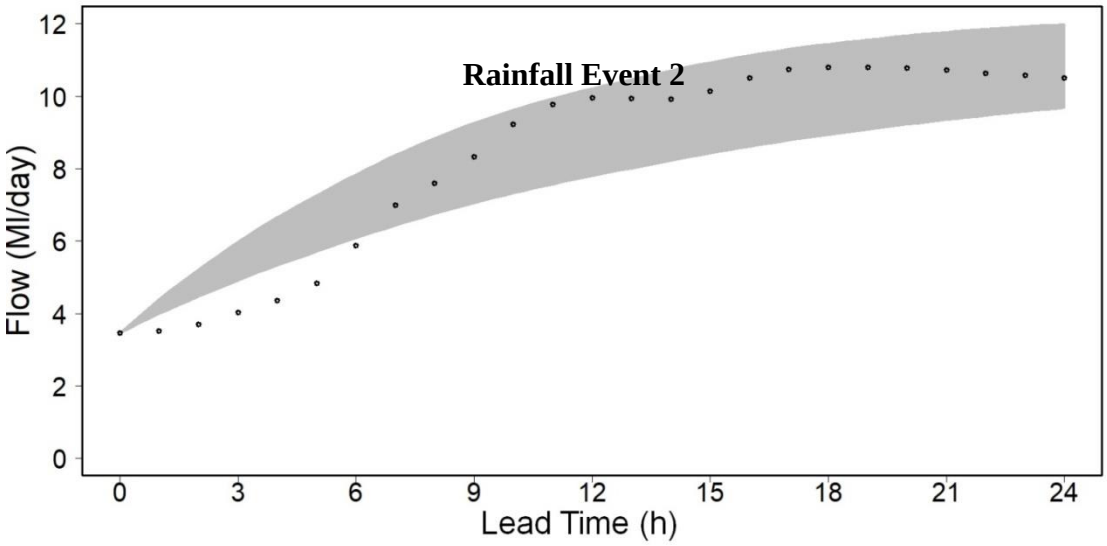
Figure 3.7. Deterministic flow forecasts using optimum parameter values for four different rainfall events.

The set of parameter values sampled from the posterior distribution resulted from DREAM are used to make probabilistic flow predictions as shown in Figure 3.8 below. On average, about 75% of the observations fall within the 90% uncertainty band of the probabilistic predictions over the four rainfall events. Moreover, the observations that lie outside the 90% prediction uncertainty are mostly observed to lie close to the boundaries of the uncertainty bounds. Thus, the overall performance of the probabilistic model is considered to be practically acceptable for the intended purpose of abstraction management application. Abstraction volumes are regulated using a set of abstraction licence conditions, which water resource operators need to carefully consider when making abstraction decisions. Thus, the use of the probabilistic flow prediction enables to make risk aware abstraction decisions and

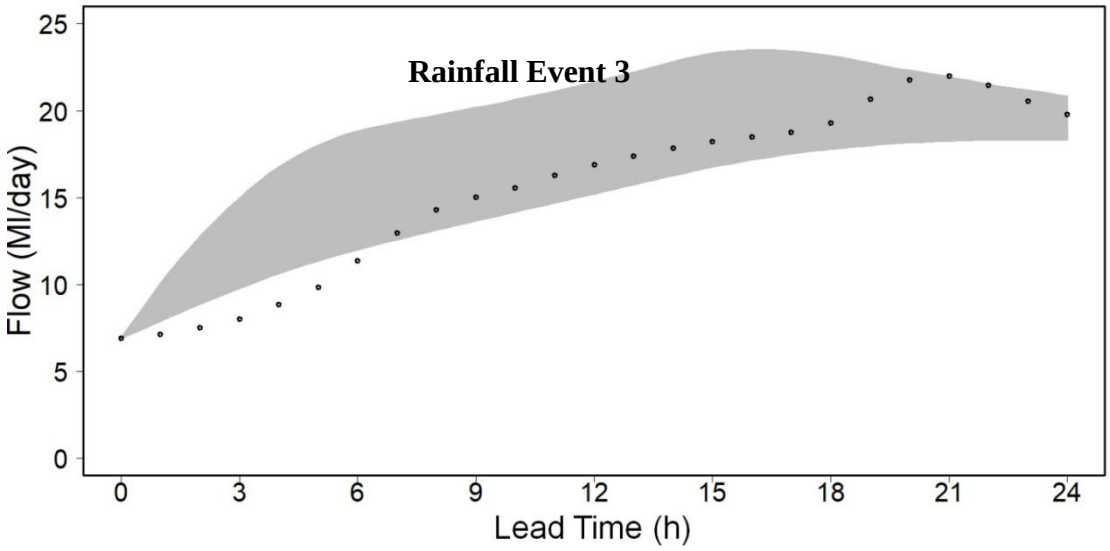
thus helps to avoid the risk of breaching abstraction licence conditions that can occur due to model errors. Furthermore, the data assimilation feature of the model enables to include newly observed flow data that helps to improve flow predictions using the state updating method discussed in section 3.2.7.



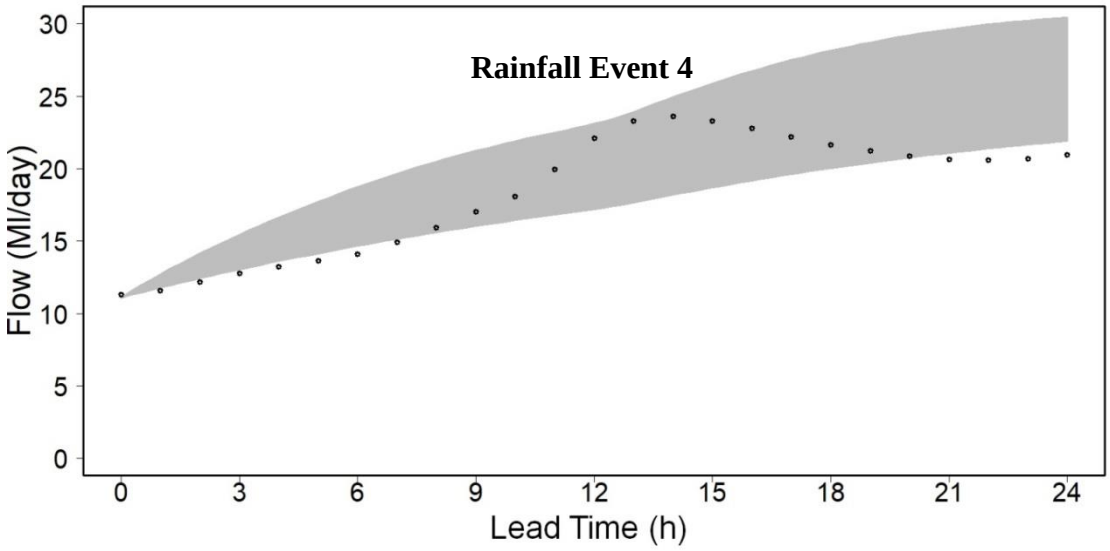
(a) Probabilistic flow forecast for rainfall event 1 (on June 12, 2006) with different lead times (h) from forecast origin of 12/06/2006 06:00AM



(b) Probabilistic flow forecast for rainfall event 1 (on August 12, 2005) with different lead times (h) from forecast origin of 12/08/2005 05:45PM



(c) Probabilistic flow forecast for rainfall event 1 (on May 13, 2007) with different lead times (h) from forecast origin of 13/05/2007 05:15AM



(d) Probabilistic flow forecast for rainfall event 1 (on January 10, 2006) with different lead times (h) from forecast origin of 10/01/2006 03:45PM

Figure 3.8. Probabilistic flow forecasts resulted from a set of parameter values sampled from posterior distribution (90 % prediction uncertainty).

Forecasted flows with acceptable level of risk will feed into the water resources management model to calculate the daily amount of water available for abstraction at the surface water abstraction site in the catchment. The water resource management model, which will be used to inform abstraction decisions and also discussed in detail in section 3.3.3, takes into account a range of real-time operational constraints in addition to the availability of water at the abstraction site. The proposed method enables to make informed surface water abstraction decisions with a view to providing a potential to abstract more water, which otherwise could be missed due to lack of data and information on the daily amount of available water that can be abstracted without breaching licence conditions (section 3.1).

3.3.3 Implications on water resources management

The potential benefits of the real-time abstraction management scheme, which uses hydrological forecasting in order to maximize the amount of water abstracted in the study catchment, is investigated in this section. Following the calibration and validation of the rainfall-runoff model, daily simulated abstraction volumes are generated from simulated flows over a period of ten years (2004 – 2014). Comparison of these daily amounts of simulated abstractions with abstraction record data (historical abstraction data) over the same period have shown that on average 67 ML of water per day have been missed at the abstraction site, which could have been abstracted sustainably without breaching abstraction licences (Figure 3.9). The missed volume of water shown in Figure 3.9 and Figure 3.10 for each day is calculated by subtracting the daily volume of water historically abstracted (recorded daily abstraction volume) at the abstraction site from the daily volume of water that could have been abstracted on that day as estimated by the model (simulated abstraction volume). Figure 3.10 includes model simulated abstraction, historical abstraction and missed volume of water calculated over the 2011 dry period and provides a graphical explanation of the computation of missed volume of water discussed above. The restrictions imposed by the licence conditions and hydraulic capacities have been applied to determine the amount of water that could have been abstracted, which is named as simulated abstraction. Water treatment works are

assumed to work at their achievable maximum capacity during the computation of simulated abstraction.

However, only some part of the 67 Ml/d water found to be missed at the abstraction site is believed to have occurred during low flow periods, during which as much water as possible is required. Surface water abstraction is mainly driven by two main factors. During low flows, surface water abstraction volumes are restricted by the availability of water in the river (based on abstraction licence conditions), whereas demand and/or storage levels in reservoirs determine abstraction levels during high flow periods. The amount of water identified here as missed over the analysis period is primarily resulted from the following two main factors, which are both associated with abstraction management.

1. Abstraction licences include conditions to protect the aquatic environment and restrict daily abstraction volumes based on specified daily flow values at abstraction sites. These are often known as 'hands off' flow conditions referring to a fixed flow in the river below which abstraction is proscribed. Thus, during low flow periods, the amount of available water for abstraction directly varies with river flow. However, current abstraction decisions are not informed on the amount of predicted daily flows in rivers and usually a conservative approach is taken when making daily abstraction decisions to avoid the risk of breaching licence conditions. This results in missing a certain amount of water daily that could have been abstracted without breaching the licence conditions.
2. During high flow periods, abstraction volumes are not limited by licence conditions, and thus water can be abstracted using maximum abstraction capacities. However, during these periods abstraction is mainly driven by demands and/or drops in storage levels in reservoirs. Consequently, some amount of water is missed daily at abstraction sites, which could have been abstracted if intakes and water treatment works operate at their achievable daily maximum capacities. Even

though there may not be any need for additional water during these periods in the study catchments, enabling to use these missed opportunities to abstract more water can help to reduce abstractions from other more strained and/or more expensive sources in the region.

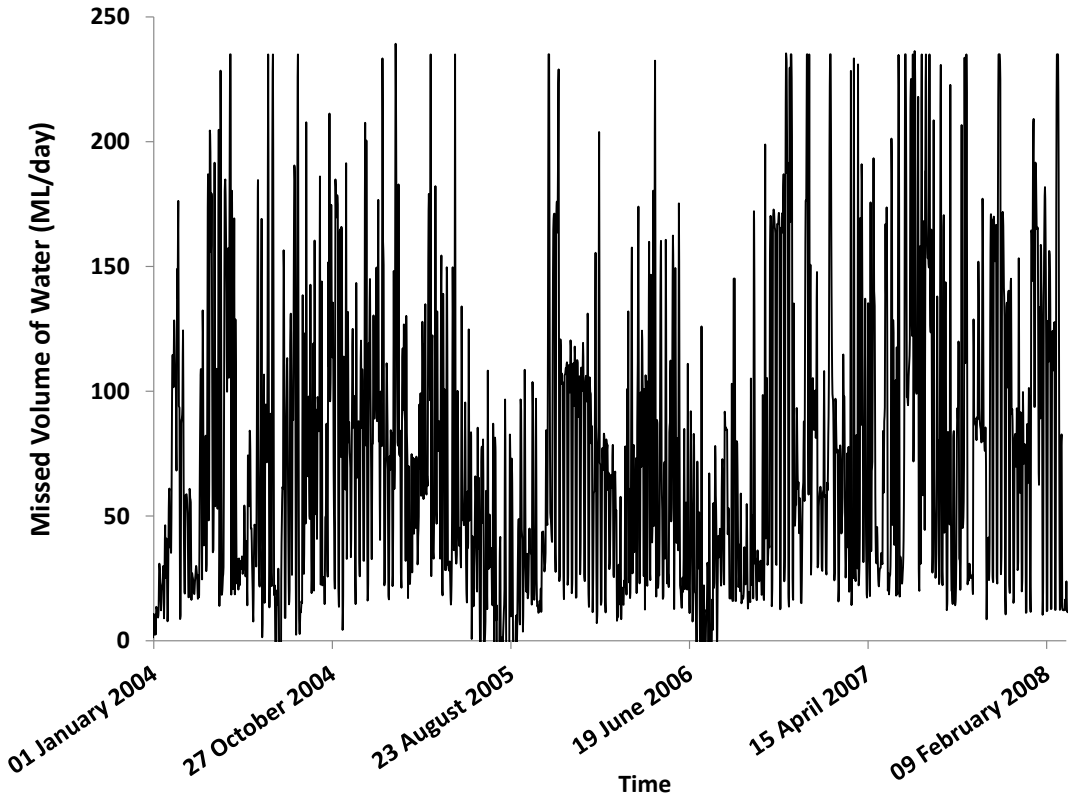


Figure 3.9. Amount of water missed over the period of 2004 – 2014 at abstraction site.

To investigate the real-time abstraction management scheme's capability to minimize the amount of water missed during low flow periods, similar assessment was required to be performed over a specified dry period. Thus, the calibrated rainfall-runoff model was coupled with the water resources management model to generate daily simulated abstraction volumes using outputs from the deterministic flow model. These are compared with historical abstractions over the 2011 dry period in the study catchment. This comparison has showed that on average a total of 30 Ml of more water per day could have been abstracted using the real-time abstraction management scheme during this period (Figure 3.10). The spikes in the missed volume of water

graphs in Figure 3.9 and Figure 3.10 can occur as a result of two main operational constraint and management decisions. During dry periods, the spikes on the missed volume of water graphs are likely to have been caused due to shut down or restrictions on some or all intake pumps due to operational outages, whereas during normal periods the spikes can be caused by either intake pump outages or operational management decisions to reduce or suspend abstractions, which is often made during high storage levels in the reservoir to reduce power consumption, due to water quality issues or low demand on the reservoir.

The integrated flow forecast and water resource management models are employed to assess implications of the use of real-time abstraction management on reservoir levels and associated water resource management decisions. The results have showed that rapid decline in reservoir levels during dry periods such as the one in 2011 can be avoided by using the real-time abstraction management scheme proposed in this study (Figure 3.11). The decline of reservoir storages beyond the control curve during dry periods activates drought management actions such as imposing water use restrictions, which affect the service level of water suppliers. These conditions also lead to drought permit applications to reduce restrictions in abstraction licence conditions, which allow abstraction of more water beyond the normal licence conditions. This exacerbates impacts of dry periods on the water environment downstream of the abstraction site. Thus, the extra amount of water that can be abstracted using RTAM scheme can also help to minimize impacts on the water environment in addition to ensuring adequate water supply during dry periods.

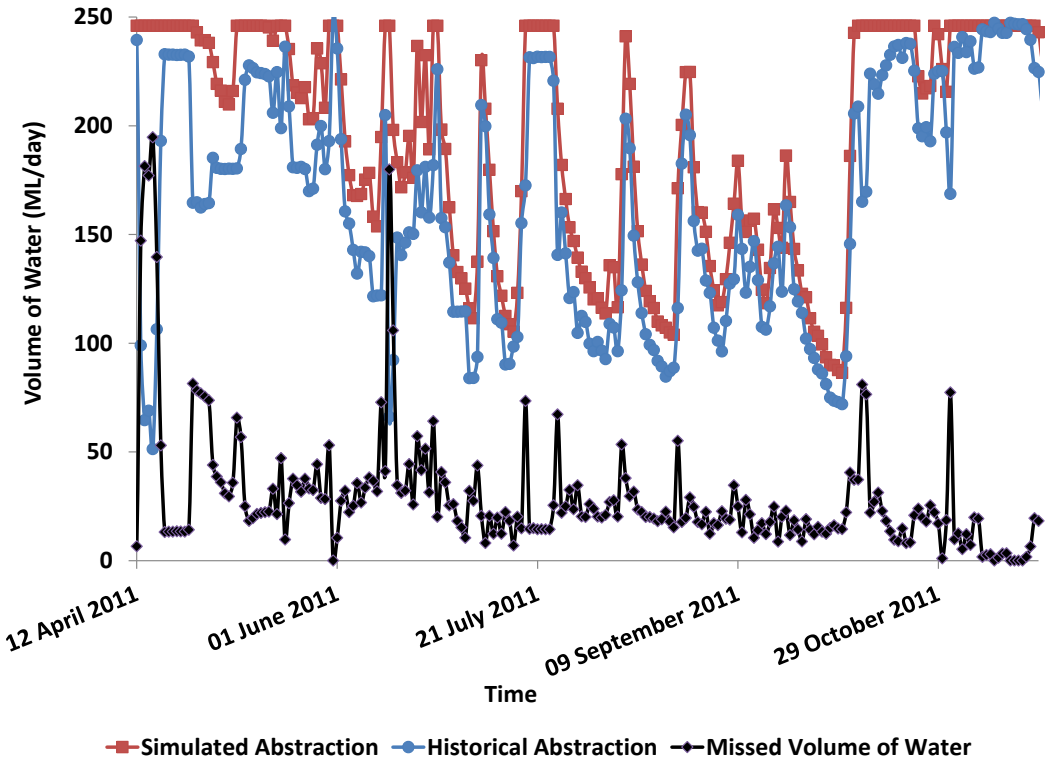


Figure 3.10. Comparison of simulated and historical abstraction and amount of water missed over the 2011 dry period at the abstraction site.

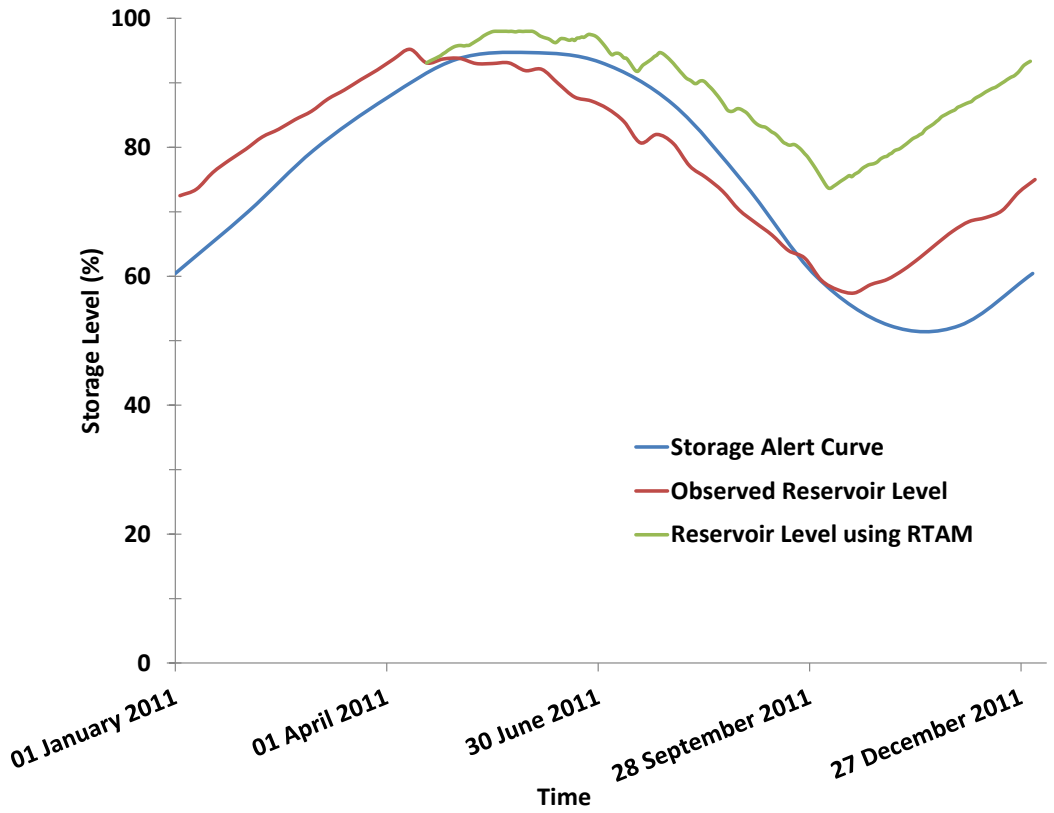


Figure 3.11. Observed reservoir levels over the 2011 dry period and reservoir control curve.

3.4 Conclusions

A real-time abstraction management scheme was devised in this study by employing hydrological forecasting and water resource management model with a view to enable to abstract more water in the study catchment. Calibration of conceptual rainfall-runoff models such as PDM involves tuning the value of parameters that lack physical basis and thus cannot be inferred from direct measurements. Thus, assessing uncertainty associated with these parameters and their impacts on predicted flow is an important task when using these models to inform water resources management decisions. DREAM, a Bayesian uncertainty analysis tool based on Markov chain Monte Carlo method, is used in this study to calibrate the PDM model in the river Dove catchment in UK. Flow prediction results from the model using the parameter sets sampled from the posterior distribution showed that PDM is able to reproduce

observed flows with reasonable accuracy for the study catchment. Deterministic and stochastic flow forecasts from the calibrated model are observed to match well with observed flows, which showed the capability of the model for the intended practical purpose of forecasting available volume of water at surface water abstraction sites and informing abstraction decisions. Simulated flow derived from PDM model using optimum values of the calibration parameters are propagated through a water resource management model over the 2011 dry period in the study catchment. This has showed the significant role that the RTAM scheme proposed in this study can play in increasing resilience of the water supply system in the study catchment and minimizing abstraction impacts on the environment. Water resources planners are required to state the number of times a system will fail to meet full demands over a specified planning period, which forms the level of service defined by water suppliers. Effective implementation of the RTAM system in surface water catchments used for water supply can help to achieve these specified levels of service by minimizing the probability of failing to meet demand during dry periods. Moreover, by dynamically linking abstraction volumes to actual availability of water in the source, the RTAM scheme helps to make surface water abstraction management systems ready for future potential abstraction reforms discussed in Chapter 2 and challenges associated with climate change and increasing demand from a growing population. This work has demonstrated a new approach to develop and integrate a probabilistic flow forecast model with water resources management model with a view to providing operationally suitable and sustainable solution to tackle emerging issues of increasing demand, climate change and associated policy reforms.

Chapter 4

Modelling of metaldehyde concentrations in surface waters: A travel time based approach

4.1 Introduction

Diffuse pollution is a significant threat to the quality of surface water systems, with agricultural runoff commonly recognised as posing the greatest risk (Grayson et al. 2008). Observed levels of diffuse agricultural pollutants in surface water have increased as pesticide application rates have intensified in most countries (Wilson & Tisdell 2001), detection methods have improved and new products emerge in the market (Loucks et al. 2005). The characteristic behavior of some of these pollutants (e.g. pesticides such as metaldehyde) mean that existing drinking water treatment processes are inadequate to reduce levels to within drinking water regulation limits and thus have recently become a recognized problem to water infrastructure operators (Lu et al. 2017). D'Arcy et al. (1998) recommends that efforts to mitigate diffuse pollution problems are best taken at catchment scale (as promoted by the Water Framework Directive) using approaches such as catchment management practices and abstraction management, which can help to avoid or reduce the need for energy and cost intensive engineered treatment solutions. However, the complex nature of the processes involved in diffuse pollutant generation and transport in rainfall runoff, along with high temporal and spatial variations in pesticide application and rainfall/runoff events pose challenges for the development and establishment of accurate and reliable modelling and mitigation strategies (Ouyang et al. 2017). Current understanding of short term pollutant dynamics in catchments caused by rainfall/runoff processes is limited due to the scarce availability of water quality data at suitable temporal resolutions (Bach et al. 2001).

The aims of this chapter are to; 1. Develop a new model to describe the fluctuation of a diffuse agricultural pollutant (metaldehyde) in surface waters caused by rainfall driven runoff; 2. Validate the model against new high resolution datasets of metaldehyde concentration within the catchment following rainfall and runoff events.

It is anticipated that the new model can be used to forecast metaldehyde concentrations and inform short term abstraction decisions such that runoff containing high levels of metaldehyde can be avoided.

Metaldehyde is an organic compound with the formula $C_8H_{16}O_4$ and has low sorption coefficient (K_{OC}) of active ingredient to organic carbon value that ranges between 34 - 240 L kg^{-1} (Kay & Grayson 2014). It is a soluble molluscicide that is used heavily in a range of agricultural products to control slugs and snails (Li et al. 2010) and has a relatively long half-life in soil that ranges between 3.17 – 223 days. In recent years high levels of metaldehyde exceeding the European and UK standards for pesticides in drinking water value of 0.1µg/l have been observed in surface waters during the application season (NFU 2013). Peak concentrations in surface waters are observed particularly following rainfall events (Kay & Grayson 2014). Water quality assessments carried out by the UK water industry on more than 2300 raw water abstraction sites in England and Wales have identified that 110 abstraction sites are at risk of metaldehyde pollution (Water UK 2013). Metaldehyde is not effectively removed using conventional drinking water treatment options such as granular activated carbon and ozone due to its high inherent stability resulting from a unique molecular structure (Webber 2014), and is hence a particular concern for water infrastructure operators.

Diffuse pollutants such as metaldehyde present on farmlands can enter river systems via a number of pathways including surface runoff, drains and groundwater flow. The dominant pathway for any particular pollutant is mainly dependent on its properties, weather conditions, soil type, land slope and network of drains in the area (Bach et al. 2001). However a number of studies have shown that runoff is the dominant pathway for most diffuse agricultural pollutants (Huber et al. 2000; Heathwaite et al. 2005; Huber et al. 1998; Bach et al. 2001). Migration of pollutants through erosion is considered significant only for highly adsorbing substances with K_{OC} values greater than 1000 L kg−1 (Kenaga, 1980). Hence metaldehyde tends not to be adsorbed by suspended solids and sediments due to its low K_{OC} value. This suggests that the transport of metaldehyde through runoff in dissolved form is more

significant than transport via soil erosion. Hence, the amount and rate of runoff generated from specific farmlands in the catchment where metaldehyde is applied combined with runoff travel time along flow paths are likely to be critically important in determining metaldehyde concentrations and dynamics in surface waters. Several studies have emphasized the significant impacts of rainfall induced runoff in mobilizing pesticides into streams (e.g., Vryzas et al. 2009; Taghavi et al. 2011; Du Preez et al. 2005; Ng & Clegg 1997). However, studies quantifying peak pollutant loads in runoff and potential exposure to downstream receivers resulting from individual rainfall events are lacking due to the need for high resolution water quality datasets, which are rarely available. Most available water quality data are in daily or coarser time resolutions that fail to capture short term fluctuations in diffuse pollution concentrations caused by individual rainfall driven runoff events. Lack of high resolution validation data has also limited the development of stormwater quality models that are capable of predicting pollutant concentrations in runoff at small time intervals, and hence be utilised in abstraction management systems. The use of automatic water samplers has been identified as a step forward towards addressing this problem (Berenzen et al. 2005; Rabiet et al. 2010).

In this study, automatic samplers were used to collect hourly surface water samples following rainfall events within a UK catchment known to be subjected to high metaldehyde concentrations. This enabled the validation of a new operationally suitable stormwater quality prediction model within the catchment. The new model aims to enable the prediction of short term fluctuations in metaldehyde concentrations arriving at a surface water abstraction site which is used for drinking water supply. Whilst a complete understanding of the transport and fate of pesticide in catchments requires consideration of numerous processes such as groundwater transport and reaction/degradation processes, the nature of the organic compound (methalydyde) as well as the focus on forecasting short term fluctuations in response to rainfall events lead us to propose a modelling approach based on the aggregation of overland surface flow travel times over the catchment, allowing a simpler and more practical model structure than a model incorporating numerous longer term

processes such as groundwater transport or erosion. The model is therefore based on the identification and routing of spatially distributed metaldehyde loads in runoff using build-up, wash-off and runoff travel time techniques. The approach proposed here provides an improvement to existing stormwater quality models by using high resolution radar rainfall data and identifying application risk areas in the catchment, which enables the consideration of spatiotemporal variations of pollutant generation and transport in the catchment. A raster based data structure is employed in the model and thus various spatially distributed catchment characteristics such as elevation, soil type, land use and rainfall are described in the model using grids. The use of the developed model in water supply catchments can help quantify potential exposures to peak metaldehyde concentrations at surface water abstraction sites with the aim of enabling better surface water abstraction management. Given the inadequacy of existing water treatment processes in removing metaldehyde, smarter abstraction management informed by predicted arrival of peak pollutant levels at abstraction sites proposed in this study provides a cost-effective and sustainable solution to tackle problems caused by diffuse pollutants.

4.2 Methodology

This section describes the study catchment as well as the development of a new process based metaldehyde transport model to forecast short term fluxes of metaldehyde in surface waters in response to individual rainfall events. The catchment is divided into 12 million grid cells of 25 square metre each and runoff generation, routing and pollutant wash-off is calculated within each cell in response to time series rainfall data collected using radar. The model is calibrated and validated using monitored flow data as well as new high resolution datasets of metaldehyde concentrations collected following rainfall events using automatic samplers.

4.2.1 Study area

The study area, River Leam catchment, is located in the sub basin of River Severn in central England and drains an area of 300 km^2 (Figure 4.1). Elevation within the

catchment ranges from 46m to 232m above sea level with a mean annual rainfall of 649mm. A UK Environment Agency flow gauging station is situated at the outlet of the catchment that records flow data, which is used to assess abstraction licence conditions against daily abstraction volumes. The normal flow depth of the River Leam at the gauging station ranges between 0.24m and 1.16m with an average flow of 1.55m^3/s. The most dominant land cover type within the catchment is arable farmland consisting of horticultural plants and cereals. Managed grassland is the second most common land use type with few urban, suburban and rural developments in the catchment. Hence, agriculture is the dominant land use in the catchment and is likely to have a significant influence on river water quality. The predominant soil types in the catchment are clayey and loamy soils, which make up approximately 65.5 % of the total area. Clay soils are vulnerable to compaction, remain wet for longer periods and have a slow natural drainage, leading to sheet runoff as opposed to channel erosion. The remainder of the catchment consists of freely draining slightly acid loamy soils or loamy and clayey soils which are not seasonally wet but suffer from impeded drainage.

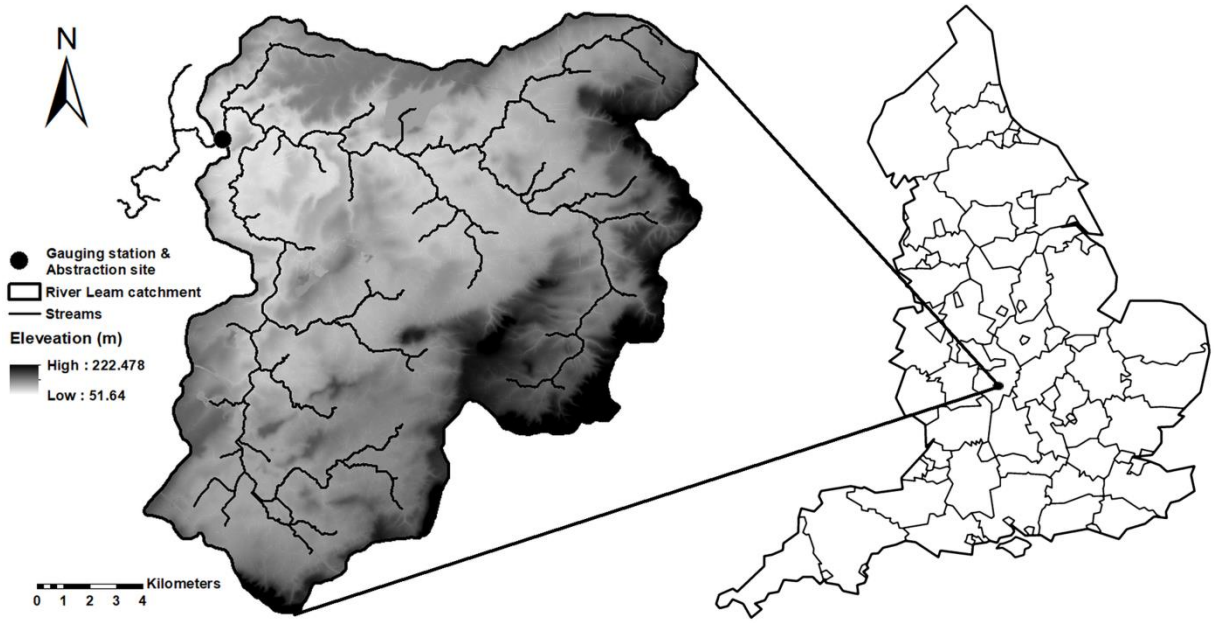


Figure 4.1. Location of River Leam Catchment, gauging stations and abstraction site.

The largest use of surface water in the catchment is for public water supply. A surface water abstraction site, located at the outlet of the study catchment as shown in Figure 4.1, is used by a water utility operator to pump water to impounding reservoirs for water supply purposes. The main water quality issues in the catchment are nutrients and pesticides from diffuse sources. Metaldehyde is typically applied in the catchment on arable farmlands that grow winter crops such as winter wheat, potatoes and oilseed rape, which usually cover about one third of the catchment area and is rotated on a seasonal basis. Because of favorable conditions for slugs during the usually wet autumn and winter seasons, metaldehyde applications are typically made between September and December. Routine monitoring conducted by the local water infrastructure operator shows that high levels of metaldehyde are present in the river during the application season (Figure 4.2) (STW 2017). The analyses in the current study focus on data collected in the catchment during the metaldehyde application season over the period 2014 -2017.

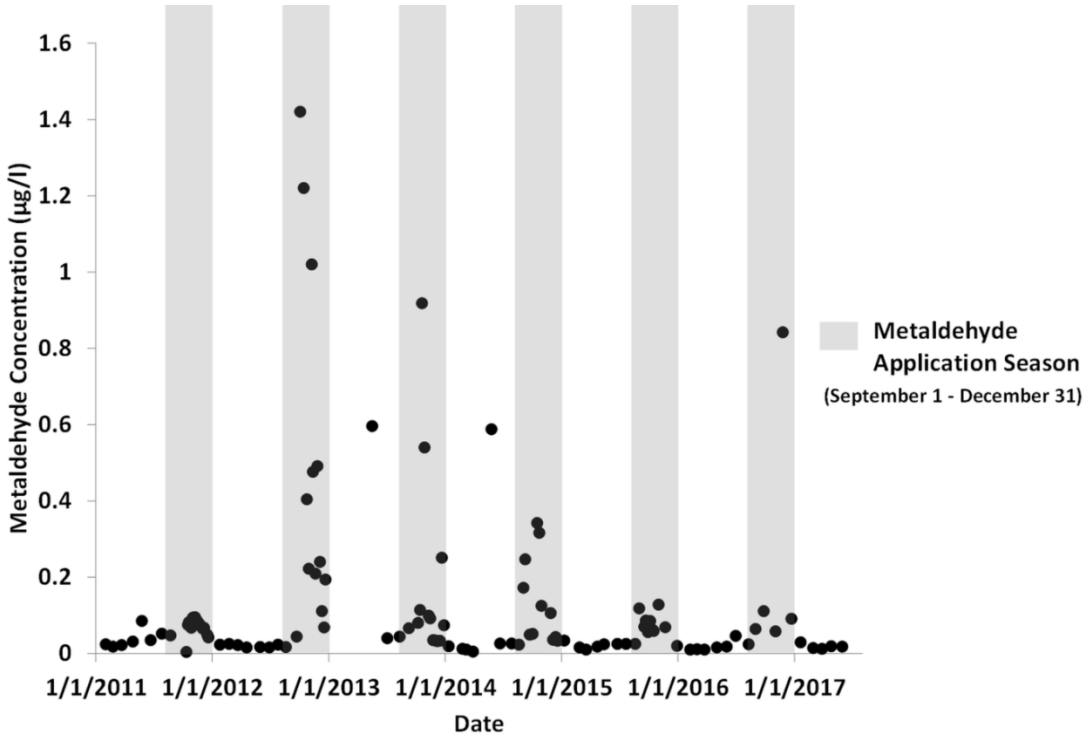


Figure 4.2. Historic seasonal variation of metaldehyde concentration in the River Leam near the catchment outlet from routine monitoring (STW 2017).

4.2.2 Development of metaldehyde prediction model

The model presented in this paper is comprised of three components: runoff generation, runoff routing and pollutant build-up/wash-off. Runoff is calculated based on overland flow generated from each grid cell in the catchment during monitored rainfall events. The travel time based runoff routing method estimates storm runoff transport from catchment grid cells to the outlet of the catchment based on geographic information system tools. The spatially distributed time variant direct runoff travel time technique employed in the model accounts for spatial and temporal variability of runoff generation and flow routing through overland flows and stream networks (Melesse & Graham 2004; Du et al. 2009) following rainfall events. The pollutant model estimates metaldehyde build-up through pesticide applications on identified metaldehyde high risk areas and its wash-off to water courses during runoff processes. The travel time based runoff routing and build-up wash-off models are integrated to enable rainfall event based prediction of metaldehyde concentrations at the catchment outlet.

4.2.3 Runoff generation

The differential form of the Soil Conservation Service (SCS) curve number (CN) method (Mancini & Rosso 1989) is used to compute spatially distributed excess rainfall in each grid cell within the study catchment. The SCS-CN runoff volume prediction method was originally developed by the United States Department of Agriculture (USDA) Soil Conservation Service (Hjelmfelt 1991). Detailed procedures of the method were originally documented in the National Engineering Handbook, Sect. 4: Hydrology (NEH-4) in 1956 and subsequently revised in 1964, 1971, 1985, 1993 and 2004 (Li et al. 2015). It is a widely used, well established technique owing to its computational simplicity and use of accessible catchment data. The differential form of the SCS-CN method to calculate cumulative excess rainfall depth I_t(mm) at time step t from each grid cell is given by:

$$I_t = \frac{(P_t - 0.2S)^2}{(P_t + 0.8S)} \quad (when\ P_t > 0.2S) \tag{4.1}$$

where $P_t\ (mm)$ is the cumulative depth of rainfall at time step *t*, calculated as

$$P_t = \sum_{i=1}^{t} p_i \Delta t \tag{4.2}$$

where p_i is the rainfall intensity at the timestep i (mm/s), Δt is time step length (s). S is the maximum soil retention potential (mm), given by $S = 25400/CN - 254$. where CN is the "curve number" ranging between 1 & 100 (Hjelmfelt 1991).

When $P_t <= 0.2S$, rainfall is completely absorbed by soils with no overland flow generation and hence resulting in zero runoff depth. Initial CN values for each study year were first determined based on hydrologic soil group (HSG), land use and hydrologic conditions data (Mishra & Singh 1999). In addition to the soil type, which mainly identifies the soil water retention capacity, antecedent moisture condition plays an important role in runoff generation (Crespo et al. 2011). In the SCS-CN method, the effect of soil moisture on runoff generation is incorporated by adjusting CN values based on antecedent moisture condition (AMC) categories. No exclusive relations or formulas are available to calculate soil moisture from antecedent rainfalls of certain preceding days, but in general the term antecedent for soil moisture calculation purpose is taken to vary from 5 to 30 preceding days (USDA 1986). AMC categories in this study were determined for each rainfall event based on cumulative rainfall volumes of the preceding 5 days. The three AMC categories are: AMC-I for dry, AMC- II for normal, and AMC-III for wet conditions. Initially assigned CN values are adjusted for each rainfall event based on their AMC categories to account for the effect of soil moisture on runoff generation.

Figure 4.3 shows a map of CN values over the River Leam catchment based on normal antecedent moisture condition (AMC – II) for the 2014 application season.

The spatially distributed CN values combined with the use of radar rainfall data enable the computation of spatially distributed runoff depths.

Once the runoff depths are computed, the runoff rate Q_t (mm/s) from each grid cell at time step *t* can be calculated as

$$Q_t = (I_t - I_{t-1})/\Delta t \tag{4.3}$$

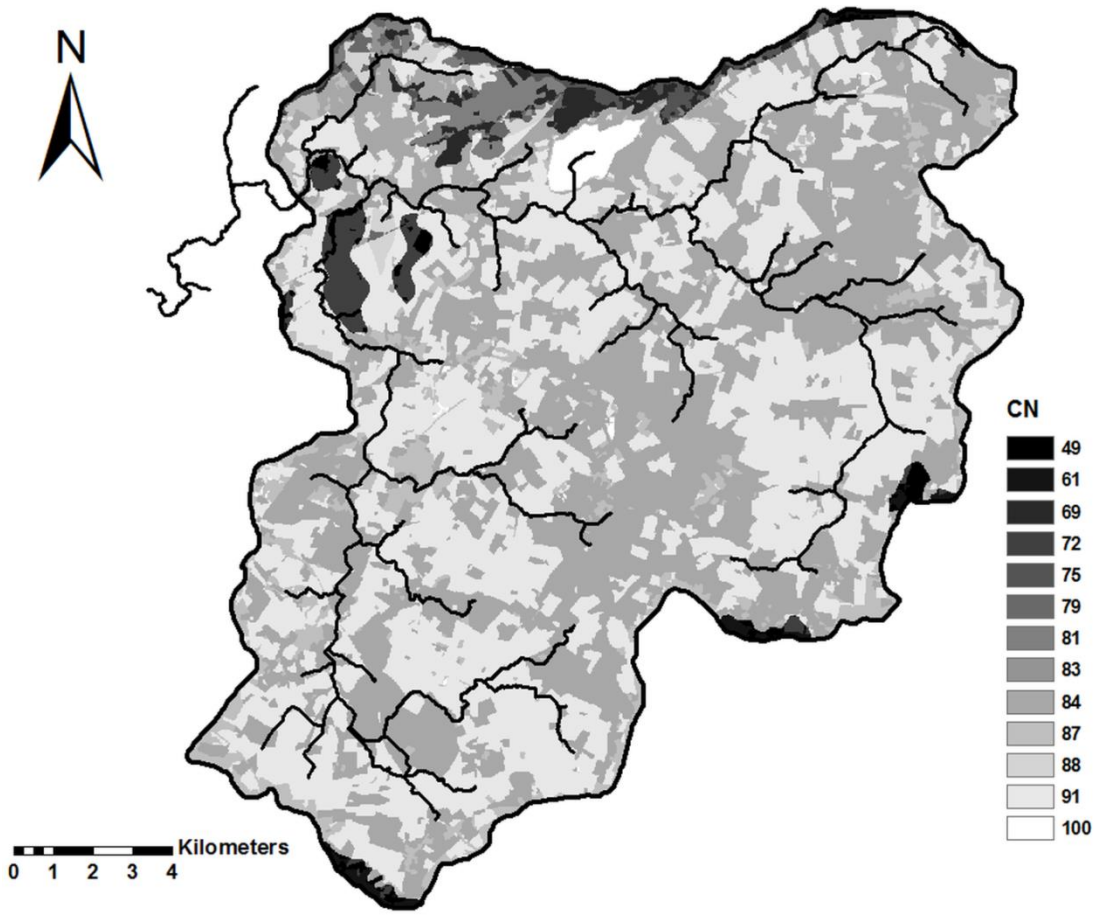


Figure 4.3. Spatial distribution of runoff Curve Numbers based on normal antecedent moisture condition (AMC – II) for the year 2014.

4.2.4 Runoff routing

In natural conditions, over land and channel travel times vary based on availability of runoff and rainfall variation in time. This is accounted in the model by employing a

time variant travel time computation technique. To determine flow pathways, a GIS flow direction tool was used to determine the steepest descent from every cell in the catchment Digital Elevation Model (DEM) along which storm runoff flows. This created unique connections between cells that enabled to define flow paths to catchment outlet and identify storm runoff flow networks in the catchment. A threshold number was set to identify cells with high flow contributing areas that form concentrated flow and were used to delineate channel networks in the catchment (Du et al. 2009). The delineated channel network density and extents were compared with stream networks from topographic maps to adjust threshold number of cells. Any cell with less upstream flow contributing cells than the threshold was considered as overland flow cell and others with more flow contributing upstream cells were classified as channel cells. Travel time computation techniques were then employed to determine travel time for each overland and channel flow cells based on available runoff in the cells as described below.

Cumulative travel times through each pathway computed from topographic data were used to route excess rainfall from each grid cell along flow paths to determine runoff hydrographs at the outlet of the catchment. First, kinematic wave theories suggested by Wong (1995, 2003) were used to derive travel time expressions for each grid cell depending on its classification i.e. overland flow cell or channel cell. For an overland flow grid cell with negligible flow backwater effect, the wave celerity (c) travelling down the grid cell was derived using kinematic wave equation and is given by (Eagleson 1970):

$$c = \frac{dx}{dt} = \alpha\beta y^{\beta-1} \tag{4.4}$$

where, α and β are parameters used in $q = \alpha y^{\beta}$ to relate discharge per unit width (q) to flow depth (y) and x is distance along flow direction.

Re-writing equation (4.1) in terms of discharge per unit width (q) gives

$$c = \frac{dx}{dt} = \alpha^{1/\beta}\beta q^{1-(1/\beta)} \tag{4.5}$$

For small period of time, it can be assumed that overland grid cells receive constant and uniform excess rainfall, i and constant upstream inflow, q_u. Thus, the unit discharge at the downstream end of the grid cell over that period can be calculated as

$$q = q_u + ix \tag{4.6}$$

Assuming α is independent of x, substituting equation (4.6) in equation (4.5) and solving the derivatives in equation (4.5) for t gives an expression for time of concentration, which is generally defined as the time required for runoff to travel along flow path from the hydraulically most distant point in the catchment to the outlet, for overland grid cells as:

$$t_c = \frac{1}{\alpha^{1/\beta}}\left[\frac{(q_u + iL)^{1/\beta} - q_u^{1/\beta}}{i}\right] \tag{4.7}$$

where t_c the time of concentration and L is the length of the grid cell in the direction of flow. In general, overland flow time of concentration for small grid areas such as that considered in this study are shorter than duration of excess rainfalls and equation (4.7) can thus be used to calculate the travel time (Eagleson 1970).

The time of concentration formula can be written as:

$$t_c = \left[\frac{Li^{1-\beta}}{\alpha}\right]^{1/\beta}\left[(\lambda + 1)^{1/\beta} - \lambda^{1/\beta}\right] \tag{4.8}$$

where λ relates upstream inflow and influx from excess rainfall as follows:

$$\lambda = q_u / iL \tag{4.9}$$

Values of friction parameters α and β can be obtained using Manning's equation as $\alpha = \sqrt{S}/n$ and $\beta = 5/3$ respectively. Thus, expression for overland flow time of concentration from equation (4.8) can be written as:

$$t_c = 7\left(\frac{nL}{S^{0.5}}\right)^{0.6} i^{-0.4}[(\lambda+1)^{0.6} - \lambda^{0.6}] \tag{4.10}$$

where, the units of parameters in equations (4.9) and (4.10) above are given as minutes for t_c, m/m for S, m^2/s for q_u, mm/h for i, and m for L. Manning's n values vary depending on the types of surface and can be selected from values recommended by Engman (1986).

The equivalent of equation (4.8) for channel flow grid cells with negligible backwater effect, a constant upstream inflow, and a uniform lateral inflow is given as

$$t_{tc} = \left(\frac{L_c}{\alpha_c q_L^{\beta_c - 1}}\right)^{1/\beta_c} \left[(\lambda_c + 1)^{1/\beta_c} - \lambda_c^{1/\beta_c}\right] \tag{4.11}$$

where t_{tc} is time of concentration, L_c is the length of the channel cell in flow direction, q_L is the uniform lateral inflow, α_c and β_c are parameters relating the discharge (Q) in the channel to the flow area (A); and λ_c relates the upstream inflow (Q_u) to the lateral inflow (q_L) as follows:

$$Q = \alpha_c A^{\beta_c} \tag{4.12}$$

$$\lambda_c = \frac{Q_u}{q_L L_c} \tag{4.13}$$

Backwater effect occurs when a water body is obstructed at downstream by a tide or any structures such as dam or bridge, which causes the water to backup along the channel and limits or stops the downstream flow currents. There are no tides or obstructions along the river in the study catchment and thus, the assumption of negligible backwater effect in this chapter is considered reasonable.

Replacing $\alpha_c = \sqrt{S}/n$ and $\beta_c = 5/3$ friction parameter values determined from Manning's equation and uniform lateral inflow ($q_L = iL_c$) in equation (4.11) above gives the channel flow time of concentration as:

$$t_{tc} = 7\left(\frac{nL_c}{S^{0.5}}\right)^{0.6} (iL_c)^{-0.4}\left[(\lambda_c + 1)^{0.6} - {\lambda_c}^{0.6}\right] \tag{4.14}$$

where, the units of parameters in equations (4.13) and (4.14) above are given as minutes for t_{tc}, m/m for S, m^3/s for Q_u, mm/h for i, and m for L_c.

To account for uncertainties introduced in the estimation of travel time, calibration parameters K_o and K_c are included in equation (4.15) and equation (4.16) as shown below to determine travel time in overland (t_c) and channel flow (t_{tc}) respectively.

$$t_c = 7K_o\left(\frac{nL}{S^{0.5}}\right)^{0.6} i^{-0.4}[(\lambda + 1)^{0.6} - \lambda^{0.6}] \tag{4.15}$$

$$t_{tc} = 7K_c\left(\frac{nL_c}{S^{0.5}}\right)^{0.6} (iL_c)^{-0.4}\left[(\lambda_c + 1)^{0.6} - {\lambda_c}^{0.6}\right] \tag{4.16}$$

The value of K_o and K_c parameters are determined by calibration. Finally, travel times calculated for each grid cells using equation (4.15) and equation (4.16) above are summed along flow paths to determine cumulative travel time of runoff from each grid cell to the catchment outlet.

4.2.5 Pollutant model

The pollutant model estimates metaldehyde build-up on high risk areas during dry days and wash-off to water courses during runoff following rainfall events. Metaldehyde risk areas in the catchment have been identified based on available land use data, which provides information on the likelihood of metaldehyde being applied to the land based on crop type during each growing season. Land growing winter crops such as winter wheat, potatoes and oilseed rape, where metaldehyde is commonly applied are identified as high risk areas. Data on land use derived from satellite imagery was acquired from the Centre for Ecology and Hydrology for each growing season used in the analysis (2014-2017). Figure 4.4 shows the identified high risk areas for the 2014 season.

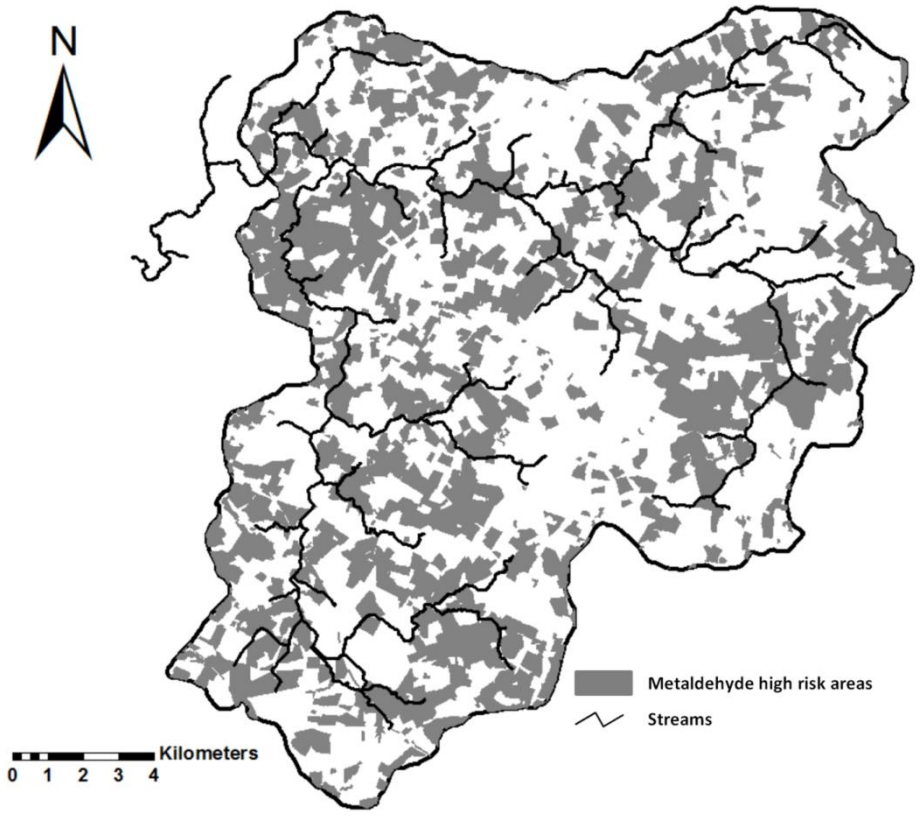


Figure 4.4. Identified Metaldehyde high risk areas in the catchment for the year 2014.

Metaldehyde application doses on these high risk areas and frequency of applications over pesticide application periods determine the accumulation of metaldehyde in the

active zone at soil surfaces (Müller et al. 2003). Moreover, the time interval between metaldehyde application and a rainfall event directly affects the amount of metaldehyde transported to water bodies through runoff. These processes are represented using build-up and wash-off components in the model. As shown in Figure 4.4, farmland in the study catchment that have high likelihood of metaldehyde being applied (metaldehyde high risk areas) are spread-out in the catchment with some parts of the catchment containing more density of high risk areas than others. As a result, the metaldehyde concentration at the catchment outlet over a specific time period is heavily dependent on the density of high risk areas within the relevant travel time isochrones. High rate of runoff generation from high risk areas increases metaldehyde levels in the river, whereas high rate of runoff generation from low risk areas have a dilution effect and can lower concentration of metaldehyde in the river. Thus, metaldehyde concentration at the outlet of the catchment significantly depends on spatial variability of a rainfall event, in relation to the distribution of high risk areas.

Pollutant build-up: Metaldehyde build-up on high risk areas occur through application of pesticides that contain metaldehyde as an active ingredient. Wet conditions during winter provide ideal environment for slugs to thrive and most metaldehyde applications are made during this period to control winter crops. Typical single slug pellet application based on guidelines from manufacturers is 5 kg/ha. This is equivalent to 75g/ hectare (0.19 g per $5m^2$ grid size used in this study) of metaldehyde based on a commonly used 1.5% slug pellet. The statutory legal requirement in the UK on metaldehyde application states that total application in a calendar year should not exceed a maximum of 700g/ha. Routine monitoring data collected by the local water infrastructure operator shows that almost all high levels of metaldehyde in the river have occurred during the September to December application season (Figure 4.2). Thus, it can be assumed that most of the 700g/ha statuary annual legal limit of metaldehyde is applied during the September to December period. Based on this assumption and the typical single metaldehyde application value of 75g/ha, a total of no more than nine applications are expected

during the winter crop growing season on any particular high risk farmland. This combined with the relatively long half-life of metaldehyde in soil suggest that metaldehyde presence on farmlands during this period is likely to be consistently high (Castle et al. 2017). In this study, it was initially assumed that metaldehyde was applied on all high risk areas 5 days before rainfall events, which was later adjusted using a calibration parameter.

Pollutant wash-off: Metaldehyde wash-off is dependent on a number of rainfall, catchment and substance characteristics. In this study, pesticide loss equation based on the ‘‘simplified formula for indirect loadings caused by runoff’’ (SFIL) (Berenzen et al. 2005; Reus et al. 1999) is used to calculate percentage loss of metaldehyde at each timestep from high risk areas through runoff.

$$L_t = \frac{Q_t}{P_t} f e^{-t_n \cdot \frac{ln2}{DT_{50soil}}} \frac{100}{1 + K_d} \tag{4.17}$$

where:

- L_t - Percentage of application dose that is washed by runoff water as a dissolved substance at timestep t,
- Q_t - Runoff depth generated at timestep *t* (mm),
- P_t – Total precipitation depth (mm),
- f - Correction factor, with $f = f_1 f_2 f_3$, f_1 (Berenzen et al. 2005; Reus et al. 1999)
- Slope factor: $f_1 = 0.02153 * slope + 0.001423 * slope^2$ if $slope < 20\%$ or $f_1 = 1$ if $slope > 20\%$,
- f_2 - Plant interception factor: $f_2 = PI/100$,
- f_3 - Buffer zone factor: $f_3 = 0.83W$ with W - width of the buffer zone (m),
- t_n - Number of days between application and a rainfall event,
- DT_{50soil} - Half-life of active ingredient in soil (days),

- K_d - Ratio of dissolved to sorbed pesticide concentrations; with $K_d = K_{OC} * \%OC * 1/100$, (Berenzen et al. 2005; Reus et al. 1999)
- K_{OC} - Sorption coefficient of active ingredient to organic carbon,
- $\%OC$ - Mass fraction of soil organic carbon content in percent.

Runoff rate (Q_t) at each model timestep and total precipitation depth (P_t) for each high risk cell are obtained from equation (4.2) and (4.3) and from rainfall data. The use of parameter K_{oc} in equation (4.17) above has some limitations as it generally refers to sorption coefficient of pesticides into soil organic matrix and does not take into account adsorptions to clay particles, which is present in the study area. However, metaldehydes' solubility and low K_{oc} value mean that this limitation is likely to have negligible impact on model outputs as peak metaldehyde concentrations are likely to be mainly due to metaldehyde transport in dissolved form.

The amount of metaldehyde available at soil surfaces during a rainfall event, which is determined by applications and the number of days between applications and a rainfall event, has significant impact on the overall wash-off load that dissolves in runoff. However, lack of data on the specific timing of metaldehyde application makes this difficult to determine. Consequently, build-up and wash-off rate parameters are difficult to be inferred from direct measurements in the catchment and are known to commonly introduce significant uncertainties in pollutant prediction models (Wijesiri et al. 2016). To account for these uncertainties an additional parameter (K), which depends on initial metaldehyde concentrations C_o in the river at the outlet of the catchment prior to rainfall events, was used in the model. The metaldehyde concentration trend in the river prior to a rainfall event provides a general indication of the level of metaldehyde application in the catchment during a particular pesticide application period (Ryberg & Gilliom 2015). Consequently, the trend is therefore used in this study to adjust computations of metaldehyde load in runoff based on measured metaldehyde presence in the catchment.

Hence, metaldehyde load in runoff from each high risk cell at each timestep is determined by

$$M_t = KL_tB \tag{4.18}$$

where: M_t - metaldehyde load in runoff at timestep t(g), $K = C_o * K_{bw}$, C_o is metaldehyde concentration in the river prior to each rainfall event (µg/l), K_{bw} is a calibration parameter (l/µg), L_t - Percentage of application dose that is present in runoff water as a dissolved substance, B - metaldehyde build-up on soil surface through applications (taken as 0.19g per $5m^2$ based on typical application of 5kg/ha using 1.5% slug pellet).

4.2.6 Model Integration

For a given rainfall event over the catchment, rate of runoff generation and travel times are computed using equation (4.3) and equations (4.15) and (4.16). The calculated travel time from each high risk cell is then used to route metaldehyde load to the outlet of the catchment. Time series of runoff (m^3/s) and metaldehyde load in runoff (g) can then be used to determine metaldehyde concentrations in runoff water arriving at the outlet of the catchment. Since metaldehyde transport in ground water is not included in the modelling structure, a measured metaldehyde concentration in the river prior to a rainfall event is used to indicate base flow concentration. Metaldehyde concentrations in base flow (C_o) during the storm runoff period is assumed to be constant whereas a constant slope method is used to increase the amount of base flow (Q_b) over the runoff period (Blume et al. 2007). These are then combined with time series of simulated concentrations in runoff and quantity of runoff water to determine total metaldehyde concentrations in the river. Accurate estimation of the arrival time of peak metaldehyde concentration at the abstraction site is important in terms of enabling smarter surface water abstraction management to avoid peak metaldehyde concentrations. Thus, time to peak (ΔT), prediction error of peak flow (ΔPF) and concentration (ΔPC) are used to evaluate the model

performance along with other commonly used criteria as shown later in sections 4.4.1 and 4.4.2.

Figure 4.5 shows runoff travel time from 2015 high risk areas computed based on a constant and uniform rainfall intensity of 1mm/hr applied for 1 hour over the whole catchment. The map of travel times from high risk areas in the catchment presented in Figure 4.5 shows that the travel time is mainly dependant on the distance of high risk areas both from the outlet of the catchment and from near by streams. Different high risk areas produce runoff depending on the rainfall amount in their local area and hence the proportion of high risk areas contributing runoff at different intervals following a rainfall event depends both on the travel time and the nature of localized rainfall events. The sum of histograms in Figure 4.5 is found to be 74.5km^2 and is in agreement with the sum of the total high risk areas in the catchment (74.5km^2), which has occurred due to the assumption of uniform rainfall over the whole catchment.

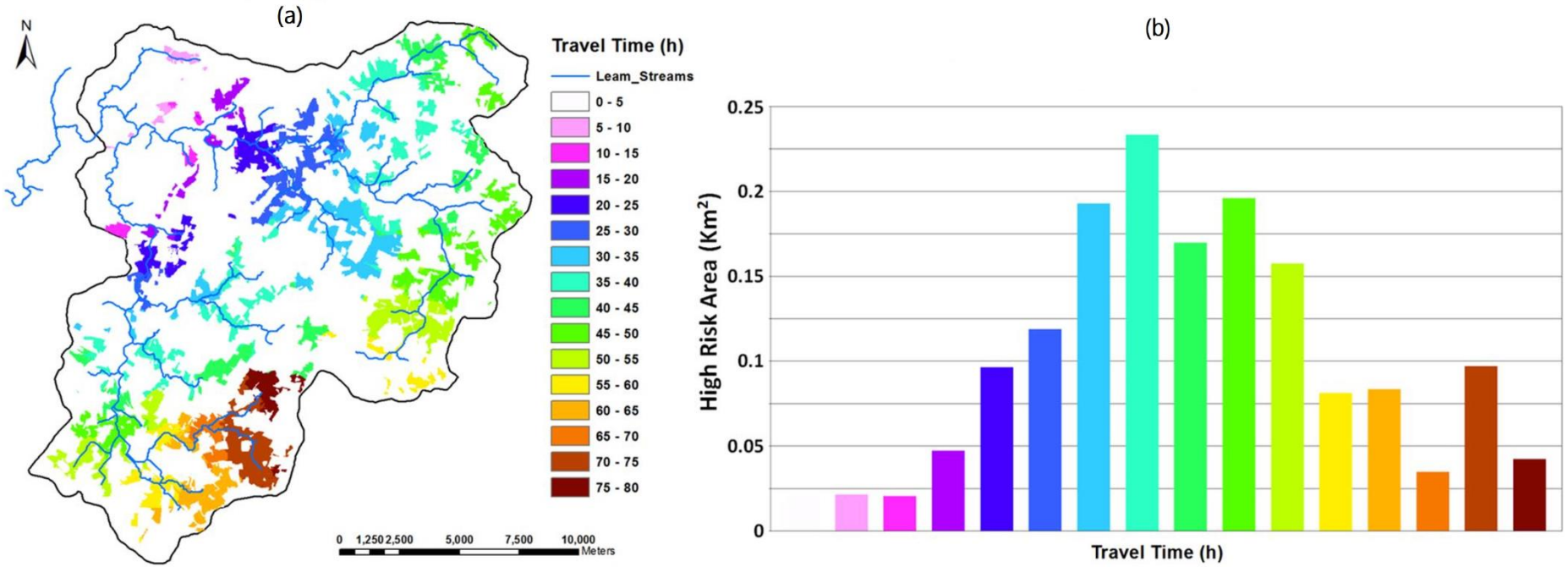


Figure 4.5. Runoff travel time and runoff contributing areas from 2014 high risk areas in the catchment based on a constant and spatially uniform 1mm/hr rainfall of one hour duration. (a) Map of travel time from high risk areas (b) Runoff contributing metaldehyde high risk areas in every 5 hours intervals following rainfall event.

4.3 Model input, calibration and verification data

4.3.1 Land use, soil type and DEM

Land use, soil type and DEM of the catchment were pre-processed to derive various spatial input datasets to the model. Direct model inputs derived from these data are land slope, flow direction, flow accumulation, length of flow pathways, Manning's coefficients (n), curve numbers (CN) and high risk areas. A vector layer of land use, which was derived from satellite imagery, was obtained from the Centre for Ecology and Hydrology, UK for each study year. The land use map classifies crop types and grassland at field level and was used to assign metaldehyde high risk areas (section 4.2.5) as well as Manning's roughness coefficient (n) values for each grid cells based on values published in the literature (Montes 1998; Brater & King 1976). Manning's roughness values assigned for overland surfaces varied between 0.06 – 0.15 whereas roughness values assigned for channel surfaces (based on the nature of the channels) varied between 0.035 - 0.04. The spatially distributed Manning's coefficient values and high risk areas were changed for each study year based on changes in land use in the catchment. The soil map for the study catchment was obtained from the UK National Soil Resources Institute (NSRI) database (NSRI 2009) for the calculation of curve numbers (see section 4.2.3). Soils in the catchment were categorized into four hydrologic soil groups (A, B, C, and D) based on the soil's runoff generating potential (USDA 1986). Hydrologic soil group A generally have the lowest runoff potential and group D have the highest potential. Hydrologic parameters for the calculation of runoff such as slope, flow direction, flow accumulation, drainage basin and stream network delineation were derived in ArcGIS using the OS Terrain 5 digital elevation model, which was obtained from the Ordnance Survey, UK (Ordnance survey 2017).

4.3.2 Rainfall

Radar rainfall data were acquired from the UK met-office's NIMROD system with spatial and temporal resolution of $1km^2$ and 5 minutes respectively (Met Office 2003). The radar rainfall data were resampled to a $5m^2$ grid and aggregated to one

hour resolution to match with the model grid and time resolution. This dataset was used as input for the calculation of runoff generation and pollutant wash-off (sections 4.2.3 and 4.2.5). Initially four rainfall events in the catchment were selected to calibrate and validate the travel time based runoff model developed in this study. Summary statistics and temporally averaged spatial variation of each rainfall event are provided in Table 4.1 and Figure 4.6. The temporal variations of each rainfall event are presented in Figure 4.8. Significant rainfall events with temporal and spatial average rainfall intensities ranging from 0.5mm/hr to 1.5mm/hr and durations ranging from 10hr to 30hr were selected to represent rainfall conditions that are likely to cause metaldehyde spikes at the outlet of the catchment. The peaks of the rainfall events used for runoff model calibration and validation varies widely between 1.5mm/hr and 5.4mm/hr. The spatial distribution of temporally averaged rainfall for the rainfall events used in runoff model calibration and validation presented in Figure 4.6 show the relatively wide variation of rainfall occurrences across the catchment area. Historical radar rainfall data was used to compute antecedent soil moisture conditions for each grid cell over the duration of rainfall events and were used to adjust grid cell curve number values.

Table 4.1. Summary statistics of rainfall events used for runoff model calibration and validation

Rainfall Event No.	**Rainfall Event Date**	**Duration (hr)**	**Temporal and Spatial Average Rainfall Intensity (mm/hr)**	**Temporal and Spatial Peak Rainfall Intensity (mm/hr)**
A1	October 28, 2013	21	1.1	4.5
A2	November 3, 2012	30	0.6	4.3
A3	September 24, 2012	10	1.5	5.4
A4	November 22, 2014	23	0.5	1.5

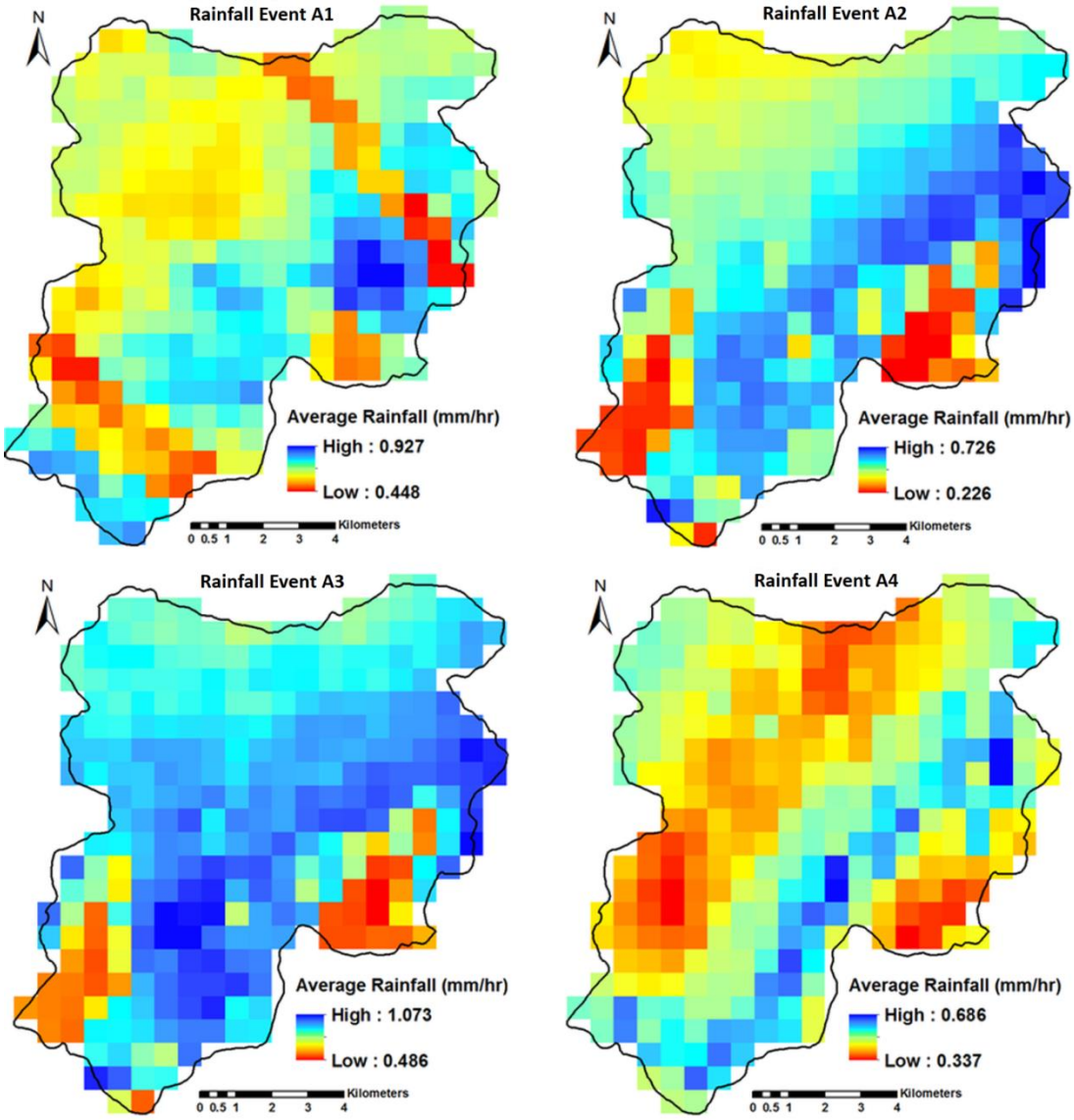


Figure 4.6. Spatial distribution of temporally averaged rainfall for the rainfall events used in runoff model calibration and validation.

Following the validation of the runoff model, radar rainfall data observed during the four metaldehyde data collection events were used to drive the metaldehyde prediction model simulations. Summary statistics and temporally averaged spatial variation of each rainfall event used for calibration and validation of the metaldehyde prediction model are provided in Table 4.2 and Figure 4.7 below. The temporal variations of each rainfall event are presented in Figure 4.9. Rainfall events with temporal and spatial average rainfall intensities ranging from 0.2mm/hr to 0.81mm/hr

and peak rainfall intensities ranging from 1.21mm/hr to 3.3mm/hr (Table 4.2) are used for the metaldehyde model calibration and validation. These rainfall events are selected based on the timings of the four metaldehyde data collection events. As shown in Figure 4.7, relatively small spatial variations across the catchment is observed, which is mainly due to temporal averaging over the longer durations and characteristics of the rainfall events.

Table 4.2. Summary statistics of rainfall events used for metaldehyde model calibration and validation

Event No.	**Event Start Date**	**Duration (hr)**	**Temporal and Spatial Average Rainfall Intensity (mm/hr)**	**Temporal and Spatial Peak Rainfall Intensity (mm/hr)**
B1	October 8 - 9, 2014	34	0.2	2.21
B2	December 12 - 13, 2015	35	0.38	1.21
B3	February 6, 2017	9	0.81	1.73
B4	November 21 - 22, 2016	35	0.55	3.3

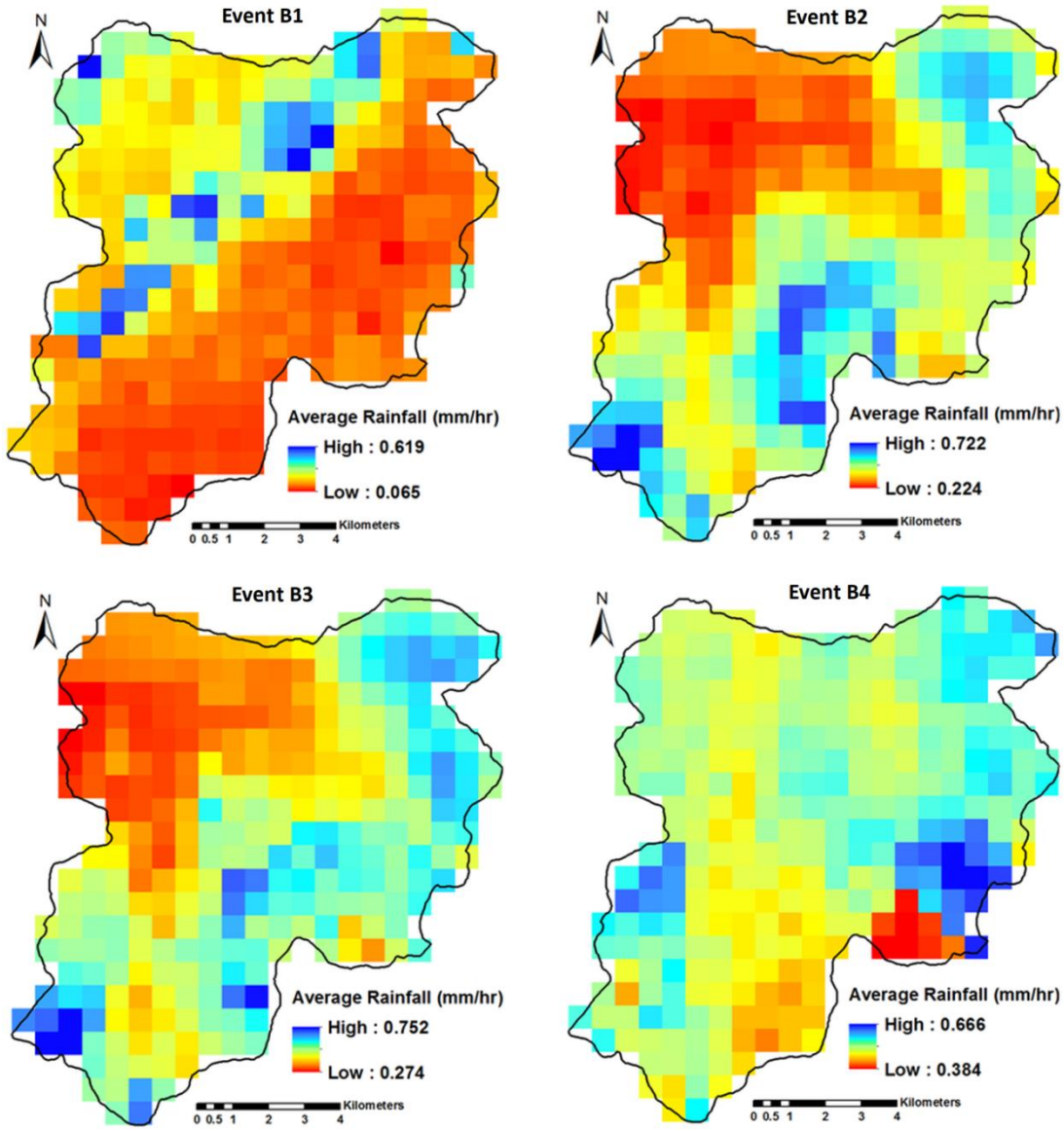


Figure 4.7. Spatial distribution of temporally averaged rainfall for the events used in metaldehyde model calibration and validation.

4.3.3 Flow

Historical hourly flow data from a flow gauging station situated at the outlet of the catchment was obtained from the UK Environment Agency. The flow hydrographs for each rainfall events were separated into base flow and direct runoff using straight line method (Reddy 2006). A straight line is drawn from the point where the sharp rise in hydrograph occurs to the end of recession limb, which is used to separate the

hydrograph into two distinct components: a fast intermittent runoff response and a slow continuous base flow response of the catchment. The fast response runoff hydrographs resulting from the selected rainfall events were used to calibrate and validate the runoff model.

4.3.4 Water sampling and metaldehyde data

We have collected water samples from river Leam using auto-samplers installed at surface water abstraction site used for drinking water supply. The use of auto-samplers enabled the continuous collection of hourly water samples during storm runoff events, which successfully captured the short term fluctuations of metaldehyde concentrations at the abstraction site. The auto-samplers were manually triggered before the arrival of forecasted rainfall events, which were judged likely to cause metaldehyde peaks due to runoff. For each event sampling was carried out for a period of 3 - 5 days, which enabled the acquisition of water samples during the full runoff period following the rainfall events. The data collection campaign was carried out over a period of three metaldehyde application seasons between September 2014 and February 2017. Collected water samples were analysed by Severn Trent Water Ltd laboratory to determine metaldehyde concentrations. Details on the metaldehyde detection method used are provided by Li et al. (2010).

4.4 Results and Discussion

This section presents the calibration and verification results of runoff and metaldehyde concentration prediction models for the rainfall events presented in Table 4.1 and Table 4.2. Comparison of simulated model results with measured flow data at the catchment outlet and metaldehyde concentration data from four water quality sampling events are discussed using various error statistics.

4.4.1 Runoff model

The accuracy of metaldehyde prediction model is dependent on runoff travel times from high risk areas to the outlet of the catchment. Thus, the runoff model, which consists of runoff generation and runoff routing components, needs to be calibrated

and validated before it is integrated to the pollutant build-up/wash-off model. Flow data recorded by a gauging station located at the outlet of the catchment is acquired from Environment Agency and is used to calibrate and validate the travel time computation technique used in the runoff model. Runoff generation and transport from the entire catchment is considered for the calibration and verification of runoff computation approach. Observed flow data from rainfall event A1 was used to calibrate parameters K_o and K_c (equation (4.15) and (4.16)), which were used in the computation of travel times in over land and channel flow cells respectively. Simulation of the runoff prediction model was carried out using eleven different combinations of K_o and K_c values (Table 4.3). The accuracy of the runoff model was evaluated using the prediction error of peak flow rate (*ΔPF*), prediction error of time to peak (*ΔT*) and volume conservation index (*VCI*), which was calculated using equation (4.19). In addition, the overall model prediction efficiency over the entire hydrograph was evaluated using model efficiency coefficient (*E*) as shown in equation (4.20). Prediction error of peak flow rate (*ΔPF*) is defined here as the difference between observed and simulated peak flows, whereas prediction error of time to peak (*ΔT*) is defined as the time difference between the arrival of simulated and observed peak flows.

$$VCI = \sum_{t=1}^{T} Q_m^t \Big/ \sum_{t=1}^{T} Q_o^t \tag{4.19}$$

$$E = 1 - \frac{\sum_{t=1}^{T}(Q_m^t - Q_o^t)^2}{\sum_{t=1}^{T}(Q_o^t - \overline{Q_o})^2} \tag{4.20}$$

where Q_m^t is predicted flow at discrete times t (m^3/s), Q_o^t is observed flow at discrete times t (m^3/s) and $\overline{Q_o}$ is mean of observed flow values over the entire period (m^3/s).

The runoff model prediction results and error statistics for rainfall event A1, which was used for model calibration, are summarized in Table 4.3. The volume conservation index (VCI) for rainfall event A1 is found to be 0.87. The results

indicated that $K_c = 1$ and $K_o = 0.8$ provide the optimum solution considering all the four evaluation criteria. The calibrated parameter value of K_c = 1 shows that the Manning's roughness coefficient values assigned to channels based on values from literature and other parameters used to compute channel travel time required no adjustment. Overall, the calibration results showed that model performance in predicting runoff is more sensitive to the computation of channel travel time than overland travel time. This is mainly due to the longer flow path that runoff travels along channels as compared to relatively much shorter flow paths of overland flows that span from runoff generating areas to the nearby water bodies. Moreover, the faster rate of flow in channels as compared to much slower flows in overland flows also mean that changes in channel travel time would have a much bigger impact on the arrival of runoff at the catchment outlet.

Table 4.3. Error statistics for rainfall event A1 with different values of K_c and K_o

K_c	K_o	ΔPF (m^3/s)	ΔT (h)	E
0.8	0.8	1.16	-7	0.67
0.9	0.8	1.07	-3	0.86
1	0.8	0.98	1	0.85
1.1	0.8	0.91	4	0.69
1.2	0.8	0.85	8	0.47
0.8	1	1.14	-6	0.72
1	1	0.96	1	0.83
1.2	1	0.83	8	0.43
0.8	1.2	1.11	-6	0.77
1	1.2	0.94	1	0.82
1.2	1.2	0.82	8	0.39

These calibrated parameter values were used in the runoff model simulations for the three remaining rainfall events. Table 4.4 summarizes the results of model simulation

and error statistics for the three rainfall events used for runoff model validation. It was observed that model simulations of all three rainfall events have efficiencies greater than 0.80 and prediction error of peak flow rate less than 10%. In addition, volume conservation index of more than 80% and time to peak error of less than 6 hours have been observed for all rainfall events. With an average efficiency of 0.87 for the rainfall events used for validation, the overall performance of the calibrated travel time based runoff model can be considered reasonable. The runoff model performed better for rainfall events with higher AMC as compared to rainfall events with low AMC. Comparison of observed and simulated runoff hydrographs for all four rainfall events are shown in Figure 4.8. The spatially averaged hourly rainfall data presented in Figure 4.8 show that various levels of rainfall intensity and durations over the four rainfall events, which are observed to cause different impacts on the runoff hydrograph characteristics. High intensity and short duration rainfall event is observed to cause large volume of runoff with significant peak as shown in Figure 4.8b, whereas low intensity and long duration rainfall is observed to cause smaller volume of runoff and peak. However, it is also necessary to note that antecedent soil moisture conditions play an important role in the rate of runoff generation and thus can significantly impact on the level of peak runoff arriving at the catchment outlet. In general, the levels of error statistics observed are practically acceptable and predicted runoff hydrographs agree well with the simulated hydrographs. Consequently, the calibrated travel time approach can be used for estimation of metaldehyde transport from high risk areas in the catchment.

Table 4.4. Model simulation results for three rainfall events

Rainfall Event no.	**VCI**	**Peak flow (m^3/s)**	**ΔPF (m^3/s)**	**Time to peak (h)**	**ΔT (h)**	**E**
A2	0.98	24.5	-2.07	45	5	0.91
A3	0.99	8.0	-0.15	38	4	0.83
A4	0.82	7.3	0.5	63	5	0.86

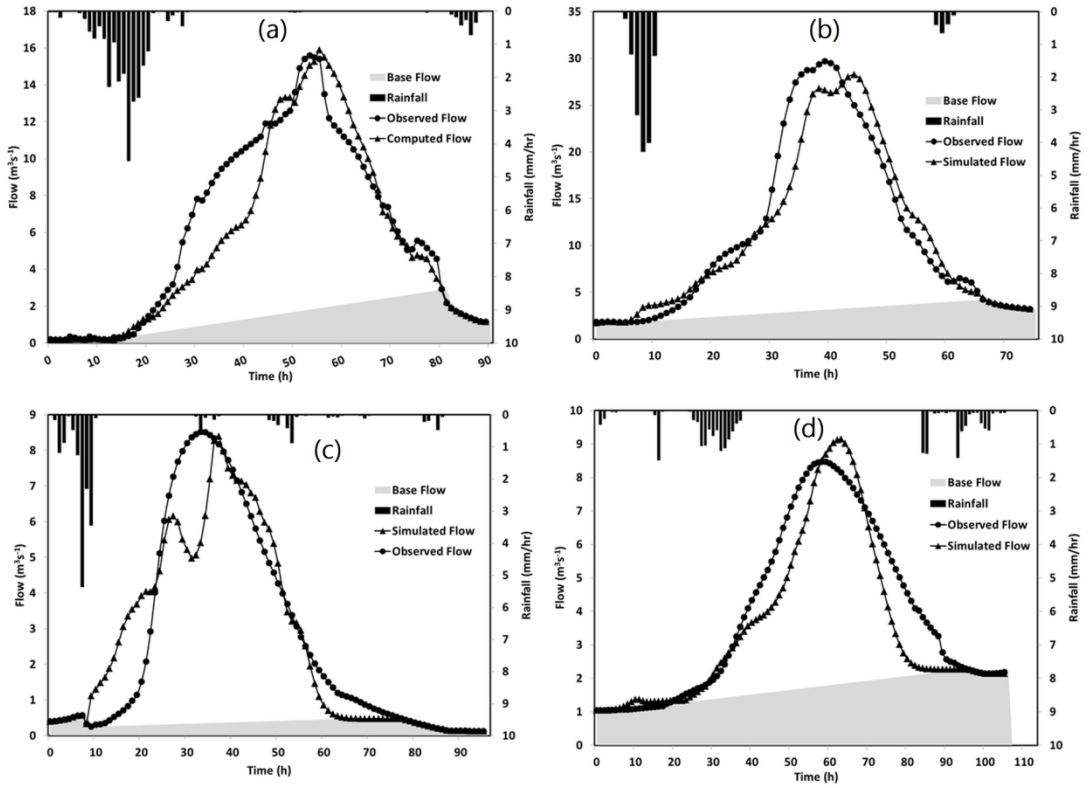


Figure 4.8. Comparison of observed and simulated runoff hydrographs and spatially averaged rainfall over the catchment (T = 0 at start of recorded rainfall). (a) Rainfall event A1 that is used for model calibration (b) Rainfall event A2 (c) Rainfall event A3 (d) Rainfall event A4.

4.4.2 Metaldehyde prediction model

The rainfall event based operation of the automatic samplers to collect hourly water samples enabled the capture of high resolution metaldehyde concentrations arriving at the outlet of the catchment following rainfall events. Results of the analysis of metaldehyde concentrations from the collected water quality samples for each event are presented in Figure 4.9. The analysis shows that relatively short lived metaldehyde peaks with event durations ranging from 12 to 48 hours occur following rainfall events (Figure 4.9). The size and nature of these short lived metaldehyde spikes are highly variable between events. For example, recorded metaldehyde concentrations rise by approximately 500% during event B2, but only by 150%

during event B3, however averaged rainfall is of the same order of magnitude for both events. The datasets therefore emphasise that runoff generation from high risk areas has a significant impact on metaldehyde concentrations in the catchment surface waters, and that pollutant dynamics is highly sensitive to temporal and spatial distributions of rainfall and land use. Moreover, soil type on the land where metaldehyde is applied combined with chemical characteristics of metaldehyde such as solubility and sorption coefficient play an important role in the process of mobilizing metaldehyde into water courses.

The metaldehyde concentration prediction model represents metaldehyde transport in runoff from high risk areas in the catchment by coupling the travel time technique calibrated in section 4.4.1 with build-up/wash-off component. This enabled forecasting of metaldehyde concentration levels following rainfall events at the outlet of the catchment, where the surface water abstraction site is located. Metaldehyde concentration data collected over data collection event B1 was used to calibrate the value of parameter K_b, which was used to account for uncertainties associated with the estimation of metaldehyde build-up and wash-off rate. Different values of parameter K_b ranging from 1 to 3.5 were set in the metaldehyde prediction model to simulate metaldehyde concentrations during data collection event one. The model performance was evaluated using four criteria i.e. prediction error of time to peak concentration (ΔTc), prediction error of peak metaldehyde concentration (ΔPC), coefficient of determination (R) of observed and simulated metaldehyde concentrations and model prediction efficiency (E). However, due to the assumption of uniform application of metaldehyde on all high risk areas (section 4.2.5), changes in parameter K_{bw} result in an overall proportional increase or decrease of predicted metaldehyde concentrations across the prediction period, hence calibration has no impact on the proportion of the variances between predicted and observed concentrations. As a result, coefficient of determination (R) values between predicted and observed concentrations are found to be insensitive to changes in parameter K_{bw}. The metaldehyde prediction model results for data collection event B1 and associated error statistics are summarized in Table 4.5. The results indicated that optimum

solution is attained with $K_b = 1.6$ considering the remaining criteria for data collection event B1. Initial concentration (C_o), which represents metaldehyde concentration in the river prior to each rainfall event, value of 0.067µg/l is used for the calibration event B1. C_o values for each event used for metaldehyde model validation are presented in Table 4.6.

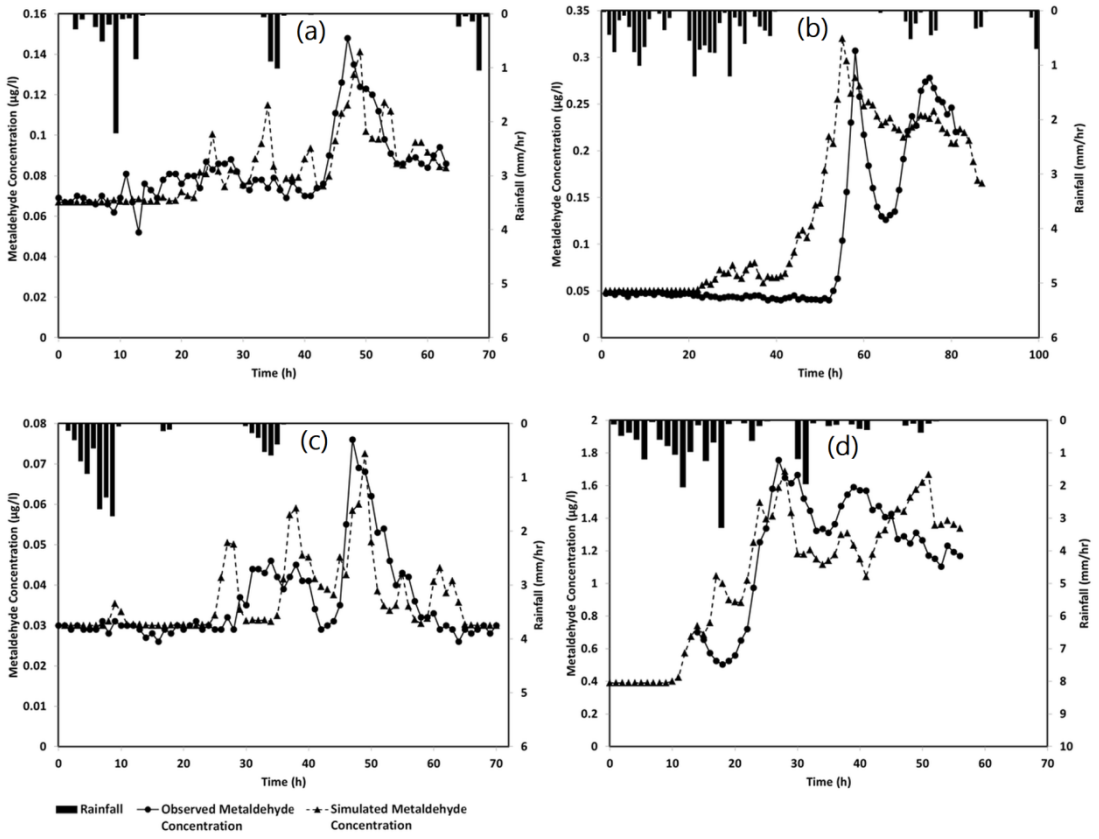


Figure 4.9. Spatially averaged rainfall and comparison of observed and simulated metaldehyde concentrations at the catchment outlet for events B1 - B4. (a) Rainfall event B1 that is used for model calibration, (b) Rainfall event B2, (c) Rainfall event B3, (d) Rainfall event B4.

Table 4.5. Error statistics for data collection event B1 with different values of K_b

K_{bw}	ΔTc (h)	Peak Metaldehyde Concentration (µg/l)	ΔPC (µg/l)	R	E
1	2	0.11	-0.04	0.77	0.10
1.3	2	0.12	-0.03	0.77	0.42
1.5	2	0.13	-0.01	0.77	0.55
1.6	2	0.14	-0.01	0.77	0.60
1.7	2	0.14	-0.01	0.77	0.54
2	2	0.15	0.01	0.77	0.47
2.5	2	0.16	0.01	0.77	0.42
3.5	2	0.20	0.05	0.77	-0.77

Verification

Metaldehyde model simulations were carried out for other three metaldehyde data collection events using calibrated parameter values. Table 4.6 summarizes model simulation results and error statistics for all three data collection events. It was observed that simulation for all three events have correlation coefficient of 0.70 or more, prediction error of peak metaldehyde concentration less than 5% and time to peak concentration error of 6 or less hours. Observed and predicted metaldehyde concentrations are shown in Figure 4.9 for all four data collection events. Figure 4.9 also shows the spatially averaged rainfall over the catchment, which is observed to have various characteristics ranging from high intensity and short duration rainfall events to low intensity and longer duration rainfall events. As shown in Figure 4.9, high intensity with short durations of rainfall and low intensity with longer durations of rainfall have been observed to cause different levels of metaldehyde peaks. Low intensity and long duration rainfall events are observed to cause high peak

metaldehyde concentrations as shown in , whereas the high intensity and short duration rainfall event in has resulted in relatively smaller peak metaldehyde concentration. But, in addition to the rainfall characteristics, the important role antecedent moisture conditions play in metaldehyde generation and routing needs to be noted as discussed in section 4.4.1. In general, metaldehyde concentrations are predicted well for all events with practically acceptable levels of errors in terms of both concentration levels and prediction of peak arrival times. The results showed the capability of the model developed in this study for the intended practical purpose of predicting the arrival of peak metaldehyde concentrations and informing surface water abstractions. Discrepancies in the prediction of the peak arrival time are likely to be caused mainly by uncertainties associated with estimation of channel travel time, antecedent conditions and the assumption of uniform metaldehyde application throughout the high risk areas in the catchment. Some of these errors may be reduced in future via the use of more calibration data and a more detailed consideration of metaldehyde applications informed by data from farmers (i.e. real time application data).

Table 4.6. Simulation results for three data collection events

Data Collection Event no.	**C_o (µg/l)**	***ΔTc* (h)**	**Peak Metaldehyde Concentration (µg/l)**	***ΔPC* (µg/l)**	**R**	**E**
B2	0.05	-3	0.32	0.01	0.81	0.45
B3	0.03	2	0.07	-0.003	0.7	0.48
B4	0.4	6	1.7	-0.06	0.74	0.45

4.5 Conclusions

Diffuse agricultural pollution is known to be a significant concern to the quality of surface water, with implications for drinking water supply. Smarter management of water resources including forecasting and prediction of pollutant spikes is a possible means to avoid contamination of drinking water supplies and reduce the cost of water treatment. This requires a detailed understanding of pollutant processes in the catchment in response to rainfall events. The occurrence, sources, transport and fate of organic compounds in the environment involve a variety of processes that determine how the compounds are initially distributed, move and react. Consequently, assessing fate and transport of contaminants in the environment is a complex issue. This chapter focuses on predicting the arrival of peak metaldehyde concentrations in runoff at abstraction sites with a view to inform surface water abstraction decisions, hence a model has been developed to describe short term dynamics and transport, primarily driven by rainfall driven runoff, rather than longer term reactions/degradation or groundwater processes. Runoff generation and routing is spatially and temporally variable and hence surface water quality responses are dependent on the spatial distribution of pesticide within the catchment (a function of land use) and the dynamics of individual rainfall events. To date the quantification and understanding of the pollutant dynamics that drive short term fluctuations has been hindered by a paucity of high resolution water quality sampling data. The physically-based distributed metaldehyde prediction approach developed in this chapter combines runoff and build-up wash-off concepts in a GIS environment, enabling the full consideration of spatially and temporally variable rainfall and land use patterns. Model parameters and input data are extracted from radar rainfall data, soil type, land use and DEMs. To address the paucity of current data we attempt to utilize automatic samplers which were triggered during rainfall events to capture the impact of forecasted rainfall events on the concentrations in surface waters. The variation in the metaldehyde concentration response between the rainfall events demonstrates the importance of a full consideration of spatiotemporal rainfall and metaldehyde application data.

In terms of practical application, it is noted that the accurate forecasting of arrival time of peaks is of more value than forecasting of the peak concentration value, as this enables surface water abstraction decision makings such as suspending abstractions temporarily in order to avoid the entrance of high metaldehyde levels into water supply systems. Given the inability of existing treatment techniques to remove high metaldehyde levels from water and the absence of direct metaldehyde detection methods, the model developed in this study provides a cost-effective and sustainable solution. When applied to the trial catchment the model was able to predict peak concentrations to within 6 hours in all cases, given the availability of water storage infrastructure in the catchment this would enable the operator to suspend abstraction for this period to allow likely periods of high concentration to pass. Given the effective utilisation of storage, such a suspension would not have a significant negative impact on water resources, especially if abstraction was increased at other times to compensate. The increasing availability of catchment scale spatial datasets combined with the relatively simple GIS based application of the model makes it suitable for use in various catchments, where prediction of metaldehyde exposures are required.

Chapter 5

Uncertainty in Metaldehyde Prediction Model

5.1 Introduction

The reliability of water quality management decisions and the associated benefits which accrue from them are dependent on the accuracy of water quality prediction models used to inform them. Calibrated parameter values and inputs are often used in deterministic water quality models without considering uncertainties and overlooking implications of this on decisions that are made based on water quality model outputs. However, uncertainties in water quality predictions can be very high due to the non-linearity and multi-dimensional nature of the process represented by models. A measure of the significance of these uncertainties in water quality model outputs is necessary to enable to inform the level of confidence in these models (Viviani 2009). Particularly, lack of measured data on spatially variable inputs and parameters are known to cause significant level of uncertainties in catchment scale water quality models. Thus, reliable water quality management decisions that aim to generate positive social, environmental and economic benefits have to carefully consider information on uncertainties of water quality models. Uncertainties in water quality predictions originate mainly from errors in model structures, inputs and parameters that inaccurately represent the various processes involved in pollutant generation and transport (Beven & Freer 2001). Numerical methods used to solve model equations and observed water quality data used for model calibration are also other factors that bring uncertainties in water quality predictions. Parameters are used in models to represent various components of hydrological processes represented by the model. Some of the parameters used in models have direct physical interpretations and can be measured, whereas values of some model parameters have to be estimated using various methods (Beven 2001). However, estimation of parameter values should preferably provide information on their uncertainties in addition to optimum parameter values (Briggs et al. 2012). Similarly, uncertainty analysis needs to be incorporated when estimating model inputs. Water quality predictions that incorporate information on the level of model uncertainties enable to make risk-

aware water quality management decisions (Hall & Borgomeo 2013). Moreover, uncertainty quantification and sensitivity analysis can help in understanding contributions of the various sources of uncertainty, which will be instrumental in planning uncertainty reduction efforts (Viviani 2009).

In this chapter, uncertainty propagation analysis of the metaldehyde prediction model, which is introduced in chapter 4, is performed with a view to enabling well informed and risk aware decisions when using the model outputs. The uncertainty analysis outputs can also potentially used to inform catchment scale data collection strategies. Monte Carlo methods are commonly used to quantify uncertainties in diffuse pollution models and analyse sources of model output uncertainties. Monte Carlo method's general applicability, easy implementation and ability to represent model output uncertainties using probability distributions make it attractive for use in uncertainty quantification and analysis (McIntyre et al. 2004). A state-of-the-art Monte Carlo based spatial uncertainty analysis tool is employed to enable efficient sampling of spatially variable model inputs and parameters. The numerically simple and physically distributed structure of the metaldehyde model combined with the availability of catchment scale high spatiotemporal resolution data have enabled the application of spatial uncertainty analysis of the metaldehyde model. A study that investigates spatial uncertainties associated with catchment scale water quality models is currently lacking in the literature. The uncertainty analysis in this chapter is focused on input and parameter uncertainties. Quantification of model structural uncertainty requires comparison of performance of different modelling tools and is not the subject of this chapter.

5.2 Methodology

Following calibration and validation of the metaldehyde prediction model (chapter 4), uncertainty quantification and analysis are performed to assess propagation of model inputs and parameters uncertainties to model outputs. Stochastic and deterministic sensitivity analyses are carried out to analyze the sensitivity of the model output to different model input and parameter uncertainties. This enabled to

identify contributions of individual sources of uncertainty to the output uncertainty. The following five major steps are followed when performing the uncertainty quantification and analysis.

1. Select model inputs and parameters to be included in uncertainty quantification
2. Define and parametrize probability distribution functions of input and parameter uncertainties
3. Generate model inputs and parameters realizations from predefined probability distributions to represent uncertainties in inputs and parameters
4. Run metaldehyde prediction model for each model input and parameter realization to propagate uncertainties through the model
5. Summarize model output uncertainties resulted from Monte Carlo runs using various measures

The metaldehyde prediction model is a catchment scale physically distributed model and thus involves large number of inputs and parameters. Consequently, there are likely to be various sources of uncertainty in the model. But, only few inputs and parameters are believed to cause larger uncertainties and are considered in the uncertainty analysis in this chapter. This may lead to underestimation of the quantified model output uncertainty as there are likely to be uncertainty contributions from other inputs and parameters to the overall uncertainty. However, inclusion of all inputs and parameters in the uncertainty analysis will make the process more complex and time consuming, while the extra quantified uncertainty is likely to be small as compared to the uncertainties caused from the major sources. Thus, information from literatures on main sources of uncertainty in stormwater quality models (chapter 2) and existing knowledge on the underlying techniques used to develop the metaldehyde model are used to select few inputs and parameters for consideration in the uncertainty assessment in this chapter. Results from the runoff prediction component of the metaldehyde model (discussed in section 4.4.1) have shown that only small uncertainties were originated from the rainfall input data. Thus, input uncertainties arising from radar rainfall data are not considered in this

study. As a result, metaldehyde applied high risk areas (*B*), time interval between applications and a rainfall event parameter *t*, build-up/ wash-off calibration parameter K_{bw}, channel flow travel time calibration parameter K_{ch} and overland flow travel time calibration parameter K_o are selected and considered in the uncertainty quantification and analysis of the metaldehyde model.

5.2.1 Uncertainty quantification of selected model inputs and parameters

Quantifications and representations of uncertainties of the five selected model inputs and parameters are discussed in this section individually. Depending on the nature of these selected input and parameters, different techniques are employed in this section to represent uncertainties associated with each.

Metaldehyde build-up (*B*): build-up of metaldehyde on high risk areas occurs through slug pellet applications mainly during the autumn and winter seasons. Typical single slug pellet application based on guidelines from manufacturers is 5 kg/ha (Metaldehyde Stewardship Group 2012a). This is equivalent to 75g/ hectare (0.19 g per the 25 square meter grid used in the model) of metaldehyde based on a commonly used 1.5% slug pellet. The statutory legal requirement in the UK on metaldehyde application states that total application in a calendar year shouldn't exceed a maximum of 700g/ha (Metaldehyde Stewardship Group 2012b). Metaldehyde Stewardship Group guidelines promote autumn restriction on metaldehyde applications, which runs from 1 August to 31 December. These restrictions recommend a maximum total metaldehyde application rate of 210g/ha during this period (Metaldehyde Stewardship Group 2012b). However, actual applications vary based on the occurrence and size of slug population during a particular period, which varies from time to time depending on wet conditions on farmlands. Lack of data on actual metaldehyde applications makes it difficult to determine which farmlands received new metaldehyde application before a particular rainfall event. In the deterministic metaldehyde prediction model it was assumed that all high risk farmlands have received new metaldehyde application before a particular rainfall event (section 4.2.5). However, only a proportion of high risk areas

are likely to receive new metaldehyde application at any one time. Metaldehyde build-up input (*B*) is used to represent build-up through applications before a particular rainfall event in the model. To account for uncertainty in estimation of metaldehyde build-up, *B* is considered as uncertain categorical variable, which takes binary values to represent whether metaldehyde is applied or not on a particular land. Hence, categorical variable *B* takes values of either 0.19 or 0, where 0.19 represent the amount of metaldehyde in grams resulted from a typical single application per the 5 square meter grid area used in the model and 0 represents no metaldehyde application on that specific farmland. The overall amount of metaldehyde application in the catchment prior to a particular rainfall event is represented using parameter *K* as discussed in section 4.2.5. Thus, the probability of metaldehyde application on each farm land prior to a specified rainfall event is represented based on parameter *K*. The probability that *B* takes a value of 0.19 or 0 (metaldehyde is applied or not) for a particular land prior to a specified rainfall event thus varies from one event to another depending on values of parameter *K*, which is derived from metaldehyde concentration in the river prior to each rainfall event (C_o) and calibration parameter (K_b). The use of values of parameter *K* to estimate probabilities used in generating realizations from *B* has enabled to reflect the fact that different proportions of farmlands in the catchment receive metaldehyde application at different periods (Figure 5.1).

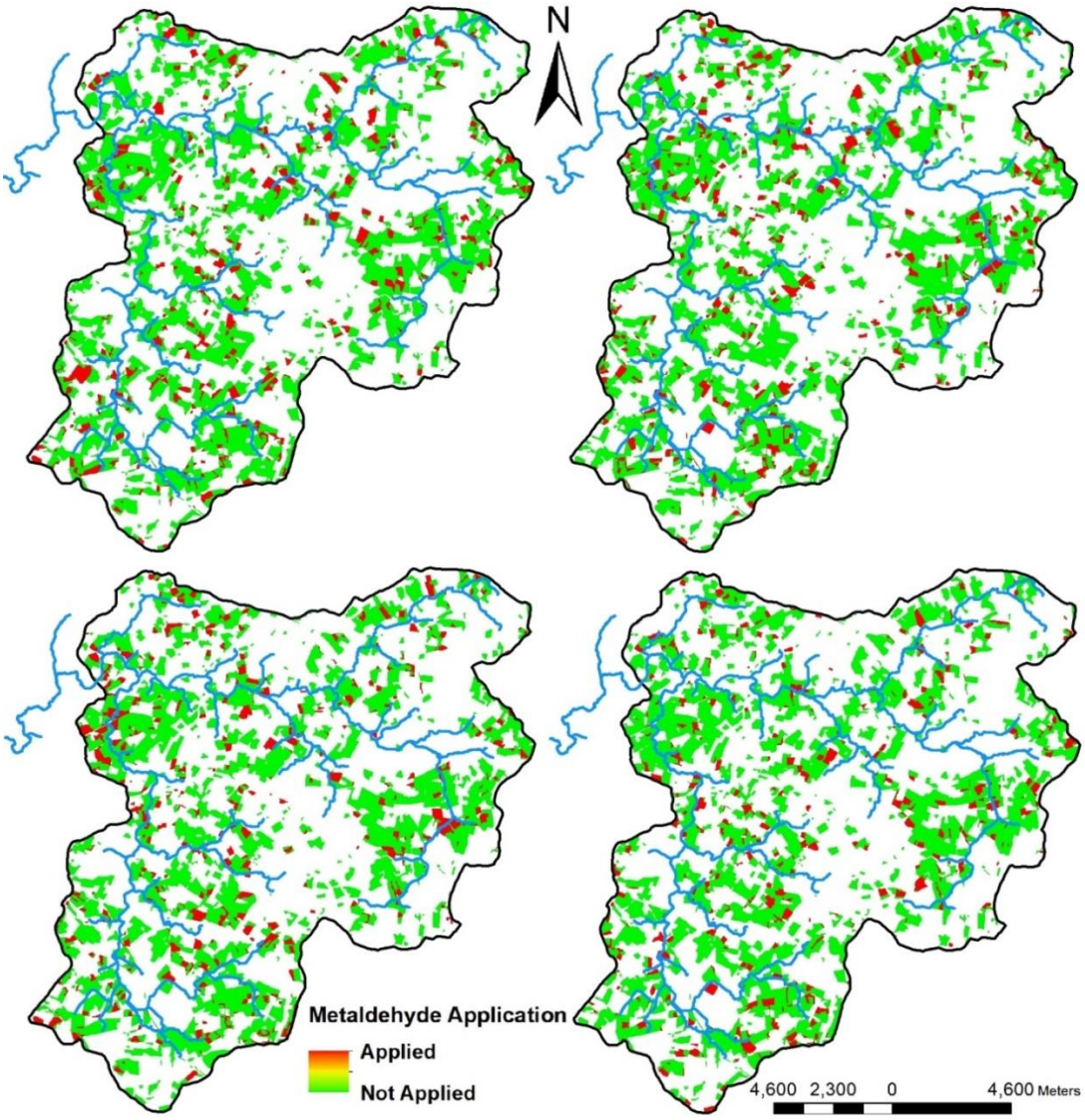


Figure 5.1. Examples of high risk area realizations with new metaldehyde applications.

Time interval between applications and a rainfall event (*t*): Recorded data on time interval parameter *t* (number of days), which is used to express the interval between metaldehyde application and a rainfall event (Figure 5.2), is rarely available and an assumed value of 5 days for all high risk farmlands in the catchment was used in the deterministic model (section 4.2.5). The values of all other parameters used in the wash-off equation were derived from measured variables. As a result, a significant proportion of uncertainty in the wash-off equation is derived from the parameter *t*. For the uncertainty analysis study in this chapter, *t* was considered as a uniformly

distributed variable between 1 and *n*, where *n* is the number of days between two consecutive slug pellet applications on high risk farmlands during the application period. In general, most of metaldehyde applications are made during September to December period as the wet soils during the autumn and winter seasons provide favorable conditions for slugs. Routine monitoring data collected by the local water infrastructure operator shows that more than 90 % of high levels of metaldehyde in the river have occurred during the September to December application season (Figure 4.2, Chapter 4). Accordingly, it is assumed that approximately 90% of the 700g/ha statuary annual legal limit of metaldehyde is applied during the September to December application period. Based on this assumption and considering the typical single slug pellet application amount of 5kg/ha (75g/ha based on a commonly used 1.5% slug pellet), frequency of application on high risk farmlands during this period is estimated to be around 14 days. This indicates that time interval between any new metaldehyde application and a rainfall event, *t*, for a particular farmland ranges between 1 to 14 days.

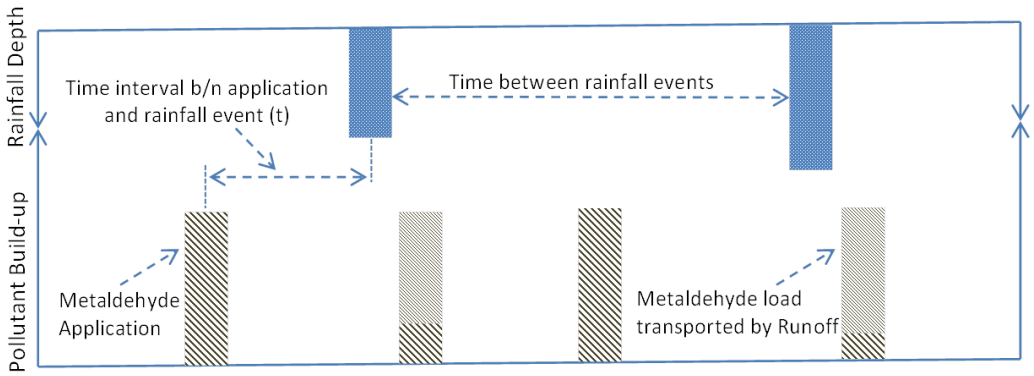


Figure 5.2. Time interval between application and rainfall events.

Build-up/ wash-off calibration parameter (K_{bw}): Parameter K_{bw} is calibrated in the model and used in combination with metaldehyde concentration trends in streams to represent the ratio of actual metaldehyde application in the catchment to the application of metaldehyde on all high risk areas. It is considered as continuous variable normally distributed with mean of the calibrated value of the parameter.

Channel travel time calibration parameter (K_{ch}): Parameter K_{ch} is used to represent uncertainties in channel surface roughness (n) and slopes used in the estimation of runoff travel time in channel networks thorough out the catchment. Uncertainty in estimated channel roughness and the digital elevation model used to generate channel slope directly propagates to parameter K_{ch} as it is used to account for these errors in the model. It is considered as continuous variable normally distributed with mean of the calibrated value of the parameter.

Overland flow calibration parameter (K_o): Overland flow calibration parameter K_o is used in the model to account for travel time estimation errors in overland flows. These errors occur due to uncertainties in DEM and land use data used to drive land slopes and roughness coefficients respectively. Hence, uncertainty in the digital elevation model and land used data used to derive roughness coefficients and land slope propagates to parameter K_{ch} as it is used to account for estimation errors in these parameters. It is considered as continuous variable normally distributed with mean of the calibrated value of the parameter.

5.2.2 Uncertainty representations

Uncertainties in selected model inputs and parameters are represented using probability distribution functions. The estimation of probability distribution function used to characterize the uncertainties depends on the spatiotemporal representations and the measurements scale of the uncertain inputs and parameters. The type and characteristics of probability distribution functions used to quantify each input and parameter analyzed in this study are described below.

Metaldehyde build-up (B) is represented in the model as a spatially variable categorical model input with two categories to represent that metaldehyde is applied or not applied on a particular farm land. The uncertainty in metaldehyde application on any specific farmland in the catchment before a particular rainfall event is characterized by a discrete probability distribution function as

$$P[C(X) = c_i] = \pi_i(X) \tag{5.1}$$

where $P(C(X) = c_i)$ is the univariate probability that variable C falls in category c_i at high risk farmland X, or shortly $\pi_i(X)$. The numerical values that will be used in the model for the two categories for metaldehyde build-up (B) are 0 (metaldehyde not applied) and 0.19 (representing thethe typical metaldehyde application).

Time interval between applications and a rainfall event parameter (t) is considered in this study as a numerical constant variable that doesn't exhibit spatiotemporal variation. Consequently, continuous probability distribution function, which quantifies the probability at which a variable takes a value in any given interval, is used to represent uncertainties in parameter t. The likelihood of metaldehyde application during any dry day between two rainfall events during the application season is reasonably assumed to be equal to the likelihood on any other day. Thus, a uniform probability distribution function is used to represent uncertainties parameter t as follow

$$P(t) = \begin{cases} \dfrac{1}{t_{max} - t_{min}} & for \;\; t_{min} < t < t_{max} \\ 0 & for \;\; t < t_{min} \quad or \quad t > t_{max} \end{cases} \tag{5.2}$$

where $P(t)$ - the univariate probability that metaldehyde is applied t days before a particular rainfall event,t_{min} – the shortest possible time interval (in number of days) between an application and a rainfall event, t_{max} – the longest possible time interval (in number of days) between an application and a rainfall event.

Build-up/ wash-off calibration parameter K_{bw}, channel travel time calibration parameter K_{ch} and overland flow calibration parameter K_o are all numerical constants that have been considered invariable both spatially and temporally. Variations in values of these parameters are mainly due to natural processes and thus the commonly used normal distribution function is used to represent uncertainties in these parameters as follow

$$P(V = v_i|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(v_i-\mu)^2}{2\sigma^2}} \quad (5.3)$$

where $P(V = v_i|\mu, \sigma^2)$ – the univariate probability that the value of parameter V = v_i, μ– base value of parameter V and σ– standard deviation of parameter V.

5.2.3 Uncertainty propagation through the model

Propagation of uncertainty through the model requires conducting a serious of model simulations with samples of inputs and parameters generated from predefined probability distributions. The sampling tool used to generate well representative realizations from the assigned probability distributions to model inputs and parameters depends on the nature of the distribution function. Hence, different tools and techniques are used to sample from the predefined probability distributions. A spatial uncertainty analysis tool called SPUP (spatial uncertainty propagation) (K. Sawicka 2016) is used to generate spatially distributed ensembles based on categorical vaues for B, which is represented as a spatially variable categorical model input as discussed in section 5.2.1. The SPUP package provides various functions to characterize uncertainties and generate realizations from uncertain spatially variable categorical and continuous model variables. Initially, quantified uncertainties of the selected model inputs and parameters are described in the SPUP tool using a function called *defineUM()*, which is given by

For categorical variables

$$defineUM(uncertain, categories, category\ probability) \quad (5.4)$$

For continuous variables

$$defineUM(uncertain, distribution, distribution\ parameters) \quad (5.5)$$

where *uncertain* - logical value (True or False) to indicate if a spatial uncertainty of a variable is considered or not, *categories* – categorical values to be assigned to uncertain variables, *category probability* – probability corresponding to each

categorical value, *distribution* - type of probability distribution function assigned to characterize the uncertain variables, *distribution parameters* – list of parameters used to describe the probability distribution function. E.g. in case of normal distribution a mean and standard deviation of the distribution are required.

A function called *genSample()* is used to generate samples from these predefined probability distribution functions. Various sampling techniques are incorporated to the *genSample()* function to enable sampling from the predefined probability distribution functions accounting for spatial variability and autocorrelation. The *genSample()* function in SPUP is expressed as:

$$genSample(UMobject, N, Sampling\_Method) \tag{5.6}$$

where *UMobject* - quantified uncertainty of model variables defined using the *defineUM()* function as shown in equation (5.4) or (5.5), *N* - number of realizations required, *Sampling_Method* - type of sampling technique to be used.

All model grids (25 square meter grids) in a single field are considered to receive similar amount of metaldehyde application at the same time. This indicates that the spatially variable model input *B*, which represents metaldehyde build-up on high-risk areas through applications, is positively spatially auto-correlated at farmland level. Thus, Monte Carlo sampling for model input *B* is made using vector map that consists all high-risk farmlands in the catchment. This insures that at every sampling run a single value is allocated to each polygon representing a high-risk farmland in the catchment. The vector maps with sampled values are later converted into gridded data to be used in the grid based model. This allowed to assign similar values for all grids in a single farmland and thus to adequately account for the auto-correlation between grids in a single farm land during sampling. However, spatial autocorrelation between any two farmlands is assumed to be zero.

Equation (5.4), which is a SPUP function used to define categorical variables, is first employed to define the metaldehyde application on high risk areas variable, B.

Categories of (0,1) and categorical probability is assigned based on parameter *K* as discussed in section 5.2.1. The spatial based sampling function of the SPUP package, *genSample(),* is then used to generate metaldehyde application on high risk areas ensembles (Figure 5.1). The Latin Hypercube method (McKay et al. 1979), which uses stratified random procedure to provide an efficient way of sampling, is employed to sample from probability distributions of model inputs and parameters. Following the sampling procedure, the model was run for each of model input and parameter realizations to propagate their uncertainties through the model. Due to the long metaldehyde model simulation time, 100 realizations of each uncertainty variable are used to evaluate the model and propagate uncertainties. This has resulted in an equal number of model output values, which in this case are multiple time series values of metaldehyde concentrations for each Monte Carlo run. Overall, five model uncertain variables are considered in the uncertainty analysis in this chapter as discussed in section 5.2.1.

5.2.4 Influence of model input and parameter uncertainties on model output

To study the overall influence of model input and parameter uncertainty on the predicted metaldehyde concentrations, samples were generated from the model input and parameter distributions discussed in section 5.2.2 using a series of forced perturbations. To do this, all previously selected model input and parameters were represented in the Monte Carlo simulations using ensembles generated from the corresponding distributions with five different ranges of probability distributions. This involved varying the characteristics of the distribution by changing its coefficient of variation. The use of coefficient of variation, which is defined as the ratio of standard deviation to the mean and often expressed as percentages, provided a suitable measure for the relative variability of the assigned input and parameter probability distributions. Mean and standard deviation of the input and parameter distributions are varied to result in five different levels of coefficient of variations (2.5%, 5%, 7.5%, 10% and 12.5%), which introduced different levels of dispersion in the input and parameters distributions. This enabled to analyze the overall influence of the selected model input and parameters uncertainties on model outputs and assess

characteristics of the output uncertainty at different levels. Calibrated values of the parameters in the deterministic model (as discussed in chapter 4) were used as base values for perturbation and Monte Carlo simulation.

Uncertainty in predicted metaldehyde concentrations (model output ensembles) are analyzed by computing various summary statistics such as confidence interval (90%) and coefficient of variation. The 90% confidence interval used here to present the results is defined as the difference between the 95^{th} and 5^{th} percentile values under the cumulative distribution curve of the predicted metaldehyde concentration. Both confidence intervals and coefficient of variation are evaluated for each series of perturbations for comparison. Gain factor (G), which is the ratio of predicted metaldehyde concentration coefficient of variation to the input and parameter coefficient of variation, is used to quantity the degree of amplification or attenuation in the perturbation transferred from the model input and parameters to the predicted metaldehyde concentration.

5.2.5 Sensitivity Analysis

Due to its catchment scale physical base and distributed nature, relatively large number of inputs and parameters are employed in the metaldehyde prediction model. Sensitivity analysis was carried out to analyze the sensitivity of the metaldehyde prediction model to selected model inputs and parameters. Quantifying uncertainty contributions from model inputs and parameters helps to identify the weakest link in the model; hence it can be used to prioritize efforts aimed at minimizing uncertainties in model predictions. In this study, three different types of sensitivity analyses were carried out to determine uncertainty contributions from different model inputs and parameters. The first two were performed using stochastic sensitivity analysis approach, which required generating and using realizations from input and parameter distributions in successive simulations. To compare uncertainty contributions from different model inputs and parameters, model output uncertainty from the first two sensitivity analyses are summarized using summary statistics such as confidence interval (CI=90%) and coefficient of variation. The level of uncertainty in model

output resulting from Monte Carlo simulations depends on the variances used in inputs and parameters distributions. Using a fixed value for one of the inputs or parameters enables to remove uncertainty contribution from that particular input or parameter. In the first stochastic sensitivity analysis approach, Monte Carlo simulation was performed initially using a full stochastic model, which considers all inputs and parameters of interest as random variables. Then, the inputs and parameters are kept constant progressively converting inputs or parameters to fixed variables through successive Monte Carlo simulations in a cumulative manner. In the first sensitivity analyses, an order of *Kbw, t, B, Ko, Kch* is used when inputs and parameters are converted to fixed variables. This order is then reversed when performing the second stochastic sensitivity analysis, which otherwise is performed similar to the first sensitivity analysis. This insured the analysis of the relative impact of uncertainty contribution from each input and parameter considered in the uncertainty analysis. In addition, combining the outputs from the two sensitivity analyses help to identify any correlations between model inputs and parameters (Benke et al. 2008).

A third sensitivity analysis is conducted, which involved increment and decrement of each input or parameter from their base values while holding all other inputs and parameters constant. A classical "spider plot" is used to analyze outputs from this deterministic sensitivity approach (Eschenbach 1992), which shows the percentage change in model input and parameters versus percentage change in predicted metaldehyde concentrations. The five different plots included in the spider plot revealed the level of predicted metaldehyde concentration sensitivity to each of the five model inputs and parameters on individual bases.

5.2.6 Assessment of uncertainty contributions

Based on the nature of physical processes involved and the type of mathematical structures used to represent them in the model, uncertainties in different model inputs and parameters are likely to impact on various attributes of the model output. In addition to the magnitude of uncertainty contribution, model inputs and parameters

roles in defining the characteristics of output uncertainty is crucial in making well informed and riskware decisions. Stochastic simulation of the model is carried out individually for each input and parameter while keeping all others constant to assess exclusive contributions to different attributes of output uncertainty. Start time, end time and the magnitude of predicted peak metaldehyde concentrations recorded in response to the variation in model inputs and parameters are considered when assessing the uncertainty contributions.

5.3 Results and discussions

5.3.1 Uncertainties in prediction of metaldehyde concentrations

Uncertainty analysis of metaldehyde prediction model carried out in this study is mainly focused on assessing input and parameter uncertainty. Model structural uncertainty is also known to cause model prediction uncertainty, but it has not been the subject of this study. The negligible difference resulted from the comparison between the output of a reference run with optimum parameter and input values and the average of outputs from Monte Carlo runs indicated that the input and parameter uncertainties have not caused systematic model errors. This is likely to be due to the approximately linear description and representation of the metaldehyde generation and transport processes used in the model.

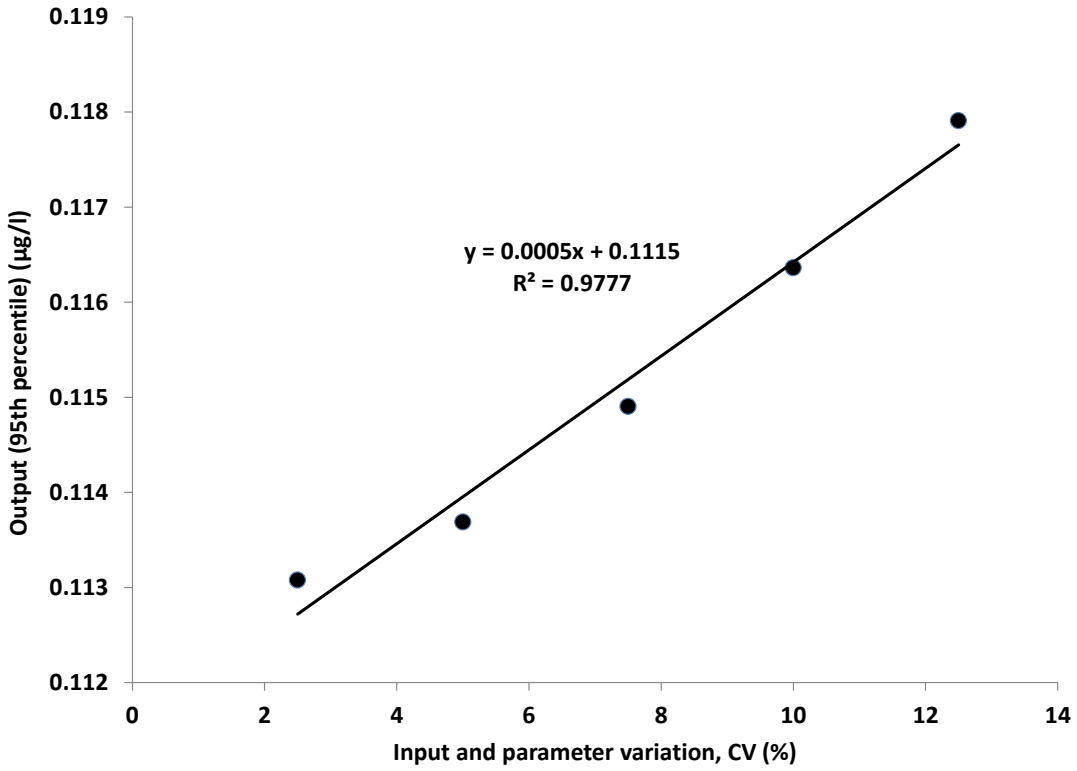


Figure 5.3. Event averaged 95th percentile metaldehyde prediction.

In general, predicted metaldehyde concentrations are found to be moderately tolerant to significant variations in the combined model input and parameter values. The rate of increase in the average 95th percentile of predicted metaldehyde concentrations against increases in input and parameter uncertainties represented as CV (%) is presented in Figure 5.3. Changes in the coefficients of variation of the model input and parameters have resulted in relatively small changes in the average 95th percentile value of predicted metaldehyde concentrations for the event as shown in Figure 5.3. An increase of 0.005µg/l in the average 95th percentile value of predicted metaldehyde concentrations is observed in response to the corresponding 10% rise in the coefficient of variations of all five inputs and parameters. Moreover, a very small gradient value of 0.0005 for the well fitted (R^2 = 0.9777) linear trend line to the data in Figure 5.3 demonstrates that the average 95th percentile of predicted metaldehyde concentrations is tolerant to significant variations in model inputs and parameters. The rate of increase in coefficient of variation and the 90% confidence interval of

predicted metaldehyde concentrations are also compared with increases in coefficient of variation of model input and parameter uncertainties as shown in Figure 5.4 and Figure 5.5 below.

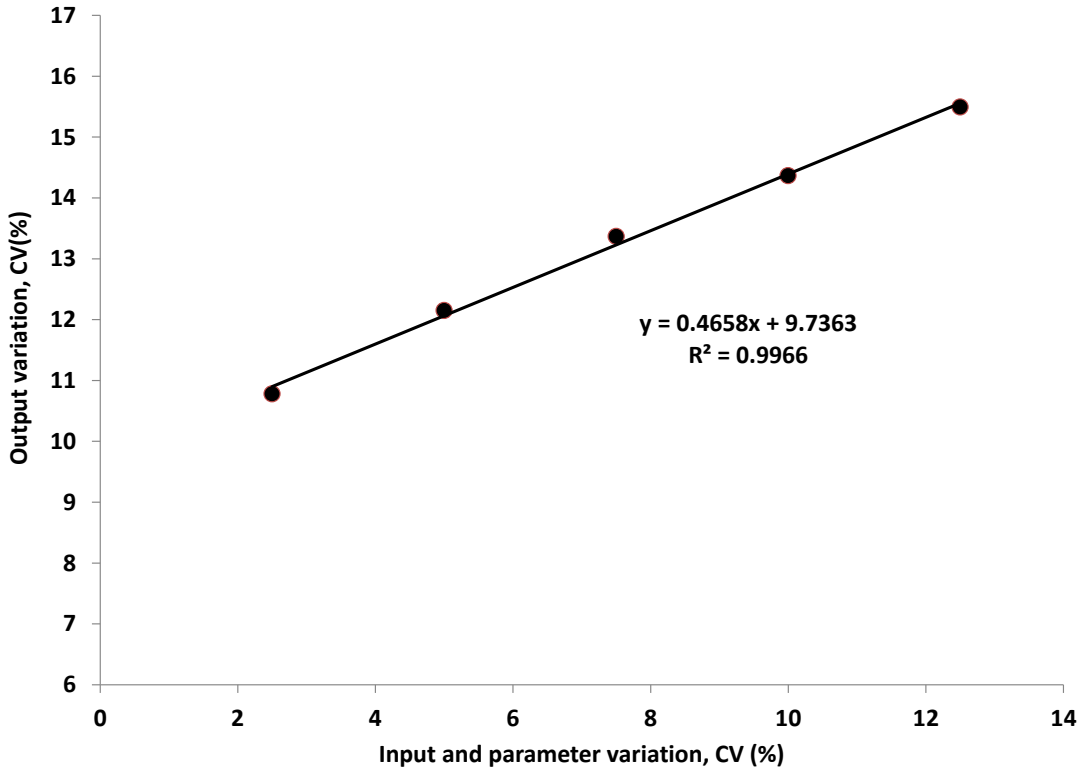


Figure 5.4. Predicted metaldehyde concentration variation in response to forced perturbations in model inputs and parameters.

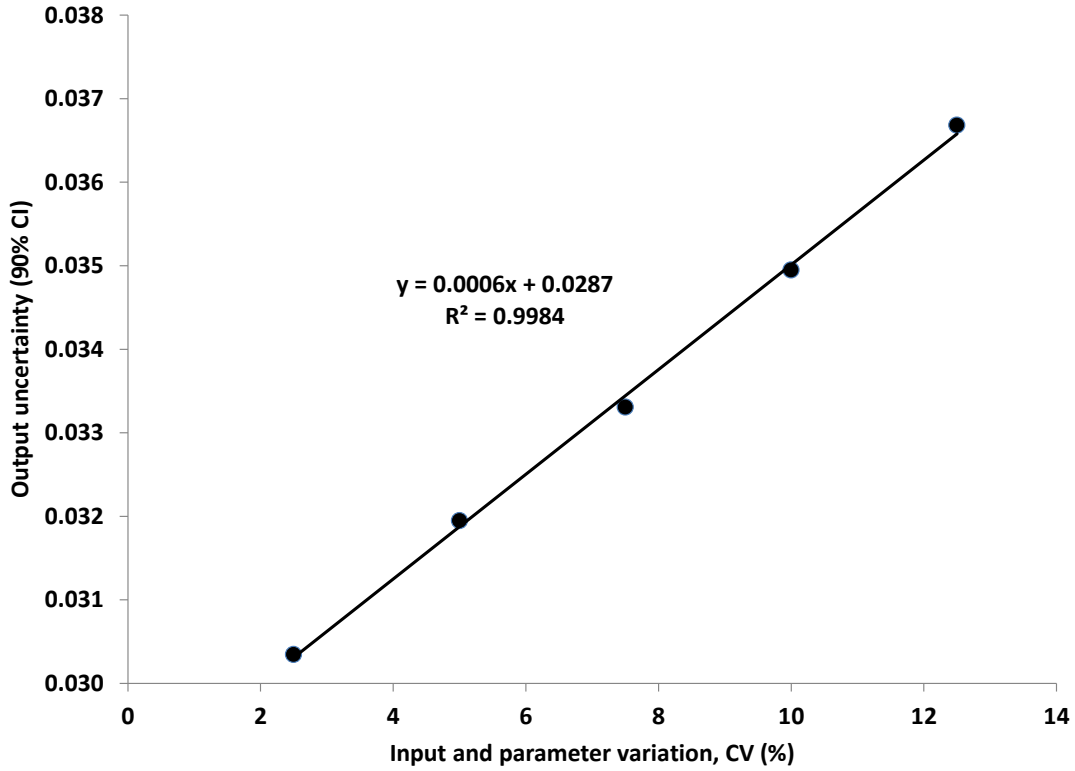


Figure 5.5. Uncertainty in predicted metaldehyde concentration (90% *CI*) in response to forced perturbations in model inputs and parameters.

A relatively small value of the gain factor (G = 0.4658), which is the slope of the plot in Figure 5.4 and measures the amplification or attenuation of output distribution as a result of change in input distributions, is observed. This further demonstrates the small changes in the scale of predicted metaldehyde concentration uncertainties as compared to significant change in input and parameter distributions. The 90% confidence interval of predicted metaldehyde concentrations is observed to increase by 0.006µg/l in response to the corresponding 10% rise in the coefficient of variations of all five inputs and parameters. This showed the reasonably acceptable overall sensitivity of the 90% confidence interval of predicted metaldehyde concentrations to changes in input and parameter uncertainties. Overall, the relatively small amplification of model input and parameter uncertainties observed in the metaldehyde model uncertainty propagation ensures that the model is reasonably tolerant of increases in the levels of input and parameter uncertainties.

5.3.2 Sensitivity Analysis

Results from all three sensitivity analysis performed in this study are presented and discussed in this section. The various techniques adopted to analyse uncertainty contributions from model inputs and parameters (discussed in section 5.2.5) have enabled to clearly identify the main source of output uncertainties in predicted metaldehyde concentrations. Two significant drops in output uncertainty (Figure 5.6) are observed when the width of the 90 % confidence interval output uncertainty dropped by 47.3% and 48.62% in response to holding the model input B and parameter K_{ch} constant respectively. This has demonstrated that the largest uncertainty is originated from parameter K_{ch} (48.62%) followed by the uncertainty in metaldehyde build-up through application B (47.3%). In comparison, less contribution to the uncertainty in predicted metaldehyde concentration are observed from parameters t, K_c and K_{bw} both in the forward and backward stochastic sensitivity analysis. The small uncertainty contribution from time interval between applications and a rainfall event parameter, t, is attributed to the relative longevity of metaldehyde in soil. The significant uncertainty contribution from the metaldehyde build-up (B), which represents metaldehyde build-up through applications on high risk areas in the catchment, indicates that more emphasis needs to be given to collect more accurate metaldehyde application data in the catchment. Accurate estimation of runoff travel time in channel networks are also required to significantly reduce uncertainties in the metaldehyde prediction model. In comparison, efforts to improve catchment data used in the computation of runoff travel time are less significant in terms of reducing the overall uncertainty in predicted peak metaldehyde concentration values. However, accurate prediction of arrival times of metaldehyde spikes is vital in terms of practical application of the model to inform surface water abstraction; hence reducing uncertainties in travel time should receive full attention. On the other hand, the results have shown that less effort is needed in improving data used to estimate the time interval between metaldehyde application and a rainfall event.

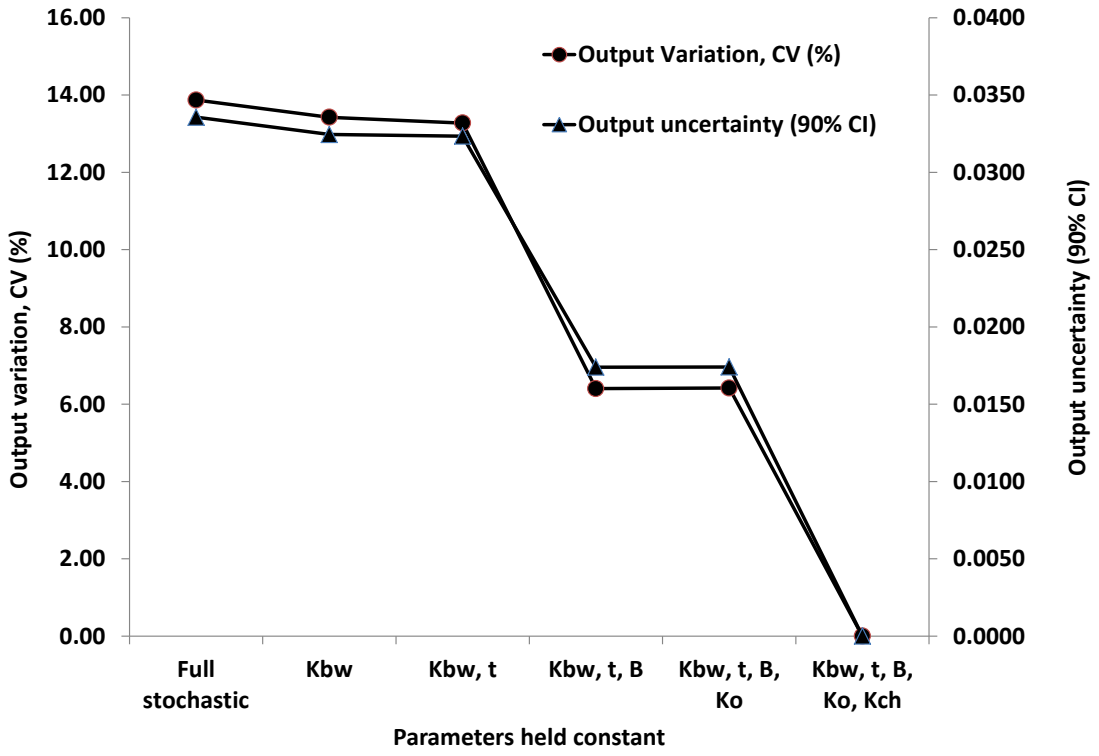


Figure 5.6. Forward sensitivity analysis.

Similar rate of change in output uncertainty is observed when the model inputs and parameters are kept constant in the reverse order during the backward stochastic sensitivity analysis. The similar pattern of output uncertainty variation observed in Figure 5.6 and Figure 5.7 revealed that no significant correlation exists between these model input and parameters. The findings that showed the two main uncertainty contributions are from metaldehyde application areas (*B*) and channel travel time parameter K_{ch} and the lack of significant interactions observed between model parameters appear to be reasonable when considering the mathematical structure of the model and associated assumptions (discussed in Chapter 4). Outputs from the deterministic sensitivity analysis (Figure 5.8) have shown that significant change in event averaged predicted metaldehyde concentration occur only for metaldehyde application areas (*B*). Event averaged concentration, which is used in the absence of any other objective function that enables to adequately show uncertainties caused from all the five inputs and parameters, is not suitable to estimate the level of

uncertainty caused by parameter K_{ch}. This is due the fact that uncertainties in parameter K_{ch} mainly impacts on the overall distributions of metaldehyde concentrations, particularly affecting the arrival time of peak concentrations. As a result, the significant level of uncertainty originated from channel travel time parameter K_{ch} , which has been showed in the stochastic uncertainty analysis method, was not observed in the deterministic sensitivity analysis output. The observed high sensitivity of the model to the metaldehyde application area (B) is mainly owing to its importance in determining the proportion of metaldehyde generated from areas bounded by a serious of isochrones in the catchment. The impact of parameter K_{ch} on the model output uncertainty, which is demonstrated in the stochastic sensitivity analysis, is attributed to its critical importance in determining the distribution of the predicted metaldehyde pollutograph across time.

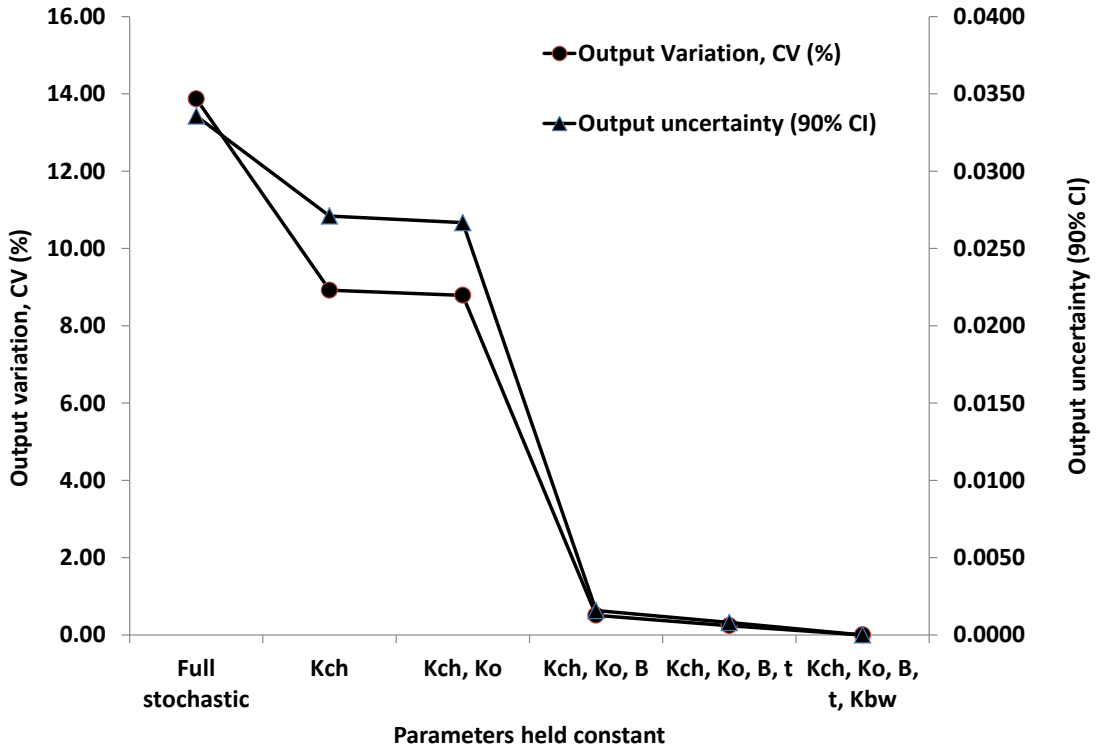


Figure 5.7. Backward sensitivity analysis.

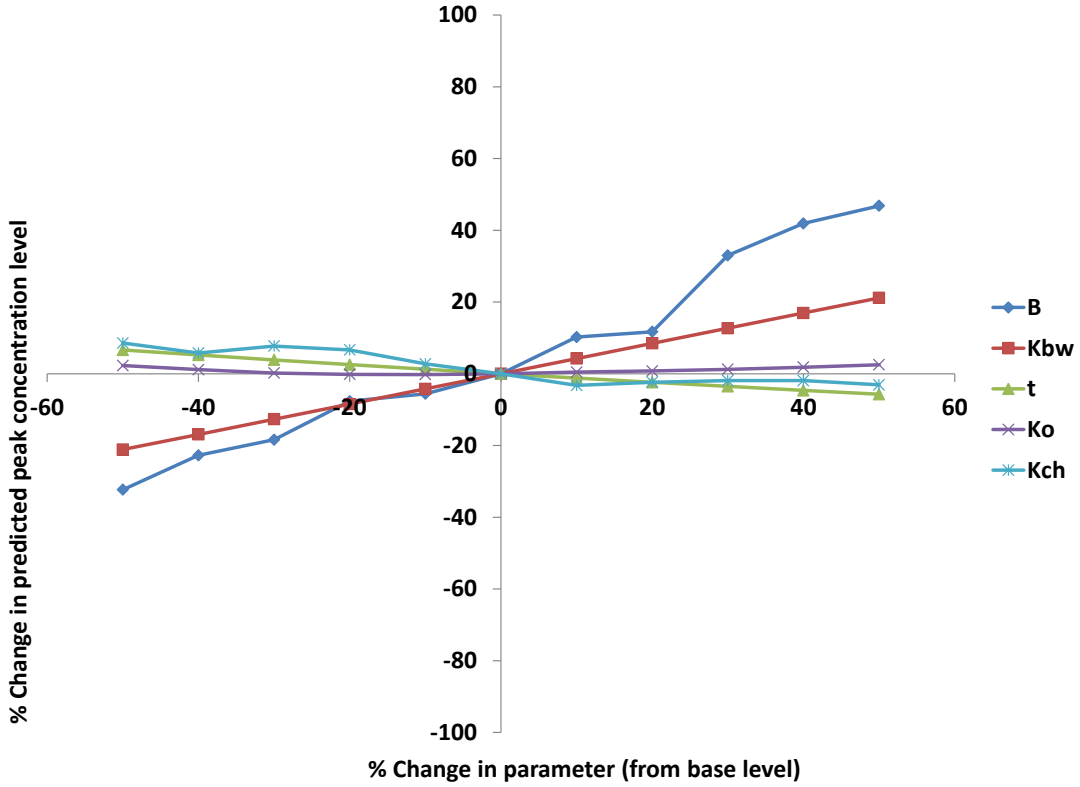


Figure 5.8. Deterministic sensitivity analysis (spider plot).

Various summary statistics including coefficient of variation are used to compare uncertainty contribution results in this chapter. However, computation of coefficient of variation based only on the selected inputs and parameters that are considered uncertain mean that its use to compare the uncertainty contributions has some disadvantages. This is because fixed values are used for the other inputs and parameters despite the uncertainty in their values and thus may have an impact on the output coefficient of variation resulted from variations in inputs and parameters. Consequently, impacts of any correlations that may exist between the selected inputs and parameters being considered and the rest of inputs and parameters on the overall output uncertainty are ignored. Various techniques can be used to tackle this problem as suggested by Jansen (1999) and Saltelli (2002), however the implementation of these techniques are computationally intensive and do not avoid these disadvantages completely.

5.3.3 Characteristics of uncertainty contributions

Identifying the characteristics of uncertainty contributions from selected model inputs and parameters will provide useful information to efficiently plan and employ efforts to reduce model uncertainties. Uncertainties in different inputs and parameters of the model are observed to dominate different features of the model output uncertainty. This section discusses outputs from the stochastic simulations of the metaldehyde model that are carried out individually for each input and parameter and used to assess exclusive contributions to different attributes of output uncertainty. Figure 5.9 showed that in general uncertainties in parameters used in the metaldehyde build-up and wash-off calculations have mainly impacted uncertainties in the level of predicted peak metaldehyde concentrations. Whereas, uncertainties in parameters used in the calculation of runoff travel times have largely impacted on the arrival time of peak metaldehyde spikes following rainfall events. Despite exhibiting high spatial variability, uncertainty in metaldehyde build-up through applications was observed to mainly impact on uncertainties in metaldehyde peak levels than arrival time of peaks as shown in Figure 5.9(c). The overall output uncertainty originated from uncertainties in estimation of parameter t (time interval between metaldehyde application and a rainfall event) is observed to be small. This is due to the relatively slow degradability of metaldehyde in soil, which is represented in the model using half-life value. This minimizes the impact of uncertainties in estimation of metaldehyde application times before a rainfall event on the model output as shown in Figure 5.9(b).

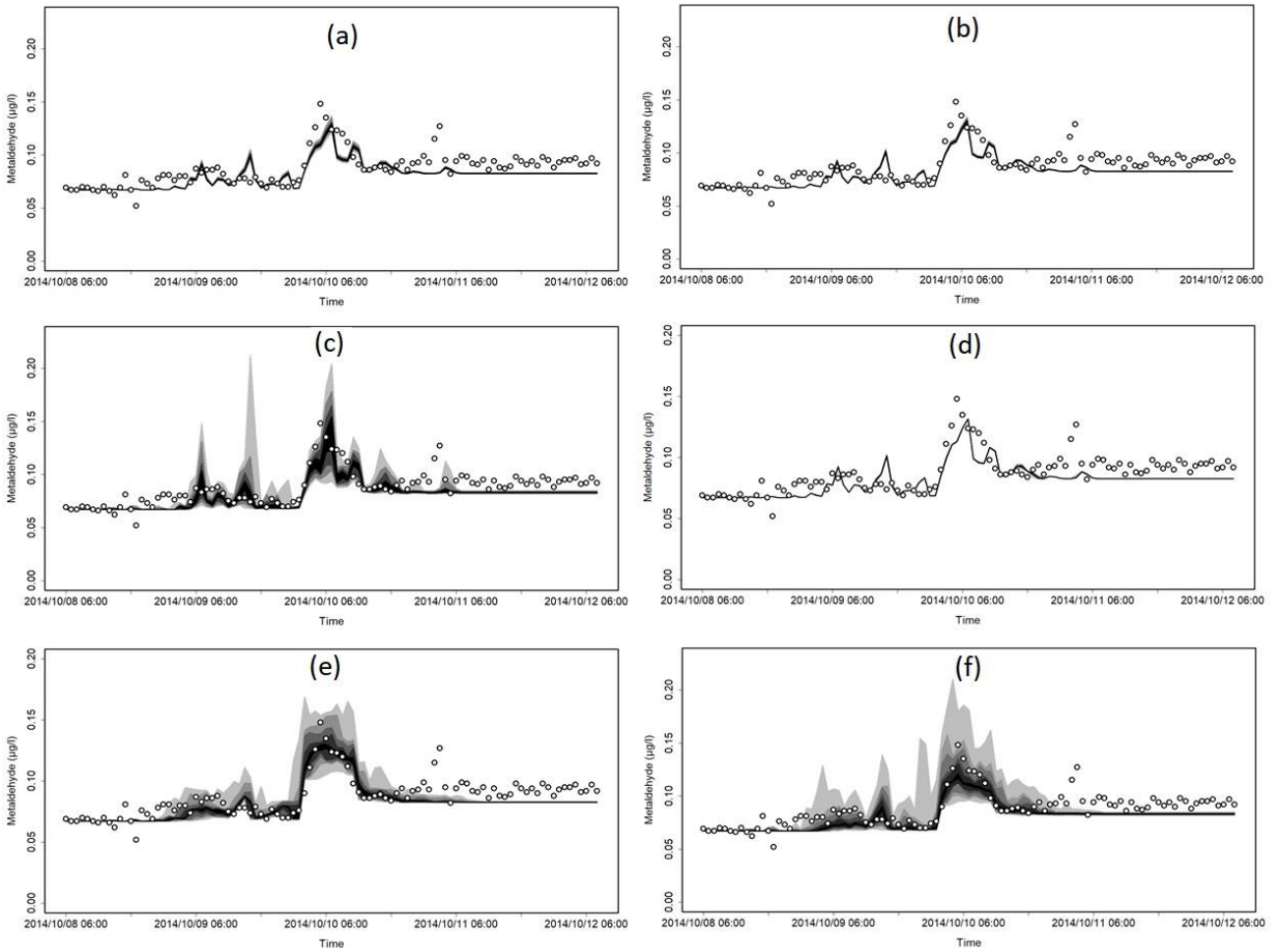


Figure 5.9. 90 % prediction uncertainty bounds of metaldehyde concentrations. (a) uncertainty due to parameter K_{bw} (b) uncertainty due to parameter t (c) uncertainty due to metaldehyde application on high risk area B (d) uncertainty due to parameter K_o (e) uncertainty due to parameter K_{ch} and (f) combined uncertainty.

Overall, the model is found to be much more sensitive to channel travel time parameter K_{ch} than overland travel time parameter K_o as shown in Figure 5.9(d) and Figure 5.9(e). This is due to the significantly higher rate of runoff that occurs in channel networks as compared to in overland flow. Uncertainties originated from travel time parameter K_{ch} is observed to mainly impact on estimation of the amount of runoff arriving at the catchment outlet per unit time during runoff periods following rainfall events. This has direct influence on the estimation of metaldehyde loads in runoff arriving at the catchment outlet per unit time following rainfall events. Hence, variations in parameter K_{ch} are likely to cause direct influence on the timely distributions of model predicted metaldehyde concentration values. Increase in the value of K_{ch} is observed to cause a widely dispersed metaldehyde concentration prediction across time with the metaldehyde spikes starting early and staying for longer. On the other hand, lower values of travel time parameter K_{ch} have caused narrowly distributed metaldehyde concentration predictions with spikes that are relatively short lived and start late. It is evident from Figure 5.9(e) that uncertainties in travel time parameter K_{ch} have influenced uncertainties of both arrival time and peak levels of predicted metaldehyde concentrations. However, the impact on the arrival time uncertainty is observed to be larger as compared to impacts on peak level uncertainty. These results have demonstrated that accurate representations of runoff transport processes in channel networks throughout the catchment is important in models used for prediction of storm water pollutions at catchment outlets.

5.4 Conclusions

Catchment scale hydrological models such as the metaldehyde prediction model developed in this study are known to incorporate many model inputs and parameters. Some of these inputs and parameters are estimated statistically whereas others have physical significance and are usually measured using resource intensive field exercises. Due to lack of sufficient data these model inputs and parameters are often estimated via calibration with no sufficient level of information on the scale of their influence on the model prediction (Beven & Binley 1992). This study has primarily

focused on assessing uncertainty propagation in the metaldehyde prediction model and establishing the relative importance of model inputs and parameters uncertainty in improving prediction accuracy. Results from uncertainty analysis conducted on the metaldehyde prediction model indicated that metaldehyde concentrations are moderately sensitive to variations (likely greater than errors during calibration) in input and parameter values. All three sensitivity analyses have clearly shown that metaldehyde application area (B) has the main influence on the model output. The stochastic sensitivity analysis has also clearly shown the significant level of model output uncertainty originated from parameter K_{ch}. Identification of inputs and parameters with the most influence on predicted metaldehyde concentration would help in the planning and design of field exercises to collect data, which would result in significant reductions in costs and effective utilization of limited resources (Kuczera & Mroczkowski 1998). Furthermore, uncertainty quantification and analysis has enabled probabilistic presentation of predicted metaldehyde concentrations, which summarizes model outputs using various statistical quantities such as percentiles and confidence levels. Unlike point estimates provided by deterministic predictions, this provides more information on model prediction error and enables risk-aware abstraction management decision making (Briggs et al. 2012). The work discussed in this chapter provides a new approach to assess spatial uncertainties associated with catchment scale water quality models, which is currently lacking in the literature. The grid based physically distributed structure of the metaldehyde model combined with the use of high spatiotemporal resolution data and efficient spatial uncertainty analysis tool have enabled us to assess spatial uncertainties of inputs and parameters in the catchment scale water quality model.

Chapter 6

Business Case and Implementation

6.1 Introduction

Abstraction from surface water accounts for 67% of STW's production. The development of a new approach to water abstraction therefore represents a major opportunity to make significant improvements to STW business. During dry weather, the amount of water available for abstraction is restricted by the need to maintain a minimum daily flow in the river. This regulatory monitoring often includes a variety of abstraction licence conditions, which are provided with the aim to meet environmental demands downstream of the abstraction point. Lack of data on flow arriving at the abstraction point mean that particularly during low flow periods water resource operators often face with uncertainty when deciding on how much water can be abstracted per day. The real-time abstraction management scheme developed in this study predicts the availability of flow in the river at the abstraction point. The approach can further be applied to other catchments in the STW's region and could help to offset the loss of 150Ml/d of abstraction by 2025 due to the WFD Restoring Sustainable Abstraction (RSA) scheme and climate change (STW 2014). Replacement of lost abstraction by developing a new water resource costs £1-2 million per Ml/d (Canal and River Trust 2015). The developed approach can also provide potential for immediate benefits from increased flexibility to further maximise the use of sources with the lowest cost to serve.

Metaldehyde is a pesticide that is poorly removed using conventional activated carbon. Water companies in the UK currently have a Department of Water Inspectorate (DWI) Section 19 undertaking relating to metaldehyde at a number of water treatment works. Although STW have an extensive catchment management programme, high levels of metaldehyde observed in rivers following rainfall events in recent years have indicated that compliance cannot be guaranteed through catchment management only. Consequently, additional investment in specialised Advanced Oxidation Processes (AOPs) and/or modified carbon adsorption

technology is planned. The model developed in this study predicts the concentration of metaldehyde in the River Leam catchment at abstraction site and enables abstraction to be suspended to avoid failures. As a result, the application of the model will reduce the scale of capital expenditure (CAPEX) outlay and the associated risks of a higher level of initial investment. By reducing the level of metaldehyde in water resource reservoirs, the scheme will also help to significantly reduce operational expenditures (OPEX) required for metaldehyde removal using the new treatment methods. The nature and operationally suitable structure of the developed model enables its application in other catchments. This chapter details the implementation of these models developed in this study at Severn Trent water Ltd and discusses the various benefits that accrue from the real-time abstraction management scheme.

6.2 Business case for exploiting the outcome of the project

In this study two different models outlined above have been developed in two different study catchments. This section aims to set out a business case for implementation and exploitation of the developed modelling approaches in catchments throughout STW region. This involves assessing the benefits of implementing the new RTAM system within a range of abstraction sites at Severn Trent Water. Additional benefits expected to emerge will include the use of the model to inform catchment management strategies. The successful implementation of the project outputs will transform the capability, flexibility and costs associated with STW's existing abstraction protocols. This is expected to lead to significant capital and operating cost savings for the company, thus generating further investment and R&D funding, and an enhancement of strategical planning. The expertise STW and Sheffield University have developed in real-time abstraction management during the course of this study is an area they may be able to exploit and sell on to other utilities. Works on implementation and system integration of the developed models in STW region is currently going on.

6.2.1 Water resource model

It is difficult to estimate the financial benefit of the water resources associated abstraction management scheme, as the development of new sources of abstraction is a complex process with a 25 year planning horizon. However, monetary benefits of enabling to abstract more water can be quantified in terms of reducing future water resource development investment CAPEX. When the licensing system was designed by the precursor to the Environment Agency in the early 1960s and issued to existing abstractors, there was a surplus of water in most areas. This has now decreased as our understanding of the impact that abstraction has on the environment has improved. In future, pressures on water availability will increase because of climate change. By the 2050s, it is anticipated summer temperatures may increase and rainfall may decrease and short duration drought conditions (12-18 months) are likely to become more frequent. Therefore the challenge faced by water resource operators is a potential shortage of future water supply and how to maximise abstraction. As shown in chapter 3, preliminary simulations using the model in the River Dove catchment identified the potential to increase abstraction by an average of 30 Ml/d during periods of low river flow. This is of strategic importance to STW as current forecasts suggest that STW will lose the ability to abstract 150Ml/d by 2025 due to new regulations and hydrological climate change effects (Figure 6.1). Assessment of water resources in STW's region over the period of twenty years shows that the scale of the supply/ demand challenge could be as much as 150Ml/d by 2025, 300Ml/d by 2030 and 340Ml/d by 2040. Replacement of these reduced water availability for abstraction by developing a new water resource infrastructure costs £1-2M per Ml/d. Hence, potentially reducing or avoiding this cost through implementation of the developed water resource modelling approach represents a significant benefit to STW. Moreover, the flow model will be essential to ensure that lost abstraction, during periods of high pesticide concentration, can be recovered by maximising abstraction during periods when pesticide (metaldehyde) concentration is low.

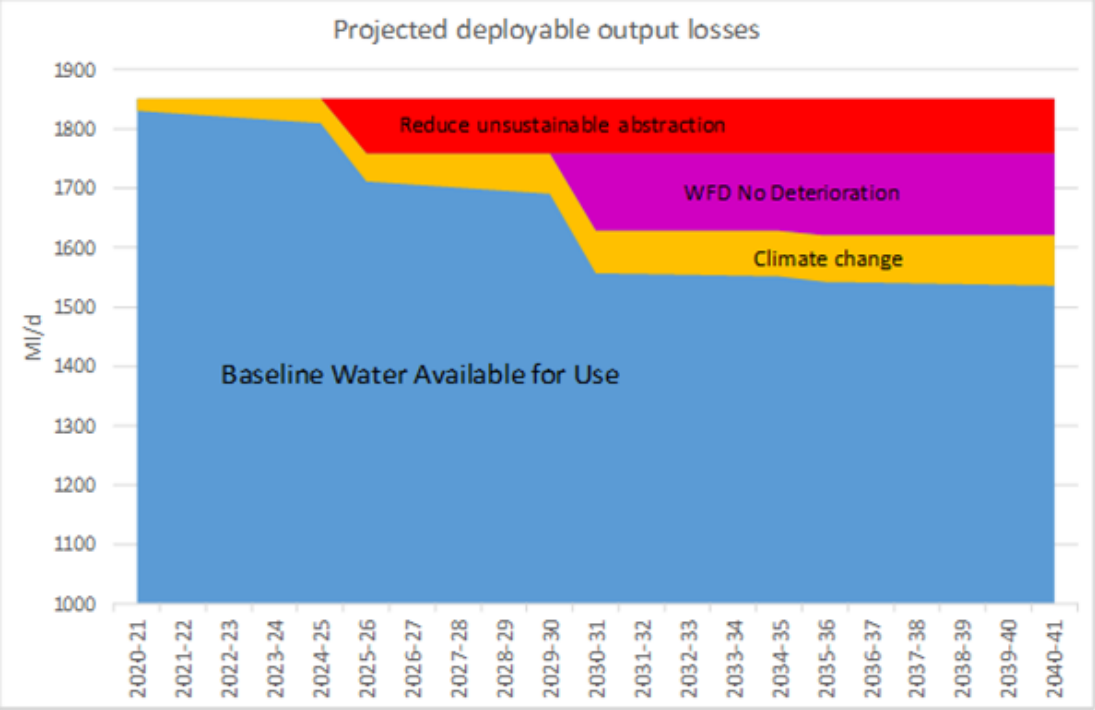


Figure 6.1. Reduction of available water for water supply use due to climate change and reduction of unsustainable abstractions to protect the environment. (STW 2017)

In addition, the water resource model enables to increase STW's performance associated with the restrictions on use outcome delivery incentive (ODI), which can generate a substantial ODI reward or penalty depending on the number of water use restrictions put on customers due to low storage levels in reservoirs. The exact amount of the penalty and reward varies depending on the timing, extent and duration of the imposed restriction on water use. However, avoiding the reputational damage associated with imposing restriction on water use is the main incentive for water companies. By enabling to improve reservoir storage levels during dry periods, the water resource model provides a potential to minimize or avoid the need to impose such restrictions on water use. As a result it helps to generate significant benefits in terms of restriction on water use ODI rewards/ penalty and avoiding reputational damages to STW.

6.2.2 Metaldehyde model

Recent years have seen increasing concentrations of pesticide in water supplies and treatment techniques are forecast to be a significant expense to STW. Of particular concern is metaldehyde which is not currently removed by existing treatment processes. Implementing the metaldehyde forecasting model at abstraction sites will allow stopping abstracting water when concentrations of metaldehyde are high. This will improve the quality of the water abstracted for supply and generates a number of business benefits to STW. In this section, the financial benefit of the developed model is derived based on the avoidance of investment in installing and operation of new treatment and online monitoring options. To comply with the regulatory standard for metaldehyde, STW have an option to install new advanced treatment processes at twelve treatment works, which are all at risk of failing regulatory standard for metaldehyde, or install monitoring equipment to detect metaldehyde and enable abstraction to be suspended during peak concentrations caused by rainfall events. Installation of the new advanced treatment processes throughout the region cost STW £60M CAPEX and £1M OPEX. The option to install new monitoring equipment, which is currently being trialled by Affinity water, is expected to require investment of £0.335M CAPEX and £00.065M p.a. OPEX per abstraction point. The implementation of the metaldehyde prediction model will avoid the need for investment in on-line monitoring at 12 points of abstraction and would save £4.02M CPAEX and £0.78M p.a. OPEX. Table 6.1 below shows comparison of installing and running costs of advanced oxidation process at different sites with and without the implementation of the metaldehyde model. The names of the sites are anonymously mentioned using alphabets due to STW security policy.

Table 6.1. Feasibility cost summary for Advanced Oxidation Treatment of metaldehyde with and without the implementation of the metaldehyde prediction model – Data supplied by STW innovation team

		Treatment costs current peak metaldehyde			**Treatment costs using model to reduce peak metaldehyde**		
Site	Production (Ml/d)	Build Cost/yr	Electricity Cost /yr	H_2O_2 Cost	Build Cost/yr	Electricity Cost /yr	H2O2 Cost
A	62	£11,070,321	£393,901	£99,514	£4,812,500	£309,494	£99,514
B	34	£6,070,821	£112,543	£55,542	£875,000	£56,272	£55,543
C	24	£4,285,285	£140,679	£55,542	£1,750,000	£112,543	£55,543
D	58	£10,356,106	£337,629	£138,857	£3,500,000	£225,086	£138,857
E	27	£4,820,946	£84,407	£64,800	£437,500	£28,136	£64,800
F	55	£9,820,446	£225,086	£127,285	£2,187,500	£140,679	£127,286
G	105	£18,748,124	£675,259	£266,142	£7,374,500	£478,309	£266,143
H	160	£28,568,570	£984,753	£386,485	£11,375,000	£731,531	£386,486
Total	**525**	**£93,740,621**	**£2,954,259**	**£1,194,171**	**£32,312,000**	**£2,082,050**	**£1,194,172**

In addition, the implementation of abstraction management using the metaldehyde model contributes to the following two ODIs associated with providing good quality drinking water.

1. Compliance with drinking water quality standards ODI is associated with a penalty of £2119.7 for every 0.01% below the target percentage of water quality compliance level per year.
2. Successful catchment management schemes ODIs is associated with a reward/penalty of £1,271,600 per scheme.

Apart from staff time, the only additional cost associated with the implementation of the model will be the availability of real-time radar rainfall data. This has been arranged to be accessed at a cost of £4K per year for the current study catchment.

When rolling out the model to other catchments in the region, additional cost for the real-time radar rainfall data per catchment will be much less than £4K per catchment per year.

6.2.3 Business risk and feasibility of implementations

The risk of implementing the real-time abstraction management models developed in this study is relatively low as the models generally inform abstraction decisions without needing any modification of the existing system. However, the metaldehyde application season coincides with the infilling period for impounding and pumped storage reservoirs. As a result, suspension of abstractions during this period poses a risk of missing opportunities to abstract more water to increase storage availability to meet demands over the summer season. Metaldehyde samples collected over the 2014-2016 application periods have been analysed to assess the extent of this impact on water resources. Based on the data collected using auto-samplers over the three application periods, metaldehyde spikes observed at abstraction point in the river leam catchment have showed durations ranging from 12 hours to about 48 hours depending on rainfall intensity and durations. But on average 12-24 hours can be considered as the average time period that intakes need to be shut down to avoid peak metaldehyde levels occurring following a typical rainfall event. The number of times these metaldehyde spikes occur during September - December (and the prediction model would suggest to shutdown intakes) is dependent on the frequency of significant rainfall events during these months. Based on the data collected over the past three years, a rainfall event with approximate intensity of more than 0.5mm/hr and with duration of more than 3 hours is likely to cause peak metaldehyde levels. It should be noted that this is very rough estimation as occurrence of peak metaldehyde levels is highly dependent on where in the catchment it rained and antecedent moisture conditions – which are all considered in the prediction model. To estimate the average number of times such kind of rainfall events occur, analysis of 12 years of radar rainfall data (2004-2015) for the River Leam catchment has been carried out, which also enabled to assess impacts on the water resource availability in reservoirs. This analysis showed that on average there will be 20 rainfall events that

are likely to cause metaldehyde peaks and require intake shutdowns over the period of 1st September to 31st December. The year 2012 would have needed the highest intake shutdowns of the 12 years analysed with 30 events (on average 23 days shutdown) and 2007 would have needed the least intake shutdowns with 17 events (Figure 6.3 & Figure 6.4). Considering an average duration of metaldehyde spike caused by a single rainfall event to be 18hrs, a rough average of 15 days intake shutdown period over the September – December period is estimated based on these assessments. Thus, abstraction management in this catchment is likely to cause an overall maximum storage reduction of 6.5% in Draycote reservoir at the end of the metaldehyde application season assuming the 100Ml/d achievable capacity of the intake at this abstraction site (Maximum capacity is 115 Ml/d) is consistently used (Figure 6.2).

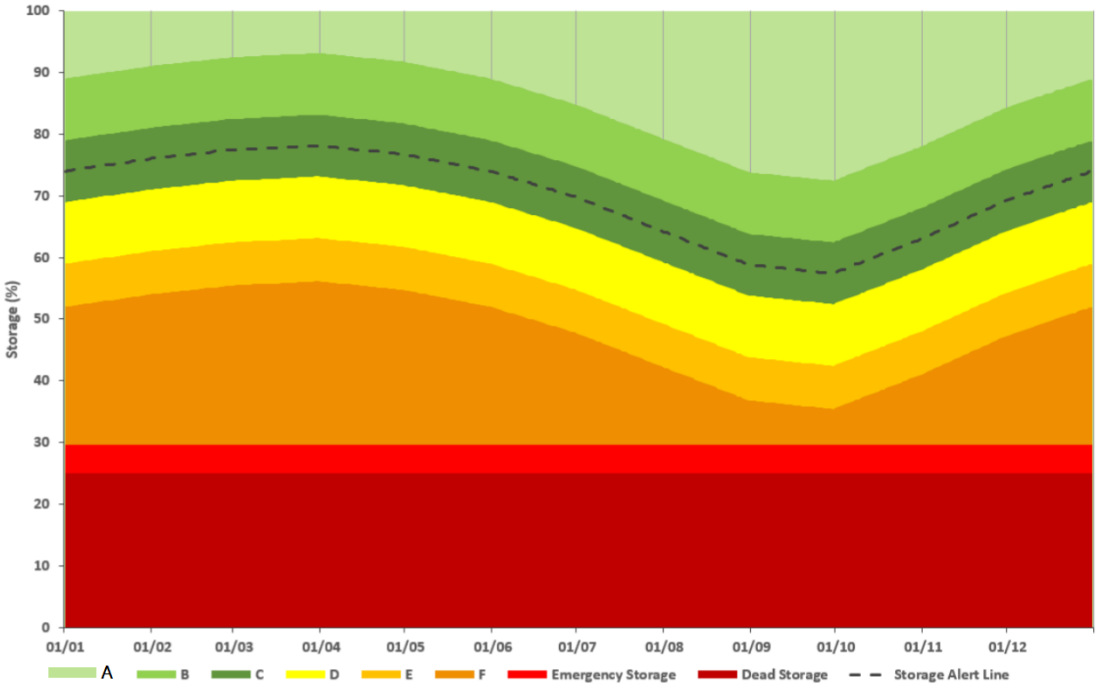


Figure 6.2. Draycote reservoir storage and drought management trigger zones (STW 2018).

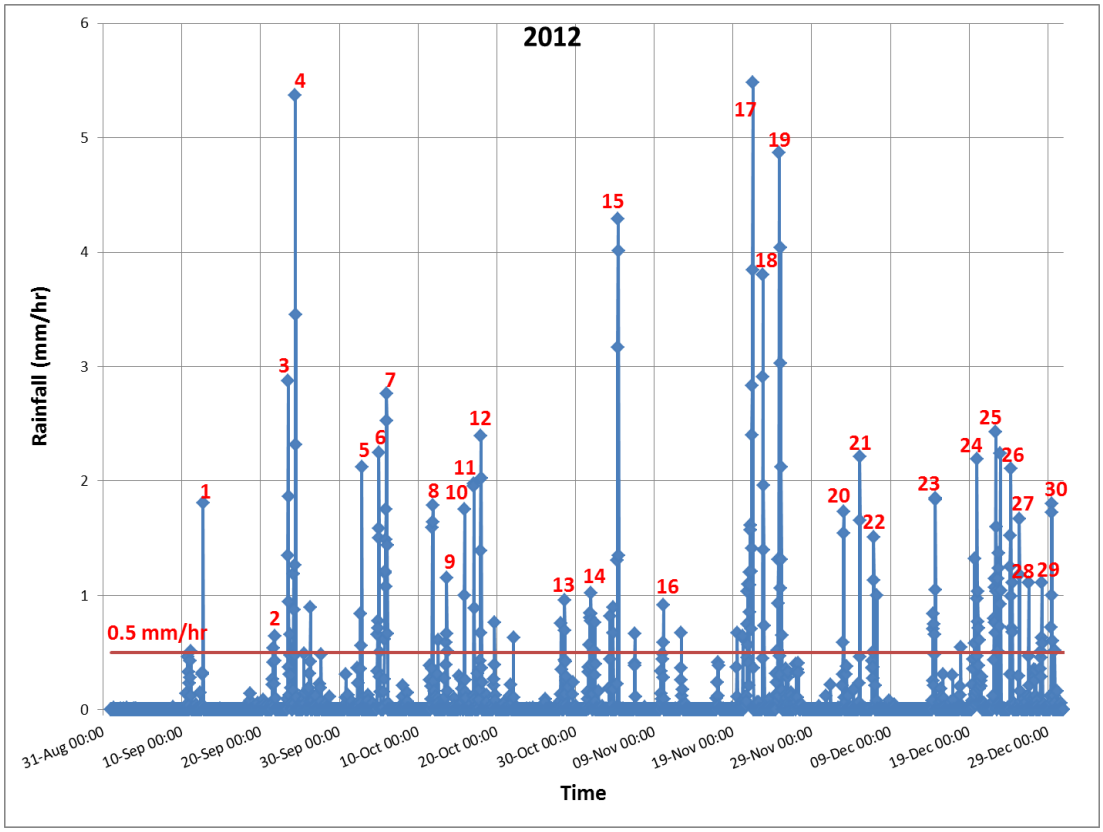


Figure 6.3. Rainfall events in the 2012 metaldehyde application season that would have required intake shutdowns.

-

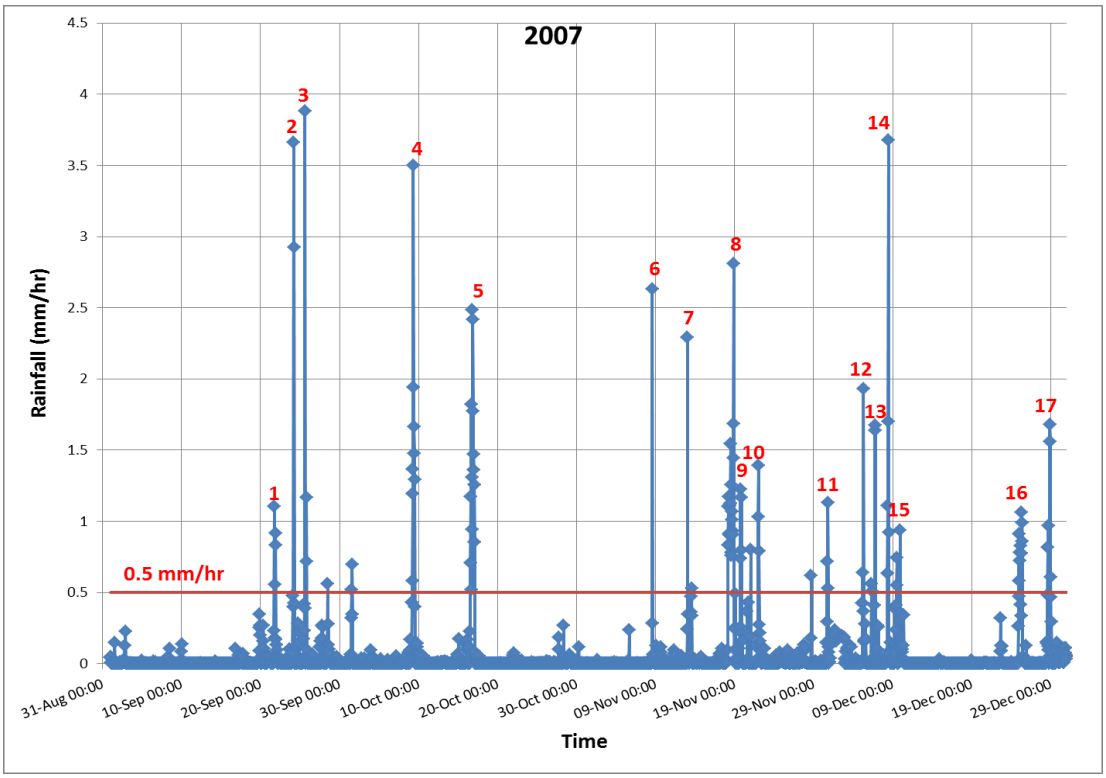


Figure 6.4. Rainfall events in the 2007 metaldehyde application season that would have required intake shutdowns.

With a view to minimize impacts on water resource availability, an abstraction management decision making framework that enables temporal suspensions of abstraction based on storage levels in the reservoir has been developed. Information on storage levels are used to determine the maximum length of period that abstraction can be suspended at a specified time. A tool to align this period to the arrival of the highest level of metaldehyde has been developed to enable to avoid the peaks while minimizing impacts on storage levels and is agreed with water resource managers at STW (Table 6.2). Figure 6.5 shows the trigger zones set out in Draycote reservoir where different levels of abstraction decision is required based on storage availability.

Table 6.2. Draycote reservoir trigger levels for the metaldehyde model based abstraction management

	Control curves		**Maximum number of hours abstraction can be suspended (h)**				
	Top of Drought Management Trigger Zone B	Tope of Drought Management Trigger Zone C	**12**	**24**	**36**	**48**	**60**
Date	Storage levels in Draycote reservoir (Percentage full)						
04/09/2017	83.5	58.5	63.5	68.5	73.5	78.5	83.5
11/09/2017	83.1	58.1	63.1	68.1	73.1	78.1	83.1
18/09/2017	82.8	57.8	62.8	67.8	72.8	77.8	82.8
25/09/2017	82.5	57.5	62.5	67.5	72.5	77.5	82.5
02/10/2017	83.5	58.5	63.5	68.5	73.5	78.5	83.5
09/10/2017	84.8	59.8	64.8	69.8	74.8	79.8	84.8
16/10/2017	86.0	61.0	66.0	71.0	76.0	81.0	86.0
23/10/2017	87.3	62.3	67.3	72.3	77.3	82.3	87.3
30/10/2017	88.6	63.6	68.6	73.6	78.6	83.6	88.6
06/11/2017	90.0	65.0	70.0	75.0	80.0	85.0	90.0
13/11/2017	91.5	66.5	71.5	76.5	81.5	86.5	91.5
20/11/2017	92.9	67.9	72.9	77.9	82.9	87.9	92.9
27/11/2017	94.3	69.3	74.3	79.3	84.3	89.3	94.3
04/12/2017	95.0	70.4	75.4	80.4	85.4	90.4	95.4
11/12/2017	95.0	71.5	76.5	81.5	86.5	91.5	96.5

18/12/2017	95.0	72.6	77.6	82.6	87.6	92.6	97.6
25/12/2017	95.0	73.6	78.6	83.6	88.6	93.6	98.6
		Max. water loss (ML)	**50**	**100**	**150**	**200**	**250**
		Max. water loss (%)	**0.22**	**0.43**	**0.65**	**0.87**	**1.09**

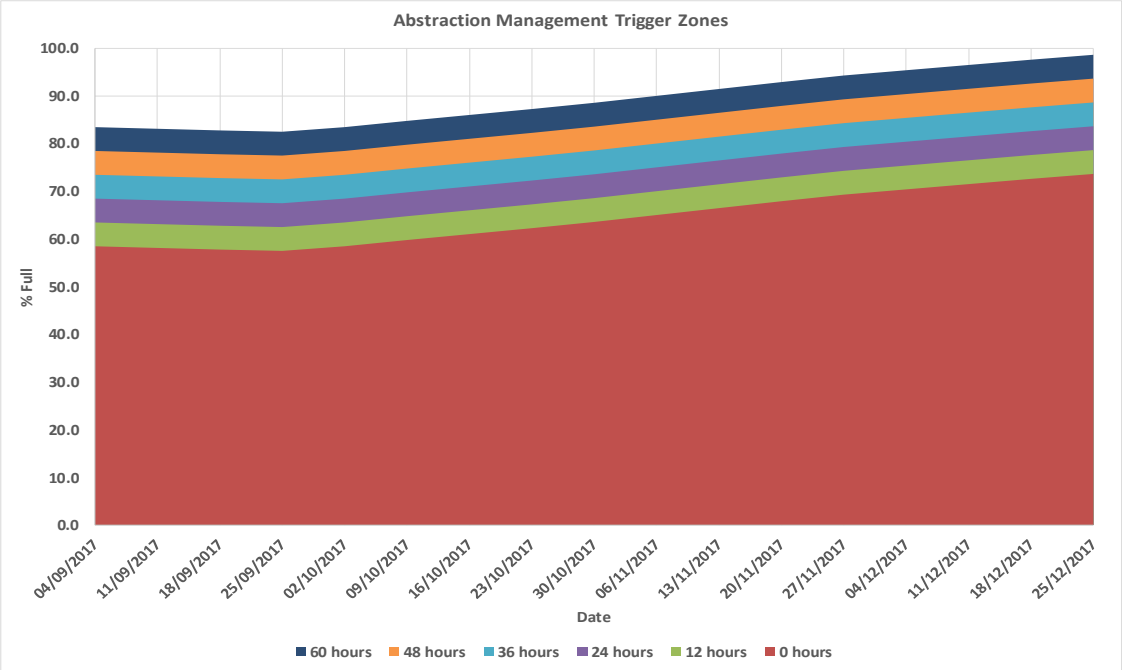


Figure 6.5. Draycote reservoir allocated trigger zones to control metaldehyde model based abstraction management.

6.3 Conclusions

Assessment of the potential benefits associated with water resources management of the real-time abstraction management scheme has demonstrated that the scheme can significantly increase resilience in surface water catchments. The sustainability and cost effectiveness of the proposed solution mean that water companies, which usually

incur high capital expenditure to increase water resources resilience to meet service levels, can generate significant financial and environmental benefits through the implementation of the scheme. The ability to predict the concentration of metaldehyde in water courses across the water sector will improve the efficiency of catchment based solutions for this problematic pesticide. The assessments presented in this chapter have demonstrated that implementation of the metaldehyde model based abstraction management scheme will avoid significant amount of unnecessary investments in new treatment processes; thus saving the cost and embodied carbon associated with their construction and operation. The fundamental research behind the development of the model and research outputs have been shared with the water industry via presentations at varies events and published paper.

Chapter 7

Summary and Conclusions

7.1 Meeting the aim and objectives

The overall aim of this study is to develop an advanced surface water abstraction management scheme that enables smarter control of surface water abstractions based on forecasted availability of water in the river, metaldehyde concentration levels in the river at abstraction sites and current storage levels. Original sub-aims and objectives of the study which are stated in Table 1.1 in section 1.3 were addressed in this study as discussed below.

a) Investigate the use of hydrological forecasting in order to maximize the amount of water abstracted in surface water catchments and develop smarter abstraction management scheme that enables to vary abstraction volumes based on availability of water in the river.
    I. Identify a rainfall-runoff modelling approach, data assimilation and uncertainty analysis methods suitable for real-time abstraction management and assess operational suitability.
    II. Develop a real-time stochastic flow forecast model by combining a conceptual rainfall-runoff model with Bayesian based uncertainty analysis method.
    III. Develop a water resource management model and integrate it with rainfall-runoff model to assess the potential benefits of the real-time abstraction management scheme and investigate implications on water resources.

A real-time abstraction management scheme is devised that is composed of a hydrological forecasting and water resource management model with the objective of enabling to abstract more water at surface water abstraction sites. A Bayesian based uncertainty analysis method is used to stochastically calibrate and validate a conceptual rainfall-runoff model. The performance of the model in the stochastic model calibration and validation method is evaluated using P-factor and R-factor

values. In addition, optimum parameter values are used in the model to evaluate its overall prediction efficiency using Nash-Sutcliffe model efficiency coefficient. Stochastic calibration and validation of the PDM is undertaken in simulation mode where the only model inputs are rainfall and potential evaporation. State correction data assimilation technique, which uses a set of rules for adjusting model states, is applied to enable to run the flow model in forecast mode and improve forecast accuracies. A water resource management model is developed to assess implications of real-time abstraction management scheme on reservoir levels and associated water resources management decisions. The integrated model outputs have showed that on average 30 Ml of more water per day could have been abstracted in the study catchment using the real-time abstraction management scheme during the 2011 dry period (Figure 3.10). The approach has also demonstrated that the implications of the additional amount of water abstracted in terms of helping to avoid the rapid decline of reservoir storage levels during dry periods (Figure 3.11).

b) Develop a new travel time based physically distributed catchment scale metaldehyde prediction model and improve understandings of short term fluctuations in metaldehyde concentrations at surface water catchment outlets caused by rainfall runoff events.
    I. Identify the limitations of current water quality modelling practices with regards to their applicability for describing short term fluctuations in pollutant concentrations at catchment scale.
    II. Enable physically distributed representation of runoff and metaldehyde generation throughout the catchment and develop improved representation of spatiotemporal variability of pollutant transport.
    III. Calibrate and validate metaldehyde model using newly collected high resolution water quality dataset in the study catchment and assess operational suitability of the model to enable smarter abstraction management.

A review of the literature on water quality models has showed the limitations of existing catchment scale pollutant models in representing short-term dynamics that is

responsible for the occurrences of peak pesticide concentrations following rainfall events. A runoff travel time based metaldehyde prediction model is developed in a GIS environment to enable prediction of short term fluctuations in metaldehyde concentrations at surface water catchment outlets. The model integrates spatiotemporally distributed runoff generation, routing and build-up/wash-off of pollutant in the study catchment. A grid based travel time computations in a GIS environment enabled an improved representation of spatiotemporal variability of pollutant transport in runoff. Water quality data collection campaign has been conducted over a period of 3 years (2014 – 2016) in the study catchment. Automatic samplers were used to collect hourly surface water samples following rainfall events, which enabled the calibration and validation of the metaldehyde prediction model. The model performance is evaluated using a set of error statistics, which showed the models' suitability for the intended purpose of quantifying potential exposures to peak metaldehyde concentrations and enabling smarter abstraction management.

c) Investigate propagation of catchment scale spatially distributed input and parameter uncertainty in the metaldhyde prediction model and enable risk aware abstraction management decisions.
    I. Define and parametrize probability distribution functions of identified model input and parameter uncertainties and generate realizations from predefined probability distributions to represent uncertainties in inputs and parameters.
    II. Analyse propagation of spatially distributed input and parameter uncertainties in the metaldehyde prediction model.
    III. Summarize results from uncertainties analysis using various measures and asses the relative significance of each input and parameter.

Large number of spatially distributed inputs and parameters are involved in the catchment scale metaldehyde prediction model. Inputs and parameters for consideration in the uncertainty analysis method are selected based on existing knowledge on the underlying techniques used to develop the model and information gathered from literatures on main sources of uncertainty in storm water quality

models. The selected parameters are believed to cause the majority of uncertainties associated with model inputs and parameters. Different techniques are employed in representing uncertainties associated with each input and parameters considered in the uncertainty analysis using probability distributions. Assigned probability distributions are parameterized using various catchment datasets and calibrated value of parameters in the deterministic metaldehyde model. Monte Carlo based uncertainty analysis tool is used to generate realizations of model input and parameters using the predefined probability distributions. This has been used to propagate inputs and parameters uncertainties of the model, which enabled to assess the influence and relative significance of each input and parameter uncertainty on model outputs. Uncertainty analysis results presented in probabilistic prediction graphs and summarised using coefficient of variation, confidence levels and gain factor have enabled to efficiently compare the various level of uncertainties. The additional information presented on uncertainties associated with the metaldehyde predicted outputs are important to enable making risk-aware decisions.

7.2 Summary of findings

This section summarizes the main research findings of the study. Initial Investigations of real-time abstraction management implications on water resources has showed that a cost effective and sustainable solution to address the growing concerns associated with the increasing pressures on the water environment can be addressed by developing and integrating a suitable flow forecast modelling approach and a water resources management model. The majorities of existing flow forecasting studies are set out with the main objective of meeting the requirements of flood warning and protection systems. As a result, wide ranges of existing real-time rainfall-runoff modelling approaches are focused on forecasting the exceedance of certain flow thresholds in rivers. However, applications of flow forecasting models to enable real-time abstraction management are required to be capable of forecasting service flow ranges, which is the variety of flow conditions ranging between the extreme dry and flood conditions. This raises specific modelling requirements and provides unique challenges due to a number of operational issues associated with

abstraction management, but these challenges can be categorized in to two main areas of abstraction management. Firstly, smarter abstraction management is required to keep the balance between enabling to abstract more water and avoid breaching abstraction licence condition, Thus, uncertainties associated with flow forecast models need to be carefully analysed and presented to enable risk aware abstraction decision making. Findings of this research have shown that, an integrated hydrological modelling framework that combines a conceptual rainfall-runoff model, Bayesian based uncertainty analysis method and data assimilation technique can be used to develop a flow forecast model suitable for surface water abstraction management purposes. The posterior probability distributions of flow forecast model parameters provided the required information to summarize simulated flow variability caused by parameter uncertainty over the entire service flow range. The probabilistic forecasts enable water resource operators to assess the risks associated with using model outputs to make abstraction decisions. Secondly, a water resource management model, which represents real-world operational constraints such as reservoir operation rules, abstraction licence conditions, available storage volumes and pump and water main capacities, was devised and integrated with the flow forecast model. This was found to enable effective implementation of real-time abstraction management scheme that helps water resource operators to quantify daily volume of water available for abstraction based on a specified level of risk. The real-time abstraction management approach developed in this study has contributed towards addressing the challenges of using flow forecast models to improve surface water abstraction decisions. An investigation of implications of the developed approach on water resources in the study catchment has demonstrated the significant role that the scheme can play in terms of recovering reservoir levels during dry periods.

The significant role runoff plays in transporting pollutants in catchments and the occurrences of peak pollutant levels at catchment outlets following rainfall events have been covered in a wide range of studies in the literature. However, most existing catchment scale pollutant models do not sufficiently represent the short-term

dynamics involved in runoff based pesticide generation and transport processes, which is mainly responsible for the occurrences of short-term peak pesticide concentrations at catchment outlets following rainfall events. This is mainly due to lack of high resolution water quality data and challenges associated with representing spatiotemporally variable distribution of pollutant generation and transport in catchments. High resolution water quality data collected in this study and the development of physically distributed metaldehyde prediction model have enabled to address these challenges. Findings of this research showed that runoff travel time based approaches can be combined with the increasing availability of catchment scale spatial datasets to quantify and understand catchment dynamics that drive short term fluctuations of pollutant concentrations following rainfall events. The approach has also showed that the likelihood of a high risk land in a catchment contributing to pollutant peak levels varies depending on the density of high risk areas available within the same travel time isochrones. This can help catchment management planners to identify and focus on high risk areas that frequently contribute to peak pollutant concentration levels arriving at catchment outlets. Spatial variation of rainfall, which governs the distributions of runoff generating across the catchment, was also found to play a significant role in determining the behaviour of peak concentration levels at catchment outlets.

Data on build-up of diffuse pollutants on high risk areas throughout the catchment are often difficult to acquire. As a result, catchment average data is commonly used despite the high spatial distribution of pollutant build-ups throughout the catchment. The physically distributed and grid based structure of the metaldehyde model enabled implementation of a Monte Carlo based spatial uncertainty analysis tool, which is used to assess uncertainties in the model associated with spatial distribution of metaldehyde build-up through applications on high risk areas. Results of the spatial input and parameter uncertainty analysis of the metaldhyde model in this study have showed that spatial variabilities in metaldehyde build-up across high risk areas in the catchment can cause significant uncertainty in predicted concentration levels. This indicates the need to carefully consider representations of pesticide

applications across the catchment in models, which are often assumed uniform throughout the catchment due to the difficulties associated with finding actual application data. It's also noted that uncertainties associated with metaldehyde transport throughout the catchment is mainly dominated by inaccurate representations of pollutant travel times in channels as compared to overland pollutant transport. Hence, improvements in channel travel time estimation technique combined with near real-time metaldehyde application data from farmers would significantly improve model performance. Identification of these inputs and parameters with the most contribution to uncertainties in predicted metaldehyde concentration helps to plan and design future field exercises to collect more data. Furthermore, probabilistic presentation of predicted metaldehyde concentrations are found to be important to investigate the various level of risks associated with different peak concentration levels and making risk aware abstraction decisions.

Ensuring compliance of water quality standards for metaldehyde requires supporting abstraction management with catchment management schemes and development of new treatment processes depending on the level of metaldehyde problem in the catchment. One of the catchment based approaches being adopted by catchment managers is replacing the use of metaldehyde with ferric phosphate, which is considered an alternative active ingredient to tackle slug problems in farmlands. However, ferric phosphate is considered less effective as compared to metaldehyde and it needs to be applied at higher rate, which results in its use to be more expensive than that of metaldehyde. In addition, a range of on-farm mitigation measures are being employed by water utilities in UK with a view to cost effectively reduce metaldehyde peaks in water courses whilst maintaining food production capacity. The metaldehyde model developed in this study can help to evaluate how effective different catchment management measures would be to reduce metaldehyde peak concentrations at the catchment outlet with a view to informing planning of catchment management schemes.

7.3 Conclusions

The research presented in this thesis has focused on developing a new real-time abstraction management scheme for surface water abstraction management. A hydrological forecasting and water resource management model are developed and integrated with a view to enable to abstract more water in the study catchment. Uncertainty associated with the model is also analysed and presented to help in making risk aware decisions. By dynamically linking abstraction volumes to actual availability of water in the source, the developed scheme helps to make surface water abstraction management systems ready for various challenges associated with climate change and increasing demand from a growing population.

A metaldehyde prediction model is also developed in this study that integrates spatially and temporally disaggregated runoff generation, routing and build-up/wash-off using a runoff travel time based approach in a GIS environment. With relatively few parameters, more practical model structure and quick simulation time, the developed model provides a suitable level of complexity for operational purposes. Given the inability of existing treatment techniques to remove high metaldehyde levels from water and the absence of direct metaldehyde detection methods, the model developed in this study provides a cost-effective and sustainable solution.

7.4 Limitations and future works

Application of the metaldehyde prediction model to inform abstraction management requires accurate representations of land use data across the catchment. Comparison of historical land use data in the study catchment has indicated the annual occurrences of land use changes across the study catchment. However, there is usually a delay in providing updated land use data which causes considerable amount of uncertainties associated with the extent and location of high risk areas used in the model for a particular metaldehyde application season. STW are currently looking into ways to acquire timely updated land use data during metaldehyde application seasons for the catchments where the metaldehyde model is planned to be applied. As discussed in chapter 4, a number of assumptions were necessary during the

development of the metaldehyde prediction model. Each of these assumptions could potentially introduce errors into the analysis. This is particularly true for the uniform metaldehyde application used across all high risk areas in the deterministic model. It is therefore advised to use the probabilistic metaldehyde prediction model to enable making risk aware abstraction decisions. Furthermore, sensitivity of the metaldehyde model is analysed in this study using five model input and parameters that are associated with metaldehyde build-up and runoff travel time processes. However, a full sensitivity analysis on a wider range of model parameters would be beneficial to assess the model's response to variations in parameters associated with a range of catchment and metaldehyde characteristics (e.g. K_{oc}). The nature of the organic compound (metalydyde) such as solubility and low sorption coefficient combined with the focus on forecasting short term fluctuations in response to rainfall events has been the drive for the development of runoff travel time based metaldehyde prediction approach in this study. Thus, further applications of the model for simulating the transport of other organic compounds needs to carefully investigate the property of the compounds and accordingly incorporate additional techniques into the model to enable accurate representation of transport processes. Furthermore, complete understanding of the transport and fate of pesticide in catchments requires consideration of numerous processes such as groundwater transport and reaction/degradation processes. Thus, applications and interpretations of the metaldehyde model outputs should be used with in context.

SCS-CN method has widespread application throughout the world and has earned a solid reputation as a well-established rainfall-runoff technique mainly due to its computational simplicity and use of accessible catchment data (Li et al. 2015). However, the method also has a number of limitations mainly owing to its empirical origins (Beven 2012; Burns et al. 2016). The SCS-CN method assumes areas with similar curve numbers as homogenous surfaces that exhibit spatially uniform runoff process (Hawkins et al. 2001). Moreover, the method doesn't provide a wide range of choice to adjust curve numbers based on antecedent soil moisture conditions, which may result in undermining effects of antecedent soil moisture conditions on runoff

generations (Boughton 1989). The use of high resolution data combined with the physically distributed technique incorporated in the metaldehyde model has helped to minimize the impact of these limitations on the metaldehyde model accuracy. The wide ranging distribution of available soil moisture capacities throughout the catchment and the variation in their runoff generating capabilities during rainfall events has been highlighted well in the PDM model as discussed in section 3.2.2. However, the PDM model represents the distribution in runoff generation across the catchment without any reference to the location of runoff generating areas in the catchment, whereas location of runoff generating areas is considered to play a significant role in simulating concentration levels of diffuse pollutants at the catchment outlet. The use of radar rainfall and development of the grid based spatially distributed catchment scale metaldehyde model in this study has enabled to provide a suitable catchment scale metaldehyde prediction model that represents both the location and rate of runoff generation.

Drain flows during rainfall events are known to significantly contribute to transportation of diffuse pollutants from farmlands to water courses in catchments that have network of tile drains (Harris & Catt 2006; Granger et al. 2010; Tiktak et al. 2012). This is mainly due to preferential water flow through macropores, which enables runoff water bypassing the soil matrix and reaching to tile drains below the ground (Y. Yuan et al. 2001; Tang et al. 2012). Section 2.6.1 above has discussed the wide range of literatures that highlighted the importance of drain flows in transporting diffuse pollution from farmlands to channel networks. In this study, the runoff computed using the SCS-CN method is assumed to be the source for rainfall event water transported at field levels both through overland (surface runoff) and drain flows. In addition, the travel time computation at field level (overland flow travel time) presented in section 4.2.4 is assumed to represent travel time of event water both in surface runoff and drains. In comparison to flow length along channel networks at catchment scale, the relatively short distances runoff flow covers at filed level is very short and is unlikely to cause significant impact on the transport of diffuse pollution as discussed in sections 5.3.2 and 5.3.3. Thus, the assumption

adopted in the development of metaldehyde model in this study to jointly represent overland and drain flow is considered a reasonable approach.

Below are several recommendations for future research to be undertaken in this area:

Future data collections and model improvement: Development of a suitable scheme for collection of near real-time metaldehyde application data from farmers can help to further improve the performance of the metaldehyde model. This can be achieved by 1) Assessing and devising a suitable technique for gathering near-real time pesticide application data from farmers. For example, a mobile application that enables to collate the exact location of farm land and time of pesticide application can be developed. 2) Developing a suitable method to update the spatially distributed model input data based on newly obtained pesticide application data with a view to improve model prediction accuracy.

Moreover, spatially distributed sampling using multiple samplers installed at optimum locations throughout the catchment can also be used to provide data that can be used for development of detailed sub-catchment scale models. Outputs from sub-catchment scale models can then be routed along the channel networks to provide metaldehyde concentrations arriving at catchment outlets. The detailed sampling data at catchment level can also be used to identify sub-catchments that suffer from high metaldehyde levels in a particular application season, which helps to inform the planning and implementations of catchment management practices. However, distributed sampling at sub-catchment level using multiple samplers can be expensive and time taking.

Enable effective catchment management practices: The techniques incorporated in the metaldehyde prediction model provide potentials to evaluate the likelihood of high risk farmland in the catchment to contribute to peak metaldehyde concentrations at catchment outlets. This involves integrated assessment of travel time and distributions of risk map locations throughout the catchment. Enabling the identification of high risk areas that frequently contribute to peak concentrations at

the outlet of the catchment on a particular pesticide application season can help to set out well targeted and effective catchment management practices. In addition to its contribution towards reducing peak concentration levels arriving at abstraction sites, it also generates immense benefits in terms of saving time, cost and resources used in catchment management.

Extend the pollutant model to capture a range of pesticides: The relatively simple structure and operationally suitable complexity of the metaldehyde model mean that, further improvements of the model to predict other compounds can generate multiple benefits. To enable this, properties of other compounds of interest needs to be investigated first. Then suitable techniques (e.g. sediment transport, reactions/ degradations, groundwater process) need to be incorporated into the model accordingly.

Enable real-time data assimilation technique: Real-time modelling provides an opportunity to incorporate newly observed data to models with a view to improve model predictions, which is known as data assimilation. In the presence of real-time metaldehyde monitoring at abstraction sites, a suitable data assimilation technique can be developed to optimally combine collected real-time meatldehyde data with model outputs to assess model errors and use the feedback to update certain states of the model. In addition to high costs associated with online monitoring techniques (costs associated with techniques that are currently in trials is discussed in chapter 6), their benefit to abstraction management is restricted by issues associated with practical operation of abstraction management. However, integration of these techniques with the metaldehyde model provides opportunities to reduce uncertainties in model forecasts while enabling to provide sufficient prediction time for operational abstraction management.

Optimisation of grab sampling: Data collected by grab sampling methods are commonly used by the water industry to monitor pesticide pollutant levels in catchments to help in the planning and development of catchment management strategies. Typically, grab samples are collected with days or weeks apart and

pesticide concentrations in between sample collection times are unknown. The ability of collected data to reflect the actual conditions at the study sites determine the validity of conclusions made based on the data. The techniques used in the metaldehyde model combined with high resolution water quality data can be used to identify optimal locations and timings for grab sampling collection. A suitable This involves the development of techniques to combine rainfall variability captured by radar rainfall data and runoff travel time technique with a view to identify strategical locations and timings to capture metaldehyde peaks using grab samples. This enables to address concerns regarding the suitability of grab sampling data to develop and evaluate catchment management practices.

References

Abbott, M.B. et al., 1986. An introduction to the European Hydrological System - System Hydrologique European, "SHE", 1: History and philosophy of a physically-based distributed modelling system. *Journal of Hydrology*, 87, pp.45–59.

Adediran, G.A., 2015. *Hydrological Forecasting with Radar and the Probability Distributed Hydrological Model (PDM)*, Universal-Publishers.

Agrawal, N. & Desmukh, T.S., 2016. Rainfall Runoff Modeling using MIKE 11 Nam – A Review. *International Journal of Innovative Science, Engineering & Technology*, 3(6)(6), pp.659–667.

Ajami, N.K., Hornberger, G.M. & Sunding, D.L., 2008. Sustainable water resource management under hydrological uncertainty. *Water Resources Research*, 44(11), pp.1–10.

Andréassian, V. et al., 2004. Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds. *Water Resources Research*, 40(5), pp.1–9.

Bach, M., Huber, A. & Frede, H., 2001. Input pathways and river load of pesticides in Germany -a national scale modelling assessment. *Water Sci. Technol*, 43(5), pp.261–268.

Bell, V.A., Carrington, D.S. & Moore, R.J., 2001. *Comparison of Rainfall-Runoff Models for Flood Forecasting, Part 2: Calibration and evaluation of models*, Bristol.

Bell, V.A. & Moore, R.J., 1998a. A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. *Hydrology and Earth System Sciences Discussions*, 2(2/3), pp.265–281.

Bell, V.A. & Moore, R.J., 1998b. A grid-based distributed flood forecasting model for use with weather radar data: Part 1. Formulation. *Hydrology and Earth System Sciences Discussions*, 2(2/3), pp.265–281.

Benaman, J., Shoemaker, C.A. & Haith, D.A., 2005. Calibration and Validation of Soil and Water Assessment Tool on an Agricultural Watershed in Upstate New York. *Journal of Hydrologic Engineering*, 10(5), pp.363–374.

Benke, K.K., Lowell, K.E. & Hamilton, A.J., 2008. Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. *Mathematical and Computer Modelling*, 47(11–12), pp.1134–1149.

Berenzen, N. et al., 2005. A comparison of predicted and measured levels of runoff-

related pesticide concentrations in small lowland streams on a landscape level. *Chemosphere*, 58(5), pp.683–691.

Berthouex, P. Mac & Brown, L.C., 2002. *Statistics for environmental engineers*, Lewis Publishers.

Beven, K., 2001. How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, 5(1), pp.1–12.

Beven, K., 2012. *Rainfall-Runoff Modelling* 2nd ed. John Wiley & Sons Ltd, ed., Oxford.

Beven, K. & Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3), pp.279–298.

Beven, K. & Freer, J., 2001. Hydrology Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, 249, pp.11–29.

Bilondi, M.P., Abbaspour, K.C. & Ghahraman, B., 2013. Application of three different calibration-uncertainty analysis methods in a semi-distributed rainfall-runoff model application. *Middle-East Journal of Scientific Research*, 15(9), pp.1255–1263.

Bingner, R.L. & Theurer, F.D., 2005. *AnnAGNPS TECHNICAL PROCESSES*, Gaithersburg, MD.

Blume, T., Zehe, E. & Bronstert, A., 2007. Rainfall – runoff response, event-based runoff coefficients and hydrograph separation. *Hydrological Sciences Journal*, 52(5), pp.843–862.

Borah, D.K. et al., 2002. DWSM - a Dynamic Watershed Simulation Model. In *Mathematical models of small watershed hydrology and applications*. Water Resources Publications, pp. 113–166.

Borah, D.K., Bera, M. & Shaw, S., 2003. Water, sediment, nutrient, and pesticide measurements in an agricultural watershed in Illinois during storm events. *Transactions of the ASAE*, 46(3), pp.657–674.

Boughton, W.; C., 1989. Soil and Water Management and Conservation A Review of the USDA SCS Curve Number Method. *Aust. J. Soil Res*, 2(7), pp.511–534.

Ter Braak, C.J.F., 2006. A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: Easy Bayesian computing for real parameter spaces. *Statistics and Computing*, 16(3), pp.239–249.

Ter Braak, C.J.F. & Vrugt, J.A., 2008. Differential Evolution Markov Chain with snooker updater and fewer chains. *Statistics and Computing*, 18(4), pp.435–446.

Brater, E.F. & King, H.W., 1976. *Handbook of hydraulics for the solution of hydraulic engineering problems.*, New York, USA: McGraw-Hill Book Company.

Breach, B., 2010. *EU drinking water legislation, european glyphosate environmental information source*, UK.

Briggs, A.H. et al., 2012. Model parameter estimation and uncertainty: A report of the ISPOR-SMDM modeling good research practices task force-6. *Value in Health*, 15(6), pp.835–842.

Brocca, L. et al., 2013. Assimilation of satellite soil moisture data into rainfall-runoff modelling for several catchments worldwide. *International Geoscience and Remote Sensing Symposium (IGARSS)*, pp.2281–2284.

Bundschuh, M., Goedkoop, W. & Kreuger, J., 2014. Evaluation of pesticide monitoring strategies in agricultural streams based on the toxic-unit concept - Experiences from long-term measurements. *Science of the Total Environment*, 484(1), pp.84–91.

Burgman, M.A., 2005. *Risks and decisions for conservation and environmental management*, Cambridge University Press.

Burn D.H. et al, 1999. Short term streamflow forecasting using artificial neural networks. *J. Hydrol.*, 214(1--4), pp.32–48.

Burnash, R.J.C., 1995. The NWS River Forecast System - Catchment Modeling. *In Computer models of watershed hydrology, Water Resources Publications, Littleton, CO*, pp.311–366.

Burns, M.A. et al., 2016. Beyond the SCS-CN method: A theoretical framework for spatially lumped rainfall-runoff response. *Water Resources Research*, 52(51), pp.5974–5997.

Cabus, P., 2008. River flow prediction through rainfall-runoff modelling with a probability-distributed model (PDM) in Flanders, Belgium. *Agricultural Water Management*, 95, pp.859–868.

Campbell, E.P., 2001. A Markov chain Monte Carlo scheme tor parameter estimation and inference in conceptual rainfall-runoff modeling. , 37(4), pp.937–947.

Canal and River Trust, 2015. Putting the water into waterways, Water resources strategy 2015-2020.

Carvalho, F.P., 2017. Pesticides, environment, and food safety. *Food and Energy Security*, 6(2), pp.48–60.

Casale, R. & Margottini, C., 2012. *Floods and landslides : integrated risk*

assessment, Springer Science & Business Media.

Casale, R. & Margottini, C., 1999. *Floods and landslides : integrated risk assessment*, Springer.

Castle, G.D. et al., 2017. Review of the molluscicide metaldehyde in the environment. *Environ. Sci.: Water Res. Technol*, 3(3), pp.415–428.

CEH, 2012. *An overview of the 2010-12 drought and its dramatic termination*, Available at: http://nora.nerc.ac.uk/20083/.

Clark, M.P. et al., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, 31, pp.1309–1324.

Cloke, H.L. & Pappenberger, F., 2009. Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3–4), pp.613–626.

Cosgrove, W.J. & Loucks, D.P., 2015. Water management: Current and future challenges and research directions. , pp.4823–4839.

Crespo, P.J. et al., 2011. Identifying controls of the rainfall–runoff response of small catchments in the tropical Andes (Ecuador). *Journal of Hydrology*, 407(1–4), pp.164–174.

Crommentuijn, T. et al., 2000. Maximum permissible and negligible concentrations for some organic substances and pesticides. *Journal of Environmental Management*, 58(4), pp.297–312.

D. N. Moriasi et al., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Transactions of the ASABE*, 50(3), pp.885–900.

D'Arcy, B.J. et al., 1998. Initiatives to tackle diffuse pollution in the UK. *Water Science and Technology*, 38(10 pt 8), pp.131–138.

DEFRA, 2012. *Draft Water Bill*, London.

DEFRA, 2014. *Making the Most of Every Drop Consultation on Reforming the Water Abstraction Management System*, London.

DEFRA, 2011a. *The Natural Choice: securing the value of nature*, London.

DEFRA, 2011b. *Water for Life*, London.

Devia, G.K., Ganasri, B.P. & Dwarakish, G.S., 2015. A Review on Hydrological Models. *Aquatic Procedia*, 4(Icwrcoe), pp.1001–1007.

Dietz, T., 2003. What is a Good Decision? Criteria for Environmental Decision Making. *Human Ecology Review*, 10, pp.33–39.

Dils, R.M. & Heathwaite, A.L., 1999. The controversial role of tile drainage in phosphorus export from agricultural land. *Water Science and Technology*, 39(12), pp.55–61.

Donigian, A. S., J. et al., 1995. *Hydrological Simulation Program - Fortran (HSPF).*, Water Resources Publications.

Dooge, J.C.I., 1973. Linear Theory of Hydrologic Systems. *Tech. Bull. 1468, Agric. Res. Service, US Dept. Agric., Washington*, p.327.

Dooge, J.C.I. & O'Kane, P., 2003. *Deterministic Methods in Systems Hydrology: IHE Delft Lecture Note Series*, CRC Press.

Du, J. et al., 2009. Development and testing of a new storm runoff routing approach based on time variant spatially distributed travel time method. *Journal of Hydrology*, 369(1–2), pp.44–54.

EA, 2011. *The case for change - current and future water availabilty*, Bristol.

Eagleson, P.S., 1970. *Dynamic Hydrology*, New York, USA: McGraw-Hill.

EC, 1998. COUNCIL DIRECTIVE 98/83/EC of 3 November 1998 on the quality of water intended for human consumption. *Official Journal of the European Communities*, L 330, pp.32–54.

Engman, E.T., 1986. Roughness Coefficients for Routing Surface Runoff. *Journal of Irrigation and Drainage Engineering*, 112(1), pp.39–53.

Erfani, T., Binions, O. & Harou, J.J., 2014. Protecting environmental flows through enhanced water licensing and water markets. *Hydrology and Earth System Sciences Discussions*, 11, pp.2967–3003.

Eschenbach, T.G., 1992. Spiderplots versus Tornado Diagrams for Sensitivity Analysis. *Interfaces*, 22(6), pp.40–46.

Feyen, L. et al., 2000. Application of a distributed physically-based hydrological model to a medium size catchment. *Hydrology and Earth System Sciences*, 4(1), pp.47–63.

Feyen, L. et al., 2007. Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the LISFLOOD model. *Journal of Hydrology*, 332(3–4), pp.276–289.

FOCUS, 2000. "FOCUS groundwater scenarios in the EU review of active substances" Report of the FOCUS Groundwater Scenarios Workgroup, EC

Document Reference Sanco/321/2000 rev.2, 202pp.

FOCUS, 2001. “FOCUS Surface Water Scenarios in the EU Evaluation Process under 91/414/EEC” Report of the FOCUS Working Group on Surface Water Scenarios EC Document Reference SANCO/4802/2001-rev2,245 pp.

Gamerman, D., 1997. Sampling from the posterior distribution in generalized linear mixed models. *Statistics and Computing*, 7, pp.57–68.

Gan, T.Y. & Biftu, G.F., 1996. Automatic calibration of conceptual rainfall-runoff models: Optimization algorithms, catchment conditions, and model structure. *Water Resources Research*, 32(12), pp.3513–3524.

Geissen, V. et al., 2015. Emerging pollutants in the environment: A challenge for water resource management. *International Soil and Water Conservation Research*, 3(1), pp.57–65.

Gelman, A. et al., 2014. *Bayesian Data Analysis*, Boca Raton, FL: CRC Press, Taylor & Francis Group.

Gelman, A. & Rubin, D.B., 1992. Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4), pp.457–511.

Gericke, O.J. & Smithers, J.C., 2014. Review of methods used to estimate catchment response time for the purpose of peak discharge estimation. *Hydrological Sciences Journal*, 59(11), pp.1935–1971.

Gleb, A., 1974. *Applied optimal estimation.*, M.I.T. Press.

Golding, B.W., 2000. Quantitative precipitation forecasting in the UK. *Journal of Hydrology*, 239, pp.286–305.

Goodson, J.M. et al., 2002. Riparian seed banks along the lower River Dove, UK: Their structure and ecological implications. *Geomorphology*, 47(1), pp.45–60.

Goswami, M. et al., 2005. Assessing the performance of eight real-time updating models and procedures for the Brosna River. *Hydrology and Earth System Sciences*, 9(4), pp.394–411.

Granger, S.J. et al., 2010. *Towards a holistic classification of diffuse agricultural water pollution from intensively managed grasslands on heavy soils* 1st ed., Elsevier Inc.

Gray, N.F., 2008. *Drinking Water Quality: Problems and Solutions*, New York: John Wiley & Sons Ltd.

Grayson, R., Kay, P. & Foulger, M., 2008. The use of GIS and multi-criteria evaluation (MCE) to identify agricultural land management practices which

cause surface water pollution in drinking water supply catchments. *Water Science and Technology*, 58(9), pp.1797–1802.

Green, P.J. & Mira, A., 2001. Delayed Rejection in Reversible Jump Metropolis-Hastings. *Biometrika*, 88(4), pp.1035–1053.

Haario, H. et al., 2006. DRAM: Efficient adaptive MCMC. *Statistics and Computing*, 16(4), pp.339–354.

Haario, H., Saksman, E. & Tamminen, J., 2001. An adaptive Metropolis algorithm. *Bernoulli*, 7(2), pp.223–242.

Hall, J.W. & Borgomeo, E., 2013. Risk-based principles for defining and managing water security. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 371(2002), p.20120407.

Hankin, B. et al., 2016. Catchment-scale sensitivity and uncertainty in water quality modelling. *Hydrological Processes*, 30(22), pp.4004–4018.

Harris, G.L. & Catt, J. a., 2006. Overview of the studies on the cracking clay soil at Brimstone Farm, UK. *Soil Use and Management*, 15(4), pp.233–239.

Hasting, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), pp.97–109.

Hawkins, R.H. et al., 2001. Another Look at SCS-CN Method. *Journal of Hydrologic Engineering*, 6(5), pp.451–452.

Heathwaite, A. L., Quinn, P.F. & Hewett, C.J.M., 2005. Modelling and managing critical source areas of diffuse pollution from agricultural land using flow connectivity simulation. *Journal of Hydrology*, 304, pp.446–461.

Henonin, J. et al., 2013. Real-time urban flood forecasting and modelling – a state of the art. *Journal of Hydroinformatics*, 15(3), p.717.

Hjelmfelt, A.T., 1991. Investigation of Curve Number Procedure. *Journal of Hydraulic Engineering*, 117(6), pp.725–737.

Holvoet, K. et al., 2008. Modifications to the SWAT code for modelling direct pesticide losses. *Environmental Modelling and Software*, 23(1), pp.72–81.

Holvoet, K.M. a, Seuntjens, P. & Vanrolleghem, P. a., 2007. Monitoring and modeling pesticide fate in surface waters at the catchment scale. *Ecological Modelling*, 209, pp.53–64.

Hough, M.N. & Jones, R.J.A., 1997. The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0-an overview. *Hydrology and Earth System Sciences*, 1(2), pp.227–239.

Huber, A., Bach, M. & Frede, H.G., 1998. Modelling pesticide losses with surface runoff in Germany. *Science of the Total Environment*, 223, pp.177–191.

Huber, A., Bach, M. & Frede, H.G., 2000. Pollution of surface waters with pesticides in Germany: Modelling non-point source inputs. *Agriculture, Ecosystems and Environment*, 80, pp.191–204.

Hutton, C.J. et al., 2014. Real-time data assimilation in urban rainfall-runoff models. *Procedia Engineering*, 70, pp.843–852.

Jansen, M.J.W., 1999. Analysis of variance designs for model output. *Computer Physics Communications*, 117(1), pp.35–43.

Jazwinski, A.H., 1970. *Stochastic processes and filtering theory*, Academic Press.

Jonsdottir, H., Madsen, H. & Palsson, O.P., 2006. Parameter estimation in stochastic rainfall-runoff models. *Journal of Hydrology*, 326(1–4), pp.379–393.

Jorgeson, J. & Julien, P., 2005. Peak flow forecasting with radar precipitation and the distributed model casc2d. *Water International*, 30(1), pp.40–49.

Kannan, N. et al., 2006. Pesticide Modelling for a Small Catchment Using SWAT-2000. *Journal of Environmental Science and Health, Part B*, 41(7), pp.1049–1070.

Kay, P. & Grayson, R., 2014. Using water industry data to assess the metaldehyde pollution problem. *Water and Environment Journal*, 28(3), pp.410–417.

Keith, L.H., 1990. Environmental sampling: A summary. *Environmental Science and Technology*, 24(5), pp.610–617.

Kenaga, E.E., 1980. Predicted bioconcentration factors and soil sorption coefficients of pesticides and other chemicals. *Ecotoxicology and Environmental Safety*, 4, pp.26–38.

Kendon, M., Marsh, T. & Parry, S., 2013. The 2010-2012 drought in England and Wales. *Weather*, 68(4), pp.88–95.

Kilburn, D., 2010. Slug pellets and rainfall have vital link to pollution. Farmers Guardian. Available at: http://www.farmersguardian.com/slug-pellets-and-rainfall-have-vital-link-to-pollution/31894.article [Accessed June 6, 2014].

Knapp, H.V., Durgunoglu, A. & Ortel, T.W., 1991. A Review of Rainfall-Runoff Modeling for Stormwater Management. *Illinois State Water Survey, Hydrology Division*, (SWS Contract Report 516), p.96.

Köhne, J.M., Köhne, S. & Šimůnek, J., 2009. A review of model applications for structured soils: a) Water flow and tracer transport. *Journal of Contaminant*

Hydrology, 104(1–4), pp.4–35.

Kuczera, G. & Mroczkowski, M., 1998. Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research*, 34(6), pp.1481–1489.

Kuivila, K.M. & Foes, C.G., 1995. Concentrations, Transport and Biological Effects of Dormant Spray Pesticides in the San Francisco Estuary, California. *Environmental Toxicology and Chemistry*, 14(7), pp.1141–1150.

Laloy, E. & Vrugt, J.A., 2012. High-dimensional posterior exploration of hydrologic models using multiple-try DREAM (ZS) and high-performance computing. *Water Resources Research*, 48(1), pp.1–18.

Lan Anh, N.T. et al., 2008. An evaluation of three lumped conceptual rainfall-runoff models at catchment scale. *The 13th World Water Congress*, p.12.

Larsbo, M. & Jarvis, N., 2003. *MACRO 5.0 : a model of water flow and solute transport in macroporous soil : technical description*, Department of Soil Sciences, Swedish Univ. of Agricultural Sciences.

Lee, H. et al., 2012. Variational assimilation of streamflow into operational distributed hydrologic models: Effect of spatiotemporal scale of adjustment. *Hydrology and Earth System Sciences*, 16, pp.2233–2251.

Lee, H., Seo, D.J. & Koren, V., 2011. Assimilation of streamflow and in situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states. *Advances in Water Resources*, 34, pp.1597–1615.

Li, C. et al., 2010. Determination of Metaldehyde in Water by SPE and UPLC – MS – MS. *Chromatographia*, 72(9-10), pp.987–991.

Li, J. et al., 2015. Two universal runoff yield models: SCS vs. LCM. *Journal of Geographical Sciences*, 25(3), pp.311–318.

Lindblom, E., Ahlman, S. & Mikkelsen, P.S., 2007. How uncertain is model-based prediction of copper loads in stormwater runoff? *Water Science and Technology*, 56(11) , pp.65–72.

Te Linde, A.H. et al., 2008. Comparing model performance of two rainfall-runoff models in the Rhine basin using different atmospheric forcing data sets. *Hydrology and Earth System Sciences*, 12(3), pp.943–957.

Loucks, D.P. et al., 2005. *Water Resources Systems Planning and Management and Applications: An Introduction to Methods, Models and Applications*, Paris: UNESCO Publishing.

Loucks, D.P. & Beek, E. Van, 2005. *WATER RESOURCES SYSTEMS PLANNING AND MANAGEMENT*, Paris: UNESCO Publishing.

Lu, Q. et al., 2017. Modelling Metaldehyde in Catchments: A River Thames Case-Study. *Environ. Sci.: Processes Impacts*, pp.586–595.

Luo, Y. & Zhang, M., 2009. Management-oriented sensitivity analysis for pesticide transport in watershed-scale water quality modeling using SWAT. *Environmental Pollution*, 157(12), pp.3370–3378.

Mancini, M. & Rosso, R., 1989. Using GIS to assess spatial variability of SCS Curve Number at the basin scale. In *In New directions for surface water modelling, Proceedines of the Baltimore Symposium*. IAHS Publ.no.181, pp. 435–444.

Manley, R.E., 1978. Simulation of flows in ungauged basins / Simulation d'écoulement sur les bassins versants non jaugés. *Hydrological Sciences Bulletin*, 23(February 2015), pp.85–101.

Mara, T.A. et al., 2016a. A comparison of two Bayesian approaches for uncertainty quantification. *Environmental Modelling & Software*, 82, pp.21–30.

Mara, T.A. et al., 2016b. A comparison of two Bayesian approaches for uncertainty quantification. *Environmental Modelling & Software*, 82, pp.21–30.

Marshall, L., Nott, D. & Sharma, A., 2004. A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research*, 40(2), pp.1–11.

McIntyre, N. et al., 2005. Ensemble predictions of runoff in ungauged catchments. *Water Resources Research*, 41, pp.1–14.

McIntyre, N. et al., 2004. Numerical efficiency in Monte Carlo simulations - a case study of a river thermodynamic model. *Journal of Enviornmental Engineering*, 130(4), pp.456–464.

McIntyre, N., Wheater, H. & Lees, M., 2002. Estimation and propagation of parametric uncertainty in environmental models. *Journal of Hydroinformatics*, 04(3), pp.177–198.

McKay, M.D., Beckman, R.J. & Conover, W.J., 1979. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2), pp.239–245.

McMillan, H.K. et al., 2013. Operational hydrological data assimilation with the recursive ensemble Kalman filter. *Hydrology and Earth System Sciences*, 17(2007), pp.21–38.

Melesse, A.M. & Graham, W.D., 2004. Storm runoff prediction based on a spatially

distributed travel time method utilizing remote sensing and GIS. *Journal of the American Water Resources Association*, 40(4), pp.863–879.

Met Office, 2003. 1 km Resolution UK Composite Rainfall Data from the Met Office Nimrod System. NCAS British Atmospheric Data Centre. Available at: http://catalogue.ceda.ac.uk/uuid/27dd6ffba67f667a18c62de5c3456350 [Accessed April 6, 2016].

Metaldehyde Stewardship Group, 2012a. *Act NOW – Get Pelletwise*, 1–2.

Metaldehyde Stewardship Group, 2012b. *Metaldehyde slug pellet stewardship, Promoting Responsible Metaldehyde Applications*, 1–5.

Mishra, S.K. & Singh, V.P., 1999. Another Look at SCS-CN Method. *Journal of Hydrologic Engineering*, 4(3), pp.257–264.

Montes, S., 1998. *Hydraulics of open channel flow*, Reston, Virginia, USA: ASCE Press.

Moore, J., 2016. *Literature Review: Tile Drainage and Phosphorus Losses from Agricultural Land, Stone Environmental, Inc.*,

Moore, R.J. et al., 1992. *Evaluation of Frontiers and Local Radar Rainfall Forecasts for Use in Flood Forecasting Models*, Wallingford.

Moore, R.J. et al., 2007. *Rainfall-runoff and other modelling for ungauged/low-benefit locations: Operational Guidelines (SC030227)*,

Moore, R.J., 2007. The PDM rainfall-runoff model. *Hydrology and Earth System Sciences*, 11(1), pp.483–499.

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. *Hydrological Sciences Journal*, 30, pp.273–297.

Moore, R.J. & Bell, V. a., 2002. Incorporation of groundwater losses and well level data in rainfall-runoff models illustrated using the PDM. *Hydrology and Earth System Sciences*, 6, pp.25–38.

Moore, R.J. & Bell, V.A., 2001. *Comparison of Rainfall-Runoff Models for Flood Forecasting*, Bristol.

Moore, R.J. & Clarke, R.T., 1981. A distribution function approach to rainfall runoff modeling. *Water Resources Research*, 17(5), p.1367.

Moradkhani, H. et al., 2005. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28(2), pp.135–147.

Moradkhani, H. & Sorooshian, S., 2009. General Review of Rainfall-Runoff Modeling: Model Calibration, Data Assimilation, and Uncertainty Analysis. In *Hydrological Modelling and the Water Cycle*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–24.

Morrison, J. et al., 2009. Water Scarcity & Climate Change: Growing Risks for Businesses and Investors. *A Ceres Report*, p.60.

Muleta, M.K. et al., 2013. Bayesian Approach for Uncertainty Analysis of an Urban Storm Water Model and Its Application to a Heavily Urbanized Watershed. *Journal of Hydrologic Engineering*, 18(10), pp.1360–1371.

Müller, K. et al., 2003. Hydrological characterisation of pesticide loads using hydrograph separation at different scales in a German catchment. *Journal of Hydrology*, 273, pp.1–17.

Neitsch, S. et al., 2002. *Soil and water assessment tool theoretical documentation, Version 2000*, TexasWater Resources Institute, College Station, TX.

NFU, 2013. Updated briefing on Metaldehyde stewardship. *NFU briefing*, PS088/13, pp.1–4.

Ng, H.Y.F. & Clegg, S.B., 1997. Atrazine and metolachlor losses in runoff events from an agricultural watershed: The importance of runoff components. *Science of the Total Environment*, 193(3), pp.215–228.

Nguyen, H.H. et al., 2017. Modelling the impacts of altered management practices, land use and climate changes on the water quality of the Millbrook catchment-reservoir system in South Australia. *Journal of Environmental Management*, 202, pp.1–11.

Noto, L.V., 2014. Exploiting the Topographic Information in a PDM-Based Conceptual Hydrological Model. *Journal of Hydrologic Engineering*, 19(6), pp.1173–1185.

NSRI, 2009. LandIS - Land Information System - Digital Soils Data. Available at: http://www.landis.org.uk/data/index.cfm [Accessed April 27, 2017].

O'Hagan, A., 2012. Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. *Environmental Modelling and Software*, 36, pp.35–48.

OFWAT, 2011. *From catchment to customer*, Birmingham.

OFWAT & EA, 2011. *The case for change – reforming water abstraction managemnet in England*,

Ordnance survey, 2017. detailed digital terrain model (DTM) of Great Britain, OS

Terrain 5.

Ort, C. et al., 2010. Sampling for pharmaceuticals and personal care products (PPCPs) and illicit drugs in wastewater systems: are your conclusions valid? A critical review. *Environmental science & technology*, 44(16), pp.6024–35.

Ouyang, W. et al., 2017. A review of diffuse pollution modeling and associated implications for watershed management in China. *Journal of Soils and Sediments*, pp.1–10.

Pappenberger, F. et al., 2006. Decision tree for choosing an uncertainty analysis methodology: a wiki experiment. *Hydrological Processes*, 20, pp.3793–3798.

Penning-Rowsell, E.C. et al., 2000. The benefits of flood warnings: Real but elusive, and politically significant. *Journal of the Chartered Institution of Water and Environmental Management*, 14, pp.7–14.

Petersen, L., Minkkinen, P. & Esbensen, K.H., 2005. Representative sampling for reliable data analysis: Theory of Sampling. *Chemometrics and Intelligent Laboratory Systems*, 77, pp.261–277.

du Plessis, A., 2017. Global Water Quality Challenges. In : Freshwater Challenges of South Africa and its Upper Vaal River. Springer Water. Springer, Cham, pp. 13–44.

Du Preez, L.H. et al., 2005. Seasonal exposures to triazine and other pesticides in surface waters in the western Highveld corn-production region in South Africa. *Environmental Pollution*, 135(1), pp.131–141.

Pullan, S.P. et al., 2016. Development and application of a catchment scale pesticide fate and transport model for use in drinking water risk assessment. *Science of the Total Environment*, 563–564, pp.434–447.

Quilbé, R. et al., 2006. Selecting a pesticide fate model at the watershed scale using a multi-criteria analysis. *Water Quality Research Journal of Canada*, 41(3), pp.283–295.

Rabiei, E. & Haberlandt, U., 2015. Applying bias correction for merging rain gauge and radar data. *Journal of Hydrology*, 522, pp.544–557.

Rabiet, M. et al., 2010. Assessing pesticide concentrations and fluxes in the stream of a small vineyard catchment - Effect of sampling frequency. *Environmental Pollution*, 158(3), pp.737–748.

Rajurkar, M.P., Kothyari, U.C. & Chaube, U.C., 2002. Artificial neural networks for daily rainfall—runoff modelling. *Hydrological Sciences Journal*, 47(6), pp.865–877.

Rangel-Peraza, J. et al., 2016. Sensitivity and uncertainty analysis on water quality modelling of Aguamilpa reservoir. *Journal of Limnology*, 75(No 1s).

Reckhow, K., 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision making. *Ecological Modelling*, 72(1–2), pp.1–20.

Reddy, P.J.R., 2006. *A text book of hydrology* 2nd ed., Laxmi Publications Pvt Limited.

Refshaard, J.C., Storm, B. & Singh, V.P., 1995. MIKE SHE. In *Computer Models of Watershed Hydrology*. Water Resources Publications, Colorado, USA, pp. 809–846.

Reichert, P., 2014. *Environmental Systems Analysis*, Zurich, Switzerland.

Reichert, P. & Borsuk, M.E., 2005. Does high forecast uncertainty preclude effective decision support? *Environmental Modelling and Software*, 20(8), pp.991–1001.

Reus, J., Middleton, D. & Bockstaller, C., 1999. Comparing environmental risk indicators for pesticides : results of the European CAPER project. *CLM (Netherlands)*, 426, pp.80–82.

Richards, R.P. & Baker, D.B., 1993. Pesticide concentration patterns in agricultural drainage networks in the lake Erie Basin. *Environmental Toxicology and Chemistry*, 12(1), pp.13–26.

Rogelis, M.C. & Werner, M., 2018. Streamflow forecasts from WRF precipitation for flood early warning in mountain tropical areas. *Hydrology and Earth System Sciences*, 22(1), pp.853–870.

Ryberg, K.R. & Gilliom, R.J., 2015. Trends in pesticide concentrations and use for major rivers of the United States. *Science of the Total Environment*, 538, pp.431–444.

Saltelli, A, 2002. Making best use of model valuations to compute sensitivity indices. *Computer Physics Communications*, 145, pp.280–297.

Sangchan, W. et al., 2012. Short-term dynamics of pesticide concentrations and loads in a river of an agricultural watershed in the outer tropics. *Agriculture, Ecosystems and Environment*, 158, pp.1–14.

Sawicka K., G.B.M.H., 2016. 'spup' – an R package for uncertainty propagation in spatial environmental modelling. *International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences*, 12, pp.275–282.

Seo, D.-J., Koren, V. & Cajina, N., 2003. Real-Time Variational Assimilation of Hydrologic and Hydrometeorological Data into Operational Hydrologic Forecasting. *Journal of Hydrometeorology*, 4, pp.627–641.

Seo, D.J. et al., 2009. Automatic state updating for operational streamflow forecasting via variational data assimilation. *Journal of Hydrology*, 367(3–4), pp.255–275.

Shamseldin, A.Y. & O 'connor, K.M., 2003. A consensus real-time river flow forecasting model for the Blue Nile River. *Water Resources Systems--Hydrological Risk, Management and Development*, (281), pp.82–89.

Shaw, E.M. et al., 2017. *Hydrology in Practice, Fourth Edition*, CRC Press.

Shi, X. et al., 2012. Comparing Nonlinear Regression and Markov Chain Monte Carlo Methods for Assessment of Prediction Uncertainty in Vadose Zone Modeling. *Vadose Zone Journal*, 11(4).

Siebert, J., 1999. *Conceptual runoff models - fiction or representation of reality? Acta Univ. Ups., Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 436. 52 pp.*, Illinois State Water Survey.

Sitterson, J. et al., 2017. *An Overview of Rainfall-Runoff Model Types An Overview of Rainfall-Runoff Model Types*, Athens, Georgia, US.

Smith, T.J. & Marshall, L.A., 2008. Bayesian methods in hydrologic modeling: A study of recent advancements in Markov chain Monte Carlo techniques. *Water Resources Research*, 44(12), pp.1–9.

Stone, W.W. & Wilson, J.T., 2006. Preferential Flow Estimates to an Agricultural Tile Drain with Implications for Glyphosate Transport. *Journal of Environment Quality*, 35(5), p.1825.

STW, 2018. *Draft Drought Plan 2019 - 2024*, Coventry, UK.

STW, 2017. Metaldehyde concnetration data analysed from samples collected in the River Leam catchment, Grab sampling data (2011-2017).

STW, 2014. *Severn Trent Water: Final Water Resources Management Plan 2014*, Coventry, UK|.

Szollosi-Nagy, A., 1976. INTRODUCTORY REMARKS ON THE STATE SPACE MODELING O F WATER RESOURCE SYSTEMS. *International institute for applied systems analysis*, (RM-76-73), p.81.

Taghavi, L., Merlina, G. & Probst, J.-L., 2011. The role of storm flows in concentration of pesticides associated with particulate and dissolved fractions as a threat to aquatic ecosystems - Case study: the agricultural watershed of Save river (Southwest of France). *Knowledge and Management of Aquatic Ecosystems*, (400), pp. 6-11.

Tang, X., Zhu, B. & Katou, H., 2012. A review of rapid transport of pesticides from

sloping farmland to surface waters: Processes and mitigation strategies. *Journal of Environmental Sciences*, 24(3), pp.351–361.

Teuling, A. J. et al., 2004. Links between river flow statistics and catchment characteristics - Implications for land surface schemes. , pp.102–104.

Thorndahl, S. et al., 2010. Uncertainty in Nowcasting of Radar Rainfall: a case study of the GLUE methodology. In *Erad 2010 - the Sixth European Conference on Radar in Meteorology and Hydrology*. p. 7.

Tierney, L., 1994. Markov Chains for Exploring Posterior Distributions. *Annals of Statistics*, 22(4), pp.1701–1762.

Tierney, L. & Mira, A., 1999. Some adaptive monte carlo methods for Bayesian inference. *Stat Med*, 18(17–18), pp.2507–2515. Available at: 0010474156.

Tiktak, A. et al., 2012. Leaching of plant protection products to field ditches in the Netherlands Development of a drainpipe scenario for arable land. *RIVM report 607407003*, p.108.

Todini, E., 2005. Rainfall-runoff models for real-time forecasting. *Encyclopedia of Hydrological Sciences*, p.28.

Tomassini, L. et al., 2007. Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods. *Journal of Climate*, 20(7), pp.1239–1254.

Uhlenbrook, S. et al., 1999. Prediction uncertainty of conceptual rainfall-runoff models caused by problems in identifying model parameters and structure. *Hydrological Sciences Journal*, 44(5), pp.779–797.

USDA, 1986. Urban Hydrology for Small Watersheds TR-55. *USDA Natural Resource Conservation Service Conservation Engeneering Division Technical Release 55*, p.164.

Vandenberghe, V. et al., 1987. Propagation of uncertainty in diffuse pollution into water quality predictions : application to the River Dender in Flanders , Belgium. , pp.347–354.

Vandenberghe, V., Bauwens, W. & Vanrolleghem, P.A., 2007. Evaluation of uncertainty propagation into river water quality predictions to guide future monitoring campaigns. *Environmental Modelling and Software*, 22(5), pp.725–732.

Vaze, J. et al., 2011. *(eWater Cooperative Research Center) Guidlines for rainfall-runoff modellingGuidelines for rainfall-runoff modelling: Towards best practice model application.*,

Vihola, M., 2012. Robust adaptive Metropolis algorithm with coerced acceptance rate. *Statistics and Computing*, 22(5), pp.997–1008.

Viviani, G., 2009. Parameter uncertainty analysis of water quality model for small river. , (July), pp.3201–3207.

Voge, W.H.F. and R.M., 2016. On the deterministic and stochastic use of hydrologic models. *Water Resources Research*, 52, pp.1–15.

Vrugt, J. A., 2003. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water Resources Research*, 39(8), pp.1201.

Vrugt, J. A. et al., 2009. Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, 10(3), pp.273–290.

Vryzas, Z. et al., 2009. Spatial and temporal distribution of pesticide residues in surface waters in northeastern Greece. *Water Research*, 43(1), pp.1–10.

Wang, D., Chen, Y. & Cai, X., 2009. State and parameter estimation of hydrologic models using the constrained ensemble Kalman filter. *Water Resources Research*, 45(11), pp.1–13.

Wang, Q. et al., 2013. A review of surface water quality models. *The Scientific World Journal*, 2013, p.231-768.

Ward, R.C., Loftis, J.C. & McBride, G.B., 1986. The "data-rich but information-poor" syndrome in water quality monitoring. *Environmental Management*, 10(3), pp.291–297.

Warren, N. et al., 2003. Pesticides and other micro-organic contaminants in freshwater sedimentary environments - A review. *Applied Geochemistry*, 18(2), pp.159–194.

Water Forum, 2015. Assessment of Severn Trent Water's engagement and the outcomes for customers in the company's business plan 2015-20. , (November 2013).

Water UK, 2013. Briefing paper on metaldehyde. *Water UK*, pp.1–5.

Webber, F., 2014. Biological Drinking Water Treatment : Challenges and Potential. *Wrc Wateronline*, (December, 2004).

Weerts, A.H. & El Serafy, G.Y.H., 2006. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models. *Water Resources Research*, 42(January), pp.1–17.

Wijesiri, B. et al., 2016. Assessing uncertainty in pollutant build-up and wash-off processes. *Environmental Pollution*, 212, pp.48–58.

Wilson, C. & Tisdell, C., 2001. Why farmers continue to use pesticides despite environmental, health and sustainability costs. *Ecological Economics*, 39(3), pp.449–462.

Wong, T.S.W., 2003. Comparison of celerity-based with velocity-based time-of-concentration of overland plane and time-of-travel in channel with upstream inflow. *Advances in Water Resources*, 26(11), pp.1171–1175.

Wong, T.S.W., 1995. Time of concentration formulae for planes with upstream inflow. *Hydrological Sciences Journal*, 40(5), pp.663–666.

Wu, Q. et al., 2004. Influences of suspended particles on the runoff of pesticides from an agricultural field at Askim, SE-Norway. *Environmental Geochemistry and Health*, 26, pp.295–302.

Xu, C., 2002. *Hydrologic Models*, Uppsala, Sweden.

Y. Yuan et al., 2001. Modified Scs Curve Number Method for Predicting Subsurface Drainage Flow. *Transactions of the ASAE*, 44(6).

Young, A. & Reynard, N.S., 2004. River Flow Simulation within Ungauged Catchments ; the Utility of Regionalised Models River Flow Simulation within Ungauged Catchments ; *Proceedings of the iEMS Second Biennial Meeting: '"Complexity and Integrated Resources Management"'. Copenhagen. International Environmental Modelling and Software Society*.

Young, A.R., 1997. *Bench marking rainfall runoff models within the Anglian Region: Final Report*, Wallingford.

Young, R.A. et al., 1989. AGNPS: A nonpoint-source pollution model for evaluating agricultural watersheds. *Journal of Soil and Water Conservation*, 44(2), pp.168–173.

Zhang, C., Wang, R. & Meng, Q., 2015. Calibration of Conceptual Rainfall-Runoff Models Using Global Optimization. *Advances in Meteorology*, 2015, p.e545376.

Zhao, X. et al., 2011. Key uncertainty sources analysis of water quality model using the first order error method. *International Journal of Environmental Science & Technology*, 8(1), pp.137–148.

Zhao, X. et al., 2011. Key uncertainty sources analysis of water quality model using the first order error method. *International Journal of Environmental Science and Technology*, 8(1), pp.137–148.