Modelling the cumulative impacts of future coal mining and coal seam gas extraction on river flows: applications of methodology

Neil R. Viney[a], David A. Post[a], Yongqiang Zhang[ab], Fazlul Karim[a], Santosh K. Aryal[a], Mat Gilfedder[c], Jorge Peña Arancibia[a], Bill Wang[a], Ang Yang[a], Ramneek Singh[a], Xiaogang Shi[ad], Russell S. Crosbie[e], Luk J.M. Peeters[e,f], Natasha F. Herron[a], Jai Vaze[a], Steve Marvanek[e], Debbie Crawford[a], Adam Ramage[g], Adrian Dehelean[h], Dennis Gonzalez[e], Lingtao Li[a], Tim Evans[h]

[a] CSIRO Land and Water, GPO Box 1700, Canberra, ACT 2601, Australia

[b] Now at: Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

[c] CSIRO Land and Water, Brisbane, Australia

[d] Now at: School of Interdisciplinary Studies, University of Glasgow, United Kingdom

[e] CSIRO Land and Water, Adelaide, Australia

[f] CSIRO Mineral Resources, Adelaide, Australia

[g] CSIRO Land and Water, Kensington, Australia

[h] Geoscience Australia, Canberra, Australia

Correspondence: david.post@csiro.au

Article info

This paper is dedicated to the memory of Neil Viney who passed away on 30 August, 2020.

Highlights

- Modelling the impacts of coal mining and coal seam gas extraction on streamflow
- Reductions in annual streamflow are proportional to the area of coal mine coverage
- For coal seam gas, reductions in annual streamflow are proportional to well density
- A zone of potential hydrological change is identified, wherein impacts may be felt

Abstract

This manuscript presents examples of the modelling of the impacts of coal mining and coal seam gas extraction on streamflow in five study catchments in Australia. The manuscript includes details on data preparation and model set-up and calibration. The modelling methodology enables the prediction of cumulative impacts from multiple future coal resource developments and distributes these predictions at multiple locations in the landscape. It is framed in terms of a structured uncertainty analysis to provide information

on the likelihoods and potential ranges of various impacts. Also included is a qualitative uncertainty analysis which subjectively assesses the likely impact on model results of various assumptions made during the modelling procedure. Model results suggest that, in the Australian context, maximum percentage reductions in annual streamflow are approximately commensurate with the proportion of coal mine coverage. In coal seam gas fields, reductions in annual streamflow are proportional to well density. The manuscript goes on to demonstrate how these modelling results can be used to identify a zone of potential hydrological change within a catchment. This zone delineates those parts of the landscape where water-dependent landscape classes and assets may be vulnerable to change associated with changes in the streamflow regime. A corollary of this is that any parts of the landscape outside the zone of potential hydrological change are unlikely to be affected by coal resource development.

Keywords: hydrological modeling; coal mining; coal seam gas extraction; cumulative impacts

1. Introduction

Over the past century, coal has been widely used around the world as a fuel source for transport and industry and to generate electricity. Modern mining methods for coal include surface and underground mining. Surface mining (also known as open cut, open cast, or mountaintop removal mining) is appropriate for seams that occur sufficiently close to the land surface that it is economical to remove the overburden in order to access the seam. Deeper coal seams require underground mining, and this is typically achieved by longwall or bord and pillar methods. Longwall mining features the controlled collapse of overlying rock once the coal has been extracted.

In recent decades, the commercial extraction from coal seams of adsorbed methane (also known as coal seam gas or coalbed methane) has increased around the world. This is

typically achieved by drilling wells to intersect the target seam and sometimes requires hydraulic fracturing to stimulate gas flow.

Coal mining and coal seam gas extraction can result in adverse hydrological and ecological consequences. These potentially include changes to surface topography, the soil profile, vegetation cover, water quantity and quality, and air quality. In particular, coal mining produces significant amounts of waste material (overburden, spoil, tailings) that is usually stored on site and the interactions of this waste with incidental water can give rise to increased erosion, acid mine drainage and high concentrations of dissolved solids, and can seep into waterways and aquifers.

The literature on applications of hydrological modelling to assess the impacts of coal resource development appears to be quite sparse. Ping et al. (2017) used MIKE-SHE to evaluate the impact of coal mining on river flows in China. They calibrated the model to pre-mine conditions (using one response gauge and a monthly time step) then compared simulations during a subsequent mining period with observed river flows. The results suggest that each ton of raw coal reduces river flow by 2.87 m^3, 8% of which is due to reductions in surface runoff and 92% due to reductions in baseflow. We note that the reliance of this methodology on observed flows means that it is not amenable to prediction of future developments. A similar methodology using the YRWBM model was reported by Guo et al. (2017) and suggested an average annual flow reduction of about 60% in a large catchment with a surface coal mine coverage of about 29%.

The focus of the current paper is very much on the cumulative impacts of multiple proposed coal resource developments and the chosen methodology reflects this. A different approach is called for to investigate the potential impacts of individual mines and coal seam gas developments, and this approach is common in Environmental Impact Assessments lodged by mine proponents. The current study does not seek to replace these approaches, but

rather to present a methodology specifically designed to assess the cumulative impacts of multiple coal resource development in close proximity.

Viney et al. (2021) presented a methodological framework for modelling the impacts of future coal mining and coal seam gas extraction on surface water resources and streamflow. This methodology enables the prediction of cumulative impacts from multiple coal resource developments and distributes these predictions at multiple locations in the landscape. It is framed in terms of a structured uncertainty analysis to provide information on the likelihoods and potential ranges of various impacts.

Viney et al. (2021) proposed a modelling structure that considers three futures or scenarios. The first of these is an undeveloped scenario, which simulates the hydrological regime that would prevail in the absence of any coal resource development. This scenario is employed largely for model calibration purposes. A second scenario, the baseline future, encompasses the impact (which may be ongoing) of any existing developments. The third scenario – the expansion future – accounts for both pre-existing and proposed developments. The impacts of the future expansion of resource development are given by the differences in modelled streamflow characteristics between the expansion and baseline futures.

This paper explores, by way of example, several aspects of the implementation of the modelling framework described by Viney et al. (2021) for assessing the cumulative impacts of multiple future coal resource developments on surface water hydrology.

2. Study background

2.1 Modelling time periods

In all the study locations outlined below, each scenario is run for the period 1983 to 2102 with common local climate inputs. For the period 1983 to 2012, the baseline and expansion

scenarios share a common pattern of resource development and include all operations that were in commercial production prior to December 2012. After 2012, both the baseline and expansion scenarios include all ongoing operations and explicitly take account of the effects of expansions that were planned and approved prior to 2013. However, in addition, the expansion scenario includes any new developments coming into operation after December 2012 and any expansions to existing operations that had not been approved before 2013. Only those proposals that were subjectively judged (as at December 2012) to be likely to proceed are included in the modelling, however, it should be noted that since the modelling commenced, some of these proposals have now been abandoned while others have since been proposed.

2.2 Study locations

The examples reported in this manuscript come from five surface water modelling domains covering various coal-bearing geological basins in eastern Australia (Figure 17). The modelling domains range in size from 2,417 km^2 to 71,532 km^2. Some host existing open-cut or underground coal mines. At the time of modelling, all modelled catchments were subject to proposals for new coal resource developments or for expansions to existing operations that required regulatory approval.

[***INSERT FIGURE Figure 17 NEAR HERE***]

The five modelling domains and their assumed development scenarios are described below.

2.1.1 Suttor River

The Galilee geological basin spans an area of about 248,000 km^2 in Queensland. The basin currently has no significant coal resource extraction. Our modelling assumes the future

development of seven coal mines, two of which will be open cut only and five of which will include both open cut and underground operations. All proposed mines are in the catchment of the Belyando River, a tributary of the Suttor River. The hydrological modelling domain thus comprises an area of 71,532 km^2 in the Suttor River catchment above Burdekin Falls Dam.

2.1.2 Gloucester and Karuah rivers

The Gloucester geological basin occupies 348 km^2 in New South Wales. It contains two existing open cut coal mines. Our modelling assumes the future expansion of both these mines, along with development of a third open cut mine and a coal seam gas field. The hydrological modelling domain encompasses the catchments of the Gloucester and Karuah rivers and covers a total area of 3,100 km^2.

2.1.3 Hunter River

The modelled catchment of the Hunter River occupies 17,787 km^2 in New South Wales. Together with many abandoned mines, it contains more than 20 existing open cut coal mines and eight underground longwall mines. Our modelling assumes the future expansion of 11 of these existing mines and the development of three new open cut mines and two new underground mines. There are no existing or proposed coal seam gas operations. A feature of the Hunter river basin is an auction and trade system that allows mining companies to acquire or transfer permits to discharge saline water to the stream network at times of peak natural streamflow.

2.1.4 Namoi River

The Namoi river basin (38,501 km^2) in New South Wales overlies the Gunnedah coal basin and is home to five existing open cut coal mines and one underground mine. We model the

impacts of expansion proposals for two of these, together with development proposals for three new open cut coal mines, two underground coal mines and a new coal seam gas field. The Namoi catchment contains three major reservoirs that supply water for agricultural, domestic and municipal users. Supply from these reservoirs is controlled by a licencing system and water sharing plan administered by the New South Wales government.

2.1.5 Richmond River

The Richmond river basin in New South Wales currently has no coal resource extraction, while our modelling assumes the future development of a coal seam gas extraction field. The hydrological modelling domain is limited to an area of 2,417 km^2 of the catchment of the Richmond River above the tidal zone and includes the tributaries Eden Creek and Shannon Brook.

2.3 Hydrological models

Surface water changes associated with coal resource development are assessed using the AWRA-L model (Vaze et al., 2019). AWRA-L is the landscape component of the AWRA modelling system (Viney et al., 2014) and is a biophysical model of the water balance between the atmosphere, the soil, groundwater and surface water stores.

AWRA-L has a flexible spatial resolution that is usually dictated by the resolution of the meteorological input data. For use in this study, AWRA-L is forced by gridded meteorological data (precipitation, solar radiation, air temperature, etc.) with a spatial resolution of 0.5 degrees (about 5 km). It operates at a daily time step.

Each spatial unit (grid cell) in AWRA-L is divided into a number of hydrological response units (HRUs) representing different landscape components. Hydrological processes are modelled separately for each HRU before the resulting fluxes are combined to give cell outputs. The

version of AWRA-L used here includes two HRUs which notionally represent (i) tall, deep-rooted vegetation (i.e., forest), and (ii) short, shallow-rooted vegetation (i.e., non-forest). Hydrologically, these two HRUs differ in their aerodynamic control of evaporation, in their interception capacities and in their degree of access to different soil layers.

In application catchments in which river regulation is prominent (Namoi and Hunter), a dedicated routing model, AWRA-R (Dutta et al., 2017), is used. AWRA-R, the river system component of the AWRA system, is a conceptual hydrological model designed for both regulated river systems. The model includes six components: a rainfall-runoff response, a routing scheme, an irrigation model, a river-groundwater interaction component, storages and a floodplain model.

Although in this paper, the AWRA-L model is used to estimate landscape fluxes of water and AWRA-R is used to route this water downstream (where needed), the methodology presented here is agnostic of the specific models used. As a result, it can be applied anywhere in the world, using whichever model(s) are required to adequately capture the hydrologic impacts of coal resource development. It could potentially even be used to assess other types of development, particularly those associated with extractive industries, although this has not been tested thus far.

As described by Viney et al. (2021) the application of AWRA-L and AWRA-R to predict the impact of coal resource development requires estimates of groundwater-mediated fluxes from a suitable groundwater model. In this study, different groundwater models are used in different modelling domains, depending on the specific requirements of each domain. They are:

- Gloucester and Karuah Rivers: regional analytic element groundwater model and MODFLOW alluvial groundwater model (Harbaugh et al., 2000).

Hunter River: a finite element mesh model.

Namoi River: regional groundwater model built using MODFLOW code.

Richmond River: regional groundwater model built using MODFLOW code.

Suttor River: regional analytic element groundwater model.

These groundwater models will not be described or discussed further here. We simply assume that they provide estimates of groundwater–surface water fluxes that are likely to be more accurate than those that can be obtained from a typical surface water model. In practice however, we acknowledge that the coarse spatial resolution of groundwater models, combined with their typically long timesteps make them unsuitable for assessing some of the impacts of coal resource development on groundwater, and particularly on surface water-groundwater interactions. These limitations were minimised by using a variable grid approach in the Hunter and Namoi regions, as well as an alluvial groundwater model in the Gloucester/Karuah region. While this allowed us to assess the impacts of coal resource development on drawdown, reductions in baseflow due to subsidence-induced cracking were unable to be represented in these models. For a discussion of a suitable groundwater modelling framework for coal resources, the reader is referred to Crosbie et al. (2016).

The points of linkage between the surface water and groundwater models and the sequencing of their application are described by Viney et al. (2021).

3. Data preparation

3.1 Disaggregation

In order to facilitate the prediction of streamflow changes at multiple locations within a modelling domain, the catchment must be discretised into smaller sub-units or sub-catchments. A model node represents the outlet point of an associated sub-catchment. Ideally, model nodes should be located at streamflow gauges, above major confluences, immediately below proposed coal mine and coal seam gas developments, and at any other locations that are required for impact modelling (e.g., at ecologically sensitive or culturally significant locations).

[***INSERT FIGURE Figure 18 NEAR HERE***]

[***INSERT FIGURE Figure 19 NEAR HERE***]

Figure 18 shows the 63 node locations chosen for modelling in the Hunter River basin. A schematic representation of the modelling network is shown in Figure 19. In this example, more than half of the 63 model nodes are co-located with streamflow gauges. A further one-third are located just above major confluences and three are immediately downstream of major impoundments. Many of the nodes (e.g., nodes 9, 21, 43) are located immediately downstream of coal mines included in the expansion scenario, while several are at locations that are important for river regulation. Similar model node locations were chosen in the other four study areas.

3.2 Climate trends

The main aim of this study is to predict the impacts of potential future coal resource developments on streamflow. As described by Viney et al. (2021), this can be achieved by comparison of two futures – one without the potential proposals and one with. As such, both scenarios should use common climate inputs. Whilst this can mostly be achieved without incorporating notional temporal changes in climate during the simulation period, it is conceivable that such changes can have an impact on streamflow changes (Chiew et al., 2018).

In this study, future climate inputs are simulated using repetitions of historical, observed climate. In particular, observed climate inputs from 1983 to 2012 are replicated for a further three thirty-year periods to 2102. In each of the replication periods, the climate inputs are scaled using local seasonal scaling factors derived from global climate models (GCMs). In each modelling domain, we use seasonal scaling factors to calculate the change in mean annual precipitation associated with a 1-degree global warming for each of 15 available GCMs. We then select the GCM with the median impact on mean annual precipitation and apply its seasonal scaling factors to the observed record. Assumed global warming in each of the three thirty-year replicates (2013–2042, 2043–2072 and 2073–2102) are 1.0 degrees, 1.5 degrees and 2.0 degrees, respectively.

[***INSERT FIGURE Figure 20 NEAR HERE***]

Figure 20 shows an example for the Richmond River domain. The median GCM has a reduction in mean annual precipitation of 1.8% per degree of global warming. The respective seasonal scaling factors are +4.3%, –5.7%, – 2.5% and –7.5% for summer, autumn,

winter and spring. In other words, projected increases in precipitation in the wettest season, summer, are more than offset by projected decreases in the other three seasons.

Figure 20 depicts the resulting time series of basin-averaged annual precipitation with increasingly trended climate change scalars. It can be seen that the decrease in precipitation from 2013 to 2102 is less than the typical interannual variability.

We acknowledge that this approach does not account for likely climate change induced increases in daily rainfall intensities or changes in wet day frequencies and associated temporal patterns. These changes may be of importance, particularly for high flow streamflow metrics. However, as the same climate sequence is used for both the baseline and development futures, the impacts of these changes would be unlikely to be observed in the model outputs in any case.

3.3 Mine footprints

Each coal mine is associated with an area of disturbance (or footprint) in which potential changes in surface hydrology can be expected. For a surface mine, the footprint area includes not just mine pits, but also roads, spoil dumps, water storages and other infrastructure. It may also include otherwise undisturbed parts of the landscape from which natural runoff is retained in storages. The footprint does not include rehabilitated areas from which runoff can enter the stream network or catchment areas upstream of drainage channels that divert water around a mine site but do not retain it.

In Australia, regulations governing surface coal mines usually prohibit the discharge of any surface water or pumped groundwater from the footprint area. The footprint fraction in a subcatchment thus provides a surrogate for the proportion of natural surface water flows that are no longer discharged to the stream network.

The footprint of an underground mine is the area above the mine workings that is potentially susceptible to subsidence. This area may experience changes in surface topography that give rise to increased ponding or increased recharge.

Mine footprint areas change over the lifetime of a mine's operations. As new parts of the lease are opened up for active use, the footprint increases. As mined parts of the lease are rehabilitated, and their runoff returned to natural drainage, the footprint decreases although not necessarily to pre-mining conditions. As well as the area of any final voids, the final mine footprint may also include the area covered by any infrastructure (e.g. dams, levee banks, roads) that are intended to remain on the site after final rehabilitation.

Time series of mine footprints for baseline and expansion mines were compiled from spatial data supplied by mining companies and government regulators, or extracted from environmental impact statements and related documents, and remote sensing imagery.

[***INSERT FIGURE Figure 21 NEAR HERE***]

Examples of the temporal evolution of mine footprint for three coal mines are shown in Figure 21. Figure 21a represents an open cut mine with baseline and expansion scenarios. Under baseline development, the mine's footprint reaches a maximum of 5.5 km^2 by 2012 before site rehabilitation commencing in 2023 gradually reduces the footprint to a permanent residual of less than 1.0 km^2 by 2032. Under the expansion scenario, the footprint increases to a peak of 19 km^2, before reducing under rehabilitation to a residual of less than 4 km^2.

Figure 21b shows the evolution of expansion footprint areas for a greenfields mine (i.e., no

baseline component) with both surface and underground workings. The mine is projected to commence surface operations in 2019, reaching a peak footprint of 143 km^2 by 2034 before rehabilitation reduces the permanently disturbed area to less than 7 km^2 by 2080. Underground operations are projected to commence in 2024 and reach a peak footprint of 86 km^2 after 40 years.

Figure 21c shows baseline and expansion footprints for a mine with both surface and underground operations. Under pre-existing plans (as at 2013) the surface footprint of the baseline scenario is assumed to continue to increase until 2019 to a maximum of 16 km^2 before rehabilitation reduces this to a permanent residual of less than 3 km^2 by 2039. Under the expansion scenario the open cut footprint increases to a peak of 24 km^2 by 2029. A relatively brief period of underground mining commencing in 2011 under the baseline scenario is supplemented by an expansion to an area of 18 km^2 by 2028.

4. Calibration

4.1 Streamflow model

Streamflow models should ideally be calibrated against observed flow records from gauging stations whose catchments are not affected by coal resource development. For a distributed model like AWRA-L the most robust calibration is achieved through using an objective function that combines goodness of fit measures from several gauging stations. It is not imperative that these stations should be located within the modelling domain; nearby stations are acceptable provided they share similar climatic, topographic and soil characteristics with the modelling domain.

In each modelling domain, separate calibrations are performed for high flow conditions and for low flow conditions. These are achieved by varying the Box-Cox transformation

parameter (lambda value). A transformation parameter value of 1.0 emphasises high flows, while a value of 0.1 places greater emphasis on low flows.

For each calibration catchment a local goodness of fit function, F, is defined by (Viney et al., 2009)

$F_H = (E_d(\lambda) + E_m)/2 - 5|\ln(1+B)|^{2.5}$ for the high flow calibration ($\lambda = 1$); and

$F_L = E_d(\lambda) - 5|\ln(1+B)|^{2.5}$ for the low flow calibration ($\lambda = 0.1$).

where $E_d(\lambda)$ is the Nash-Sutcliffe efficiency of daily flows (Nash and Sutcliffe, 1970) calculated with a Box-Cox transformation parameter of λ, E_m is the Nash-Sutcliffe efficiency of monthly flows and B is the overall prediction bias (total prediction error divided by total observed streamflow).

For each separate calibration, the objective function for the entire domain is then given by the mean of the 25th, 50th, 75th and 100th percentiles of all the F values in the domain.

As an example, for modelling in the Gloucester-Karuah basin, observed streamflow data from 16 gauging stations was used in calibration. Four of these stations are in the catchments of the Gloucester and Karuah rivers, while the remainder are in adjacent catchments. The resulting median F values are 0.62 for the high flow calibration and 0.58 for the low flow calibration.

These two calibrated parameter sets were then assessed in the same 16 catchments for their ability to predict a range of hydrological response variables that reflect different components of the flow regime. The response variables, which are all accumulated to annual time series, are

1. Annual streamflow

2. Daily streamflow at the 99th percentile
3. Number of days with streamflow above the long-term 90th percentile
4. Interquartile range of daily streamflow
5. Daily streamflow at the 1st percentile
6. Number of days with streamflow below the long-term 10th percentile
7. Number of continuous spells with streamflow below the long-term 10th percentile
8. Longest low streamflow spell
9. Number of zero-flow days (defined for practical reasons as flow below 1 ML/d)

The first four of these represent high flow characteristics of the hydrological regime and the last five represent low flow characteristics. In general terms, response variable 9 is only relevant for intermittent streams, or streams which become intermittent under the expansion scenario. For such streams, response variable 5 is likely to be zero. In general, for a particular location, only one or the other of response variables 5 and 9 is likely to produce useful information, but not both.

[***INSERT FIGURE Figure 22 NEAR HERE***]

For the Gloucester-Karuah modelling domain, the range in bias for each parameter set and for each response variable is shown in Figure 22. While both parameter sets predict the high flow variables better than the low flow variables, it is clear that the high flow parameter set provides better predictions of the high flow variables and the low flow parameters provide better predictions of the low flow variables.

The two resulting deterministic model predictions are not used directly in reporting streamflow changes. Instead they are used to (i) inform prior parameter distributions for

the uncertainty analysis (Subsection 5.1); (ii) provide recharge estimates for surface water – groundwater modelling; and (iii) provide system inflows for calibration of the river system model.

4.2 River system model

Unlike the streamflow model, a river system model can only be calibrated using observed streamflow records from within the modelling domain itself. Also, unlike the streamflow model which is calibrated regionally, calibration of the river model is done on a reach by reach basis. One consequence of this is that river system calibration can partly compensate for errors in inflows from the streamflow model. This means that prediction bias in river model output is typically quite small. It also means that even where two quite distinct parameter sets (e.g., a high flow parameter set and a low flow parameter set) for the streamflow model are used to generate inflows into the river model, the resulting predictions from the two realisations of the river model tend to converge.

[***INSERT FIGURE Figure 23 NEAR HERE***]

Examples of this are shown in Figure 23 for the Namoi modelling domain. Here the calibrated biases of predictions of the hydrological response variables across the 23 gauging stations is much smaller, particularly for annual flow (response variable 1 in Figure 23). Secondly, there is much less divergence between the high flow and low flow simulations than is the case with the streamflow model. Note however that the biases in Figures 6 and 7 are not directly comparable as Figure 6 shows results for the Gloucester-Karuah region and Figure 7 shows results for the Namoi region. Despite this, these reductions in bias between AWRA-L and AWRA-R are consistent across all five regions.

5. Uncertainty

5.1 Quantitative uncertainty

The aim of the quantitative uncertainty analysis is to provide probabilistic estimates of the changes in the hydrological response variables due to coal resource development. A large number of parameter combinations are evaluated and, in line with the Approximate Bayesian Computation outlined by Peeters et al. (2016) for propagating uncertainty through models, only those parameter combinations that result in acceptable model behaviour are included in the parameter ensemble used to make predictions.

Model parameters are sampled from a prior distribution that takes account of the values and spread of the optimised parameters from the two calibrated parameter sets, as outlined by Viney et al. (2021).

Acceptable model behaviour is defined for each hydrological response variable based on the capability of the model to reproduce historical, observed time series of the hydrological response variable. For each hydrological response variable, a goodness of fit between model simulated and observed annual hydrological response variable, as well as an acceptance threshold, are defined.

In each modelling domain, 3000 parameter combinations are generated from the AWRA-L and AWRA-R model parameters, together with the parameter combinations for the groundwater model. The acceptance threshold for each hydrological response variable is set to the 90th percentile of the average goodness of fit between observed and simulated hydrological response variable values obtained from model nodes at available streamflow gauging sites. This means that out of the 3000 model replicates, the 300 best (10%) are selected for each hydrological response variable.

The selection of the 10% threshold is based on two considerations: (i) guaranteeing enough prediction samples to ensure numerical robustness, and (ii) the sample's prediction performance is close to that obtained from the high- and low-streamflow model calibrations. Furthermore, it is expected that the full 3000 replicates contain many with infeasible parameter combinations (caused, for example, by parameter correlations that are not considered in the independent random sampling) and that these are likely to be filtered out by sampling only the best 10% of replicates.

The ensemble of predictions are the changes in hydrological response variables simulated with the parameter combinations for which the goodness of fit exceeds the acceptance threshold. The resulting ensembles are presented in Subsection 6.1.

5.2 Qualitative uncertainty

However comprehensive the uncertainty quantification, there will always be aspects of the chain of models that cannot be accounted for, due to data availability, constraints on time and budget or technical limitations. The uncertainty quantification can therefore be complemented by a qualitative uncertainty analysis (Kloprogge et al., 2011; Peeters, 2017). Such a qualitative uncertainty analysis systematically discusses the assumptions and model choices made, scores the extent to which the assumptions were affected by data availability, budget and time constraints or technical limitations, and most importantly assesses the extent to which the assumption may affect the predictions.

[***INSERT TABLE Table 2 NEAR HERE***]

The major assumptions and model choices underpinning the surface water modelling

454 described here are listed in Table 2. The goal of this qualitative uncertainty analysis is to
455 provide a non-technical overview of the model assumptions, their justification and effect on
456 predictions, as judged by the modelling team. This also facilitates an open and transparent
457 review of the modelling.

458 Each assumption in Table 2 is rated against three attributes (data, resources and technical)
459 and their effect on predictions.

460 1. The data rating is the degree to which the question 'If more or different data were
461 available, would this assumption or choice still have been made?' would be
462 answered positively. A low rating means that the assumption is not influenced by
463 data availability, while a high rating indicates that this choice would be revisited if
464 more data were available.

465 2. The resources rating reflects the extent to which resources available for the
466 modelling, such as computing resources, personnel and time, influenced this
467 assumption or model choice. Again, a low rating indicates the same assumption
468 would have been made with unlimited resources, while a high rating indicates the
469 assumption is driven by resource constraints.

470 3. The technical rating reflects the extent to which the assumption is influenced by
471 technical and computational issues. A high rating is assigned to assumptions and
472 model choices that are predominantly driven by computational or technical
473 limitations of the model code. These include issues related to spatial and temporal
474 resolution of the models.

475 The most important rating relates to the effect of the assumption or model choice on the
476 predictions. This is a qualitative assessment by the modelling team of the extent to which a
477 model choice will affect the model predictions, with low indicating a minimal effect and high

a large effect.

A detailed discussion of each of the assumptions, including the rationale for the scoring, follows.

Selection of calibration catchments

The parameters that control the transformation of rainfall into streamflow are adjusted based on a comparison of observed and simulated historical streamflow. Only a limited number of the model nodes have historical streamflow. To calibrate the surface water model, it may be necessary to use data from a number of catchments outside the modelling domain with the parameter combinations that achieve an acceptable agreement with observed flows being deemed suitable for all catchments in the subregion.

The selection of calibration catchments is therefore almost solely based on data availability, which results in a medium rating for this criterion. As it is technically trivial to include more calibration catchments in the calibration procedure and as it would not appreciably change the computing time required, both the resources and technical columns have a low rating.

The regionalisation methodology is valid as long as the selected catchments for calibration are not substantially incompatible with those in the prediction domain in terms of size, climate, land use, topography, geology and geomorphology. The majority of these assumptions can be considered valid and the overall effect on the predictions is therefore deemed to be low.

High-flow and low-flow objective function

AWRA-L simulates daily streamflow. High-streamflow and low-streamflow conditions are governed by different aspects of the hydrological system and it is difficult for any streamflow model to find parameter sets that are able to adequately simulate both extremes of the

hydrograph. In recognition of this issue, two objective functions are chosen: one tailored to medium and high flows and another one tailored to low flows.

Even with more calibration catchments and more time available for calibration, a high-flow and low-flow objective function would still be necessary to find parameter sets suited to simulate different aspects of the hydrograph. Data and resources are therefore scored low, while the technical criterion is scored high.

The high-streamflow objective function is a weighted sum of the Nash–Sutcliffe efficiency and the bias. The former is most sensitive to differences in simulated and observed daily and monthly streamflow, whereas the latter is most affected by the discrepancy between long-term observed and simulated streamflow. The weighting of both components represents the trade-off between simulating short-term and long-term streamflow behaviour. It also reflects the fact that some parameters are more sensitive to daily behaviour and some are more sensitive to long-term hydrology.

The low-streamflow objective function is achieved by transforming the observed and simulated streamflow through a Box-Cox transformation (Box and Cox, 1964) which ensures that a small number of large discrepancies in high streamflow will have less prominence in the objective function than a large number of small discrepancies in low streamflow. Like the high-streamflow objective function, the low-streamflow objective function consists of two components, the efficiency transformed by a Box-Cox power of 0.1 and bias, which again represent the trade-off between short-term and long-term accuracy.

The choice of the weights between both terms in both objective functions is based on the experience of the modelling team (Viney et al., 2009). The choice is not constrained by data, technical issues or available resources. Although different choices of the weights will result in a different set of optimised parameter values, previous with the calibration of AWRA-L on

a continental scale, has shown the calibration to be fairly robust against the weights in the objective function (Vaze et al., 2013).

Although the selection of objective function and its weights is a crucial step in the surface water modelling process, the overall effect on the predictions is marginal through the uncertainty analysis, hence the low rating.

Selection of goodness-of-fit function for each hydrological response variable

The goodness-of-fit function for each hydrological response variable for uncertainty analysis has a very similar role to the objective function in calibration. Where the calibration focuses on identifying a single parameter set that provides an overall good fit between observed and simulated values, the uncertainty analysis aims to select an ensemble of parameter combinations that are best suited to make the chosen prediction.

Within the context of assessing the hydrological impacts of coal resource development, the calibration aims to provide a parameter set that performs well at a daily resolution, whereas the uncertainty analysis focuses on specific aspects of the yearly hydrograph.

The goodness-of-fit function is tailored to each hydrological response variable and averaged over a number of selected catchments that contribute to flow in the modelling domain. This ensures parameter combinations are chosen that are able to simulate the specific part of the hydrograph relevant to the hydrological response variable, at a local scale.

Like the objective function selection, the choice of summary statistic is primarily guided by the predictions and to a much lesser extent by the available data, technical issues or resources. This is the reason for the low rating for these attributes.

The impact on the predictions is deemed minimal (low rating) as it is an unbiased estimate of model mismatch and because it summarises the same aspect of the hydrograph as is

needed for the prediction.

Selection of acceptance threshold for uncertainty analysis

The acceptance threshold ideally is independently defined based on an analysis of the system for propagating uncertainty through models. For the surface water hydrological response variables, such an independent threshold definition can be based on the observation uncertainty, which depends on an analysis of the rating curves for each observation gauging station as well as at the model nodes. There are limited rating curve data available, hence the medium rating. Even if this information were available, the operational constraints might prevent such a detailed analysis – although it is technically feasible. The resources column therefore receives a high rating while the technical column receives a medium rating.

The choice of setting the acceptance threshold equal to the 90th percentile of the summary statistic for a particular hydrological response variable (i.e. selecting the best 10% of replicates) is a subjective decision made by the modelling team. By varying this threshold through a trial-and-error procedure in the testing phase of the uncertainty analysis methodology, the modelling team learned that this threshold is an acceptable trade-off between guaranteeing enough prediction samples and overall good model performance. Although relaxing the threshold may lead to larger uncertainty intervals for the predictions, the median predicted values are considered robust to this change. A formal test of this hypothesis has not yet been carried out. The effect on predictions is therefore scored a medium rating.

Interaction with the groundwater model

The coupling between the results of the groundwater model and the surface water model represents a pragmatic solution to account for surface water – groundwater interactions at a

regional scale. Even if a suitable algorithm for integrated coupling of fluxes between the surface water and groundwater models were available, the differences in spatial and temporal resolution would require non-trivial upscaling and downscaling of spatio-temporal distributions of fluxes. For these reasons and also for practical reasons related to run times and computational storage issues, the modelling methodology for the modelling domains described here involves a one-directional feed of changes in the groundwater flux to streams from the groundwater model, rather than a fully coupled implementation. Thus, the rating for the technical attribute is high.

The data and resources columns are rated medium because even if it were technically feasible to fully integrate the models, the implementation would be constrained by the available data and the operational constraints. In an integrated model, a simulation would likely involve multiple iterations between the groundwater and stream components and increase the computational load significantly.

The overall effect on the predictions is assumed to be small, as the change in baseflow due to coal mining is small compared to the other components of the water balance and the effect of rainfall interception by mine sites.

Implementation of the expansion scenario

The coal resource expansion plan is implemented through the interaction with the groundwater model and by removing the fraction of runoff in the catchment that is intercepted by the mine footprint from the total catchment runoff. The key choices that are made in implementing this scenario are (i) determining which mining developments are included, and (ii) deciding on the spatial and temporal development of their hydrological footprints.

In catchments in which the mine footprint is only a small fraction of the total area of the

catchment, the precise delineation of the spatial extent of the mine footprint is not crucial to the predictions. In catchments in which the footprint is a sizeable fraction, accurate delineation of mine footprint becomes very important.

Similarly, the temporal evolution of mine footprints is crucial as it will determine how long the catchment will be affected. This is especially relevant for the post-mining rehabilitation of mine sites, when it becomes possible again for runoff generated within the mine footprint to reach the streams.

The accuracy with which mine footprints are represented in the model depends largely on the accuracy of the planned mine footprints published or provided by the mine proponents. This therefore is one of the crucial aspects of the surface water model as it potentially has a high impact on predictions and it is driven by data availability rather than availability of resources or technical issues. The data attribute is therefore rated high, while the resources and technical columns are rated low. The effect on predictions is rated high.

6. Examples of prediction outcomes

6.1 Metrics of hydrological change

The impacts of coal resource development on the streamflow regime at each model node are evaluated using the nine hydrological response variables outlined in Section 4.1. These hydrological response variables were chosen to be able to quantify changes across the entire flow regime. For each of these hydrological response variables a time series of annual values for the period 2013 to 2102 is constructed for each model node.

In order to more directly represent physically significant streamflow stages, three further response variables are introduced. Unlike the first nine variables, which are calculated as annual time series, the remaining three are calculated as average occurrence frequencies

over the three 30-year time periods (2013–2042, 2043–2072 and 2073–2102) and are referenced to the modelled average baseline occurrences in the reference period, 1983–2012. They are therefore designed to facilitate direct comparison of changes between the four time periods. These variables are:

10. The average number of events per year over a 30-year period where peak daily flow in flood events exceeds the modelled flow with a return period of 0.3 years in the reference period. This metric is designed to be approximately representative of over-bench flow events.
11. The average number of events per year over a 30-year period where peak daily flow in flood events exceeds the modelled flow with a return period of 3 years in the reference period. This metric is designed to be approximately representative of over-bank flow events.
12. The average number of days per year with streamflow below 10 ML/d during a 30-year period. This threshold is designed to be approximately representative of the flow rate at which all river pools will join up and form a continuously flowing reach.

Each of these three variables has specific ecological significance. Response variable 10 is significant for vegetation seeding and recruitment. Response variable 11 is significant for floodplain inundation. Response variable 12 is significant for the mobility of flow-dependent fauna. These three response variables are used in a method outlined by Hosack et al. (2018) to make inferences about the impact of coal resource development on water-dependent ecological assets.

For each model node, 3000 sets of randomly selected parameter values were used to generate 3000 replicates of development impact. From these, the best 300 replicates for each hydrological response variable – as assessed by their ability to predict that hydrological response variable at a number of observation sites – were chosen for further analysis. The

assessment nodes are chosen for their availability of suitable observational data. Results are presented using a series of boxplots for each hydrological response variable. Each boxplot was generated from the resulting 300 samples and shows differences between the predicted expansion and baseline scenarios. The boxplots (examples of which are shown for five of the response variables selected across the five study regions in Figure 24 to Figure 28) show the distributions over the 300 replicates of the maximum raw change (*amax*) for selected response variables between the baseline and expansion predictions, the corresponding maximum percent change (*pmax*) and the year of maximum change (*tmax*). In general, the most meaningful diagnostic for the flux-based metrics (1, 2, 4 and 5) is *pmax*, while the most meaningful diagnostic for the frequency-based metrics (3, 6, 7, 8 and 9) is *amax*.

It is important to recognise that the *amax* and *pmax* values give the largest annual departure between the expansion and baseline predictions for the respective hydrological response variables. As such, *amax* and *pmax* represent extreme responses. They do not represent the magnitudes of responses that would be expected to occur every year.

[***INSERT FIGURE Figure 24 NEAR HERE***]

Changes in annual flow (hydrological response variable 1) for model nodes in the Suttor River basin are shown in Figure 24. Changes in *amax* accumulate with distance downstream. For the 300 model replicates, the median of the maximum annual reduction in streamflow reaches a peak of 74 GL at the bottom of the basin. However, much of this reduction is sourced from further upstream, in the parts of the basin where maximum development occurs. The biggest changes in local streamflow occur closer to the location of the developments themselves and reach a peak median reduction of 21% at node 3 in Sandy

Creek. In most nodes, the year of maximum impact on annual streamflow is predicted to be between 2038 and 2051. This timing reflects the approximate time of maximum coal mine development footprint for a number of the modelled mines.

[***INSERT FIGURE Figure 25 NEAR HERE***]

Changes in the magnitude of peak flow events (response variable 2) in the Hunter River catchment show a similar pattern (Figure 25), with *amax* accumulating with distance downstream and *pmax* being greatest in those tributaries where maximum development occurs. The greatest median reductions in peak flow occur in the smaller, heavily developed sub-catchments with a maximum of 68% at node 52, which has a 25 km^2 catchment and a maximum open cut mine footprint covering 66% of the sub-catchment area. Once again, the timing of the maximum change in peak flow reflects the approximate time of maximum coal mine development.

[***INSERT FIGURE Figure 26 NEAR HERE***]

Changes in the annual number of days with low flow (response variable 6) in the Gloucester-Karuah modelling domain are shown in Figure 26. In general, there is an increase in the number of low flow days as a result of mining development, but in some sub-catchments there is a modelled reduction at the 5th percentile. The greatest median increase in low flow frequency occurs at node 14 and represents an extra 12 days per year. However, the

timing of the biggest impacts on response variable 6 show considerable spread with medians sometimes occurring very late in the modelling period.

[***INSERT FIGURE Figure 27 NEAR HERE***]

Changes in the annual number of low flow spells (response variable 7) in the Namoi River basin are shown in Figure 27. The number of low flow spells is predicted to increase at most nodes, but at the 5th percentile, some nodes show decreases as shorter spells coalesce. The impact on number of spells is relatively small in the main river channel (median impact is no greater than one spell per year), but the median of the maximum change rises to 23 extra spells per year at the tributary node 25. There is considerable uncertainty in the projections of *tmax* in Figure 27, although most median *tmax* values occur before 2050.

[***INSERT FIGURE Figure 28 NEAR HERE***]

Changes in the length of the longest low flow spell (response variable 8) in the Richmond River basin are shown in Figure 28. The Richmond River modelling domain is the only one of the five discussed here for which changes in streamflow are mediated entirely by coal seam gas developments. The largest change in the median of the maximum spell length is six days at node 6. The timing of maximum change tends to be relatively late in the modelling period, if not beyond 2102, because groundwater changes tend to increase over time and groundwater levels have often not reached equilibrium by the end of the modelling period.

[***INSERT FIGURE Figure 29 NEAR HERE***]

The average increase in stream disconnectedness (as assessed by response variable 12) for three 30-year time periods in the Suttor River basin is shown in Figure 29. Mines are located along Sandy Creek and above node 34 on the Belyando River. The response in the first 30 year period is for substantial increases in the occurrence of stream disconnectedness in reaches of the relatively drier Sandy Creek, coupled with smaller increases further downstream. By the second 30-year time period, substantial increases in stream disconnectedness have propagated further downstream and this pattern persists in the final 30-year time period. However, in Sandy Creek there is a recovery in stream connectedness after 2042 when most of the mining operations in that area have ceased and mine footprints have reduced.

[***INSERT FIGURE Figure 30 NEAR HERE***]

[***INSERT FIGURE Figure 31 NEAR HERE***]

It is not unreasonable to expect that the overall impact on streamflow characteristics of coal seam gas extraction and coal mining should relate to the magnitude and density of the extractive operations. In Figure 30 and Figure 31, we assess the maximum changes in annual streamflow as a function of extraction density. Figure 30 shows the relationship between well density and predicted maximum change in annual streamflow for different model nodes in and near the proposed coal seam gas field in the Richmond River basin.

There is a clear linear relationship with the greatest reductions in streamflow occurring at nodes with the highest density of wells within their catchments. It should be noted, however, that the changes in annual streamflow are very small and represent percentage changes of less than 1%.

Similarly, Figure 31 shows that modelled streamflow in the Hunter, Namoi and Suttor modelling domains, which are dominated by coal mining, is strongly dependent on the proportion of a sub-catchment that comprises mine footprints. The greater the footprint proportion in a sub-catchment, the greater the corresponding reduction in percentage streamflow. This is not entirely unexpected since the model algorithm adopted here for coal mining mandates that reductions in surface runoff generation are equal to the footprint area. Departures from this notional one-to-one relationship in Figure 31 are caused by the added imposition of groundwater-mediated changes and of water extraction for irrigation and industry, and by cases where the modelled maximum change in streamflow does not exactly coincide with the year of maximum mine footprint. The latter anomaly is more likely to occur if the maximum footprint occurs in a year with small baseline streamflow.

The impacts on annual streamflow that are modelled here for coal mining appear to be slightly less than those reported in the modelling study by Guo et al. (2017). This difference might be associated with a greater impact of groundwater on streamflow changes in the latter study.

6.2 Thresholds of acceptable change

In order to rule out water-dependent landscape assets that are very unlikely to be impacted by changes in surface water hydrology, it is necessary to define what a significant change in hydrology is and which reaches of the stream network are and are not showing a significant hydrological change. A significant hydrological change is defined conservatively for each of

the eight hydrological response variables. For:

- the high-flow flux-based hydrological response variables 1, 2 and 4, this is a greater than 5% chance of a 1% or greater change in the variable (i.e. if more than 5% of model replicates show a maximum difference between the expansion and baseline scenarios of at least 1% of the baseline value).

- the low-flow flux-based hydrological response variable 5, this is a greater than 5% chance of a 1% or greater change in the variable and the change in runoff depth is greater than 0.0002 mm. The addition of a runoff depth threshold is designed to exclude reaches where the absolute change in runoff is negligible.

- the frequency-based variables 3, 6, 8 and 9, this is a greater than 5% chance of there being a change in the variable of at least 3 days in any year.

- the frequency-based variable 7, this is a greater than 5% chance of there being a change in the variable of at least two spells in any year.

Significant, as defined here, does not mean that changes that exceed these thresholds are necessarily large or will have a noticeable impact; the thresholds have been defined very conservatively. Rather it is used to delineate areas where streamflow is *very unlikely* to be impacted by the additional coal resource development. If results from the surface water modelling indicate that for all eight variables at a model node there is a less than 5% probability that the hydrological changes will exceed the thresholds, then the landscape classes and assets that depend on streamflow at that location can be considered very unlikely to be impacted. Thus these significance thresholds form the basis for defining the *zone of potential hydrological change* (Subsection 6.3), outside of which the potential for impacts is very unlikely.

Using these definitions, at 18 nodes in the Hunter River catchment, there are no significant hydrological changes due to the additional coal resource development; at 10 nodes, there are significant changes in all eight hydrological response variables; at all other nodes, there are significant changes in some hydrological response variables, but not others. The majority of nodes (44 of 63) experience changes in three of the low-streamflow hydrological response variables (variables 6, 7 and 8) and in the high flow response variable 4 (41 of 63); about half (31 of 63) experience a significant change in annual streamflow (variable 1).

Zone of potential hydrological change

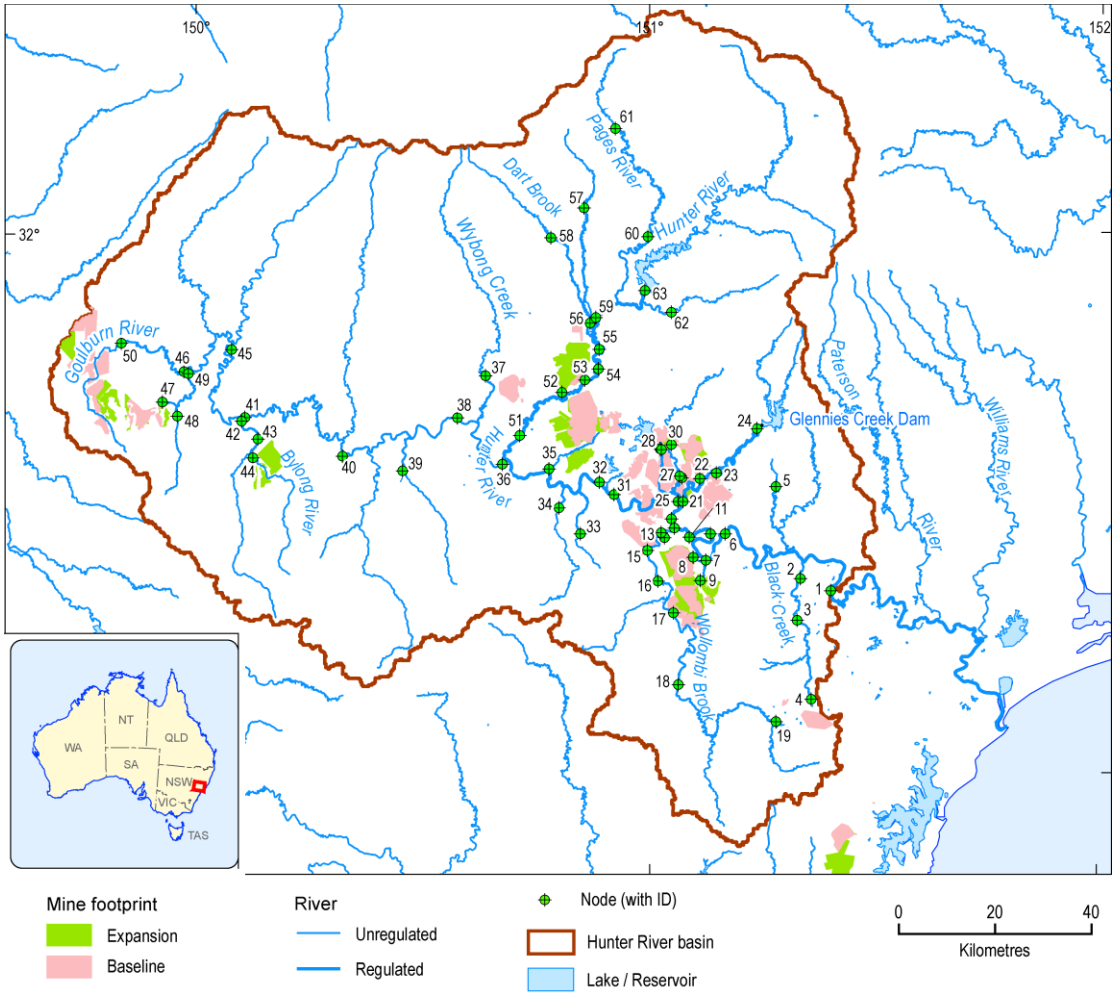


Figure 18 enables results at a model node to be applied to some length of reach upstream

and downstream of the node, as appropriate to do so. The information in can be used to identify the reaches of the Hunter blue line river network which are likely to have a significant hydrological change from additional coal resource development shown.

[***INSERT FIGURE Figure 32 NEAR HERE***]

Figure 32 shows reaches predicted to experience a significant change in at least one hydrological response variable due to additional coal resource development. For some reaches (e.g. node 18 to node 19; node 55 to node 59), the change from a significant hydrological change to a non-significant hydrological change occurs somewhere between the two nodes. These reaches are shown as dashed pink lines and other information is needed to determine where to delineate the change from significant to non-significant hydrological change. Note that these streams can show potential impacts upstream of the mine locations due to groundwater drawdown potentially impacting streamflow. In both cases shown here, it is likely that significant impacts only extend for a few kilometres upstream of the node immediately above the coal mines, but in the absence of further information, this cannot be verified. Similarly, upstream of the pink headwater model nodes in Figure 32 (i.e. those showing a significant change in hydrology), there will be some length of stream that is also potentially affected by coal resource development. The potentially affected reaches comprise more than 600 km of the modelled Hunter River stream network.

Note that Figure 15 shows that the impact of coal mining on streamflow decreases moving further downstream from the mine site, however Figure 16 shows that the Hunter River is potentially impacted all the way to the end of the modelling domain at Node 1. This is because one or more of the hydrological response variables has a greater than 5% chance of

exceeding the specified change all the way to the end of the modelling domain (as can be seen in Figure 9 showing results for the 99th percentile of flow in the Hunter River. As a result, we are unable to rule these parts of the river out in terms of potential impacts from coal mining.

To define the zone of potential hydrological change for any impact and risk analysis – that is, the area outside of which it is very unlikely that the water-dependent landscape classes and assets will be impacted – we need to determine the upstream extents of the stream network likely to experience a significant hydrological change. This final step is addressed by Post et al. (2020) where drawdown results from the groundwater modelling and mine footprint data are used to identify stream reaches that are not explicit in the surface water model node-link network and where hydrological changes from the expansion scenario are potentially significant.

Conclusions

This manuscript presents examples of the modelling of the impacts of coal mining and coal seam gas extraction on streamflow in five study catchments in Australia. The modelling methodology enables the prediction of cumulative impacts from multiple future coal resource developments and distributes these predictions at multiple locations in the landscape. It is framed in terms of a structured uncertainty analysis to provide information on the likelihoods and potential ranges of various impacts.

A qualitative uncertainty analysis, which subjectively assesses the likely impact on model results of various assumptions made during the modelling procedure, indicates that model predictions are most sensitive to uncertainty in the implementation of the expansion scenario for future resource development. It further suggests that this uncertainty can be ameliorated, in part, by the incorporation of better data on resource development. The

qualitative uncertainty analysis also indicates that model results are relatively insensitive to model choices involving model calibration and interactions with the groundwater model.

Predictions of hydrological change associated with coal resource development suggest that, in the Australian context, maximum percentage reductions in annual streamflow are approximately commensurate with the proportion of coal mine coverage. Departures from this notional one-to-one relationship are caused by the added imposition of groundwater-mediated changes and of water extraction for irrigation and industry, and by cases where the modelled maximum change in streamflow does not exactly coincide with the year of maximum mine footprint.

In coal seam gas fields, reductions in modelled annual streamflow are proportional to well density. However, these reductions are substantially smaller than those associated with coal mining.

The predictions of hydrological change associated with future coal resource development in the Hunter River basin are used to identify a zone of potential hydrological change within the catchment. This zone delineates those parts of the landscape where water-dependent landscape classes and assets may be vulnerable to change associated with changes in the streamflow regime. A corollary of this is that any parts of the landscape outside the zone of potential hydrological change are unlikely to be affected by coal resource development. Model results suggest that as much as 600 km of the Hunter River basin's stream network may be vulnerable to significant flow changes. Such changes might include reductions in the magnitude of annual flow and in the frequency of flood events and increases in the prevalence of low flow frequency and persistence.

Acknowledgement

The work described in this manuscript was partly funded by the Australian Government's Department of the Environment as part of the Bioregional Assessments Programme.

References

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. Royal Statistical Soc., 26, 211–243.

Chiew, F.H.S., Fu, G., Post, D.A., Zhang, Y., Wang, B., Viney, N.R., 2018. Impact of coal resource development on streamflow characteristics: influence of climate variability and climate change. Water, 10, w10091161.

Crosbie, R.S., Peeters, L., Carey, H., 2016. Groundwater modelling. Submethodology M07 from the Bioregional Assessment Technical Programme. Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia, 61pp.

Dutta, D., Vaze, J., Kim, S., Hughes, J., A. Yang, A., Teng, J., Lerat, J., 2017. Development and application of a large scale river system model for national water accounting in Australia. J. Hydrol., 547, 124–142.

Guo, Q., Su, N., Yang, Y., Li, J., Wang, X., 2017. Using hydrological simulation to identify contribution of coal mining to runoff change in the Kuye River Basin, China. Water Resour., 44, 586–594.

Harbaugh A.W., Banta, E.R., Hill, M.C. and McDonald, M.G., 2000. MODFLOW-2000, The U.S. Geological Survey modular ground-water model – user guide to modularization concepts and the ground-water flow process. U.S. Geological Survey Open-File Report 00-92. U.S. Department of the Interior, Reston, Virginia.

Hosack, G.R., Ickowicz, A., Hayes, K.R., Dambacher, J.M., Barry, S.A., Henderson, B.L., 2018. Receptor impact modelling. Submethodology M08 from the Bioregional Assessment Technical Programme, Department of the Environment and Energy, Bureau of Meteorology, CSIRO and Geoscience Australia, Australia, 89pp.

Kloprogge, P., van der Sluijs, J.P., Petersen, A.C., 2011. A method for the analysis of assumptions in model-based environmental assessments. Environ. Mod. Softw., 26, 289–301.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. Part I – a discussion of principles. J. Hydrol., 10, 282–290.

Peeters, L. J. M., 2017. Assumption hunting in groundwater modeling: find assumptions before they find you. Groundwater, 55, 665–669.

Ping, J., Yan, S., Gu, P., Wu, Z., Hu, C., 2017. Application of MIKE SHE to study the impact of coal mining on river runoff in Gujiao mining area, Shanxi, China. PLOS ONE, 12, e0188949.

Post, D. A., Crosbie, R. S., Viney, N. R., Peeters, L. J. M., Zhang, Y., Herron, N., Janardhanan, S., Wilkins, A., Karim, F., Aryal, S., Pena-Arancibia, J., Lewis, S., Evans, T., Vaze, J., Chiew, F. H. S., Marvanek, S., Henderson, B., Schmidt, B. and Herr, A. 2020. Impacts of coal resource development in eastern Australia on groundwater and surface water. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2020.125281.

Vaze J., N.R. Viney, M.P. Stenson, L.J. Renzullo, A.I.J.M. van Dijk, D. Dutta, R. Crosbie, J. Lerat, D. Penton, J. Vleeshouwer, L.J.M Peeters, J. Teng, S. Kim, J. Hughes, W. Dawes, Y. Zhang, B. Leighton, J. Perraud, J.Y.A. Joehnk, B. Wang, A. Frost, A. Elmahdi, A. Smith and C.

Daamen, 2013. The Australian Water Resource Assessment modelling system (AWRA). MODSIM2013, 20th International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand, 3015–3021.

Vaze, J., Viney, N.R., Crosbie, R., Dawes, W.R., Wang, B., Frost, A., Donnelly, C., Elmahdi, A., 2019. Australian Water Resources Assessment Landscape model (AWRA-L). Hydrol. Proc. (submitted).

Viney, N.R., Perraud, J., Vaze, J., Chiew, F.H.S., 2009. The usefulness of bias constraints in model calibration for regionalisation to ungauged catchments. Proc. 18th World IMACS/MODSIM Congress, Cairns, Australia, 3421–3427.

Viney, N.R., Post, D. A., Crosbie, R.S., Peeters, L.J.M., 2021. Modelling the impacts of future coal mining and coal seam gas extraction on river flows: a methodological framework. J. Hydrol. (this issue).

Viney, N.R., Vaze, J., Vleeshouwer, J., Yang, A., van Dijk, A., Frost, A., 2014. The AWRA modelling system. Proc. Hydrol. and Water Resour. Symp., Perth, Australia, 1018–1025.

Figure captions

Figure 1. Location of the five study catchments in eastern Australia.

Figure 2. Subcatchment disaggregation and node locations for the Hunter River modelling domain.

Figure 3. Schematic representation of the node-link network for the Hunter River basin. Model nodes number from most downstream node upstream. Blue nodes correspond to stream gauging stations; orange nodes correspond to model-specific nodes. The thicker blue line depicts the regulated part of the river system.

Figure 4. Time series of observed (1900–2012) and projected (2013–2102) annual precipitation averaged over the Richmond River modelling domain (blue line); the red line is a centrally weighted moving average.

Figure 5. Assumed temporal variation of footprint area for three individual coal mines: (a) an open-cut mine in the Namoi River basin with baseline and expansion scenarios; (b) a proposed open-cut and underground mine in the Suttor River basin; and (c) an open-cut and underground mine in the Hunter River basin with both baseline and expansion scenarios.

Figure 6. Biases of eight hydrological response variables in 16 calibration catchments for the streamflow model AWRA-L in the Gloucester-Karuah modelling domain. Boxplots show 10th, 25th, 50th, 75th and 90th percentiles for the high flow calibration (blue) and the low flow calibration (red).

Figure 7. Biases of eight hydrological response variables in 23 calibration catchments for the river model AWRA-R in the Namoi modelling domain. Boxplots show 10th, 25th, 50th, 75th and 90th percentiles for the high flow calibration (blue) and the low flow calibration (red).

Figure 8. Maximum absolute change (amax), maximum percentage change (pmax) and time of maximum change (tmax) in hydrological response variable 1 (annual flow) at selected model nodes within the Suttor River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

Figure 9. Maximum absolute change (amax), maximum percentage change (pmax) and time of maximum change (tmax) in hydrological response variable 2 (99th percentile flow) at selected model nodes within the Hunter River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

Figure 10. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 6 (number of low flow days) at 19 model nodes within the Gloucester-Karuah modelling domain. The streamflow network accumulates towards the left of the figure to nodes 1 and 30. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

Figure 11. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 7 (number of low flow spells) at selected model nodes within the Namoi River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

Figure 12. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 8 (length of longest low flow spells) at 14 model nodes within the Richmond River catchment. The streamflow network accumulates towards the left of

the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

Figure 13. Average increase in hydrological response variable 12 (stream disconnectedness) over three 30-year time periods, expressed in average number of days per year, for a contiguous transect of model nodes in the Suttor River (nodes 27–25), the Belyando River (nodes 55–11) and Sandy Creek (nodes 8–3). The streamflow network accumulates towards the left of the figure.

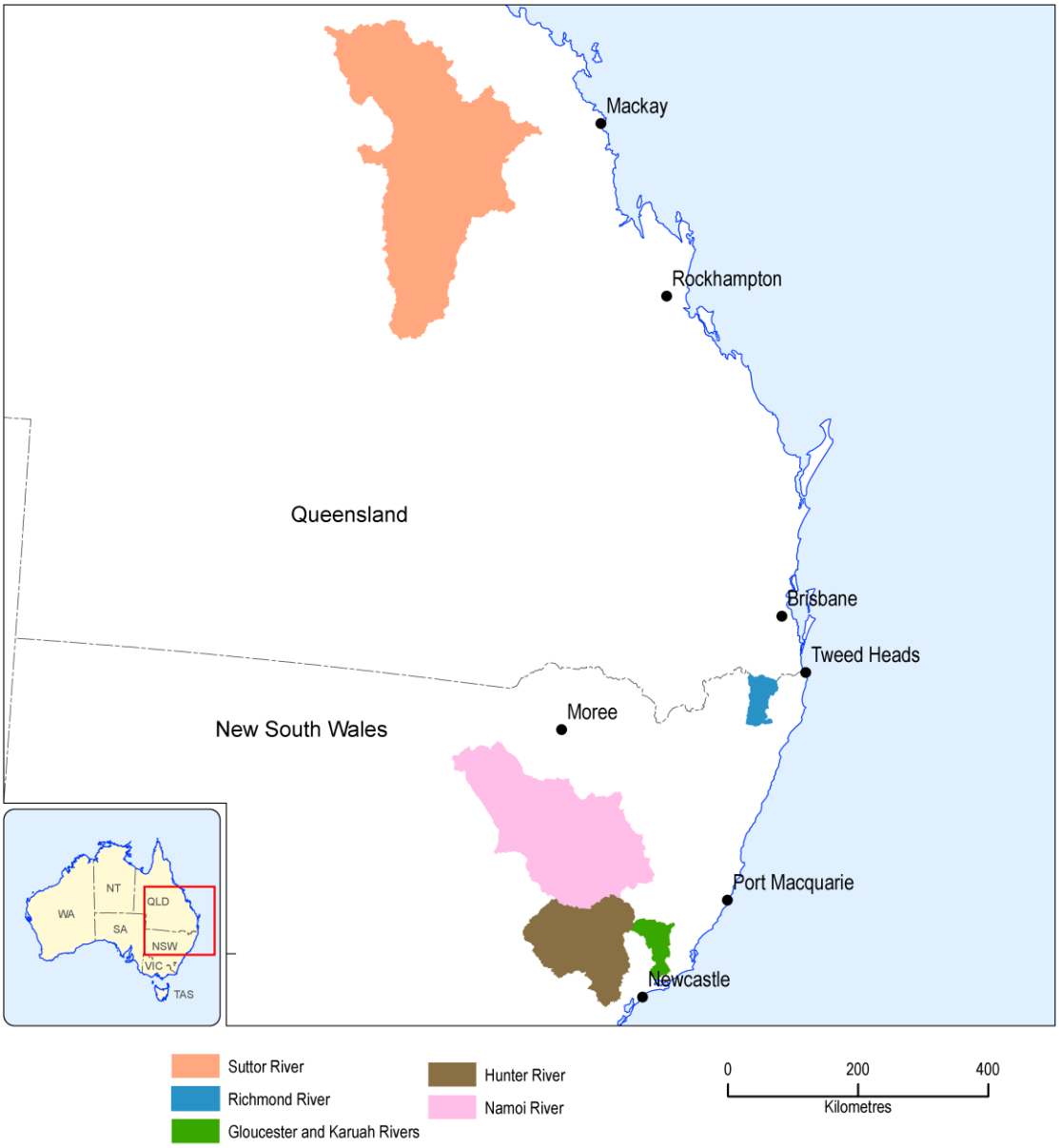
Figure 14. 95th percentile of maximum change in annual flow at model nodes in the Richmond River catchment as a function of the density of coal seam gas wells within 2 km of the node subcatchment.

Figure 15. Predicted maximum change in annual streamflow as a function of the maximum proportion of mine footprint in a catchment for model nodes in the Hunter, Namoi and Suttor river basins. For the Namoi River basin, only model nodes with catchments further than 20 km from a proposed coal seam gas field are included.

Figure 16. Reaches in the Hunter River catchment that are potentially susceptible to changes in streamflow characteristics under the expansion scenario.

Table captions

Table 1. Qualitative uncertainty analysis for the surface water modelling in the Namoi modelling domain.



989

990 **Figure 17. Location of the five study catchments in eastern Australia.**

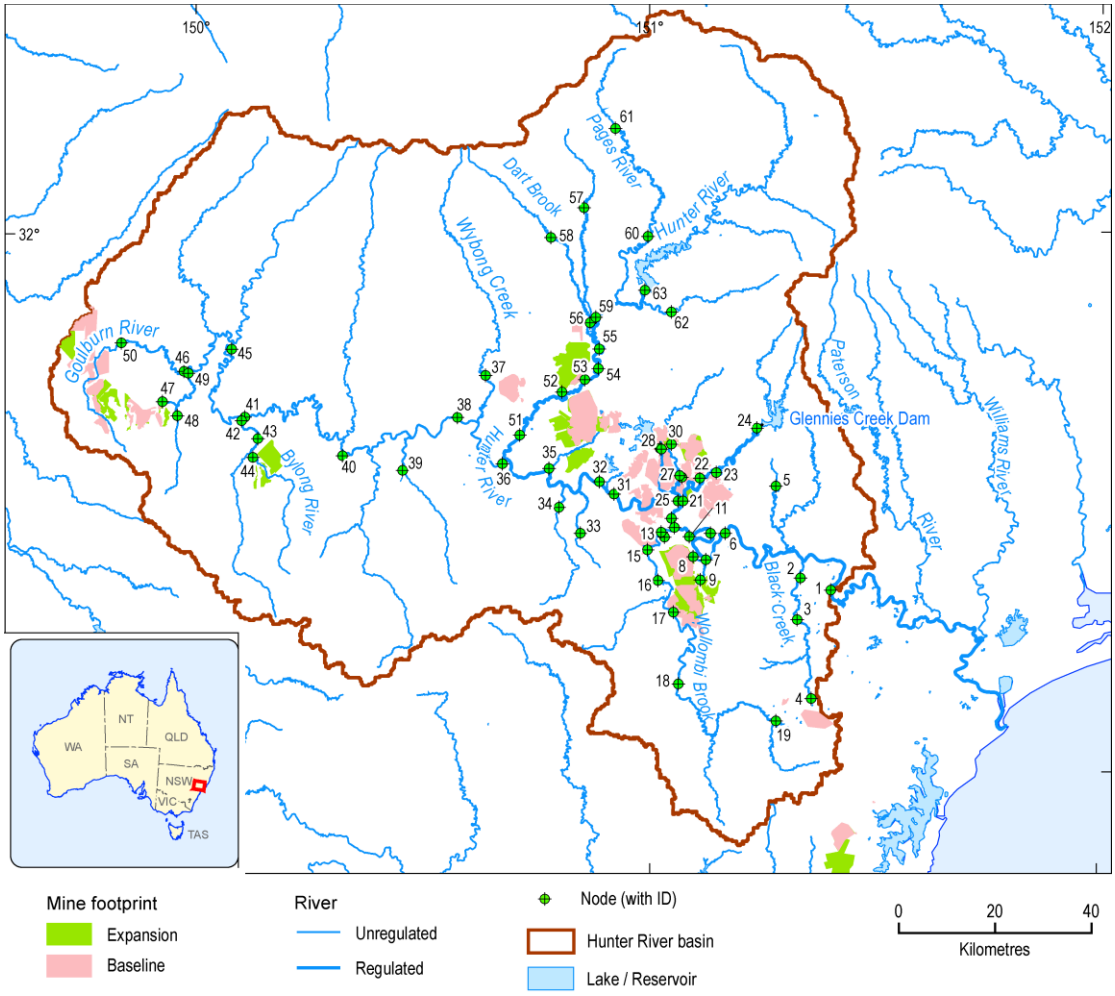
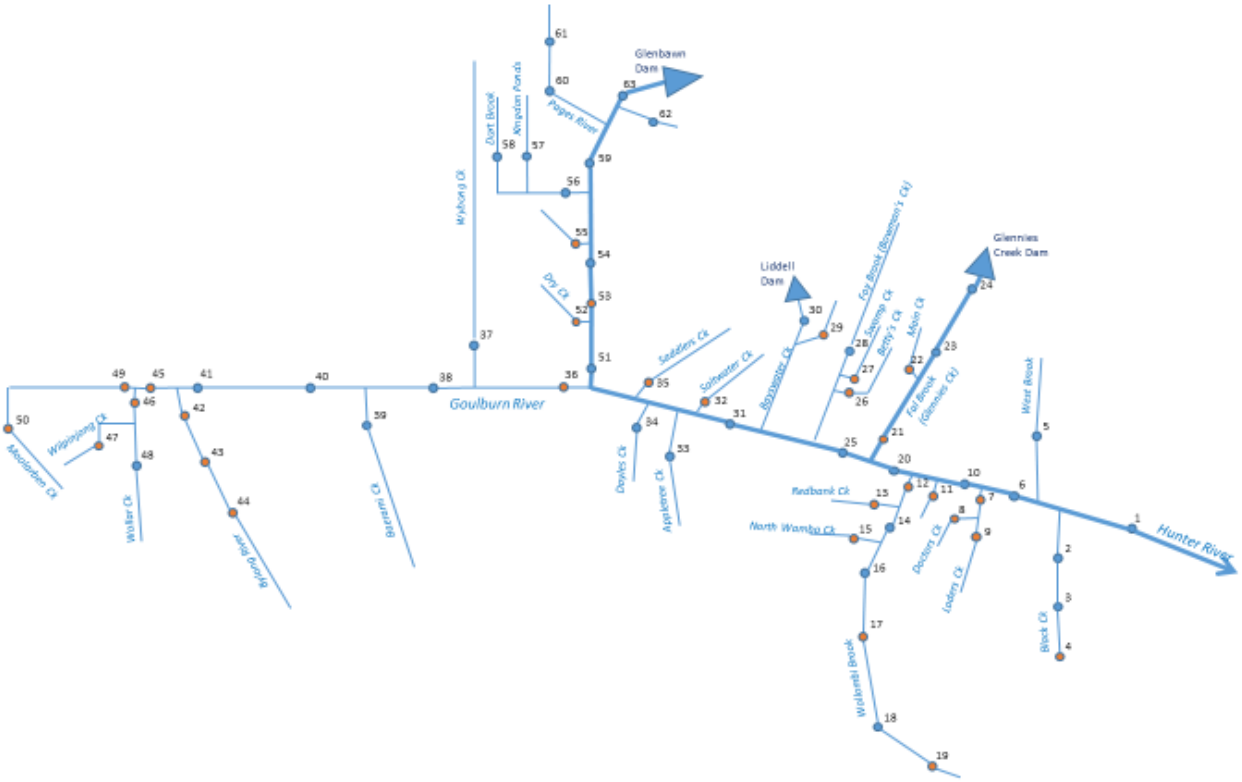


Figure 18. Subcatchment disaggregation and node locations for the Hunter River modelling domain.



994

995 **Figure 19. Schematic representation of the node-link network for the Hunter River**

996 **basin. Model nodes number from most downstream node upstream. Blue nodes**

997 **correspond to stream gauging stations; orange nodes correspond to model-specific**

998 **nodes. The thicker blue line depicts the regulated part of the river system.**

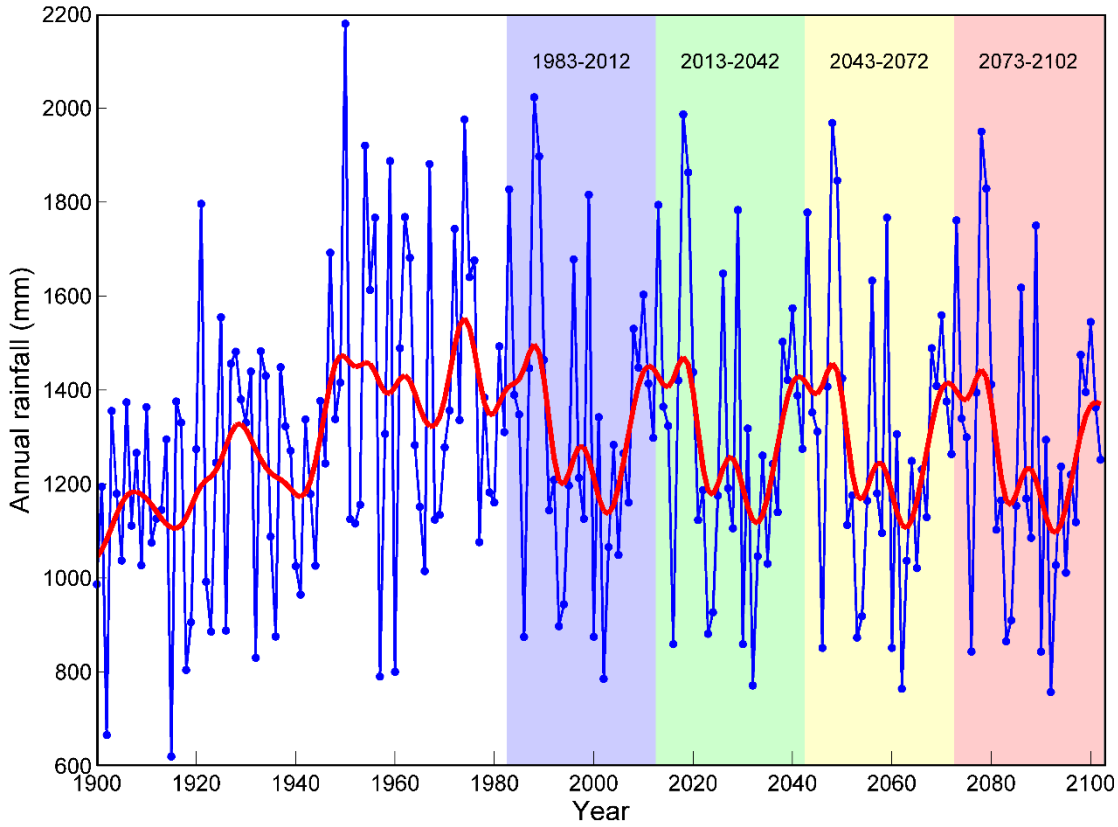
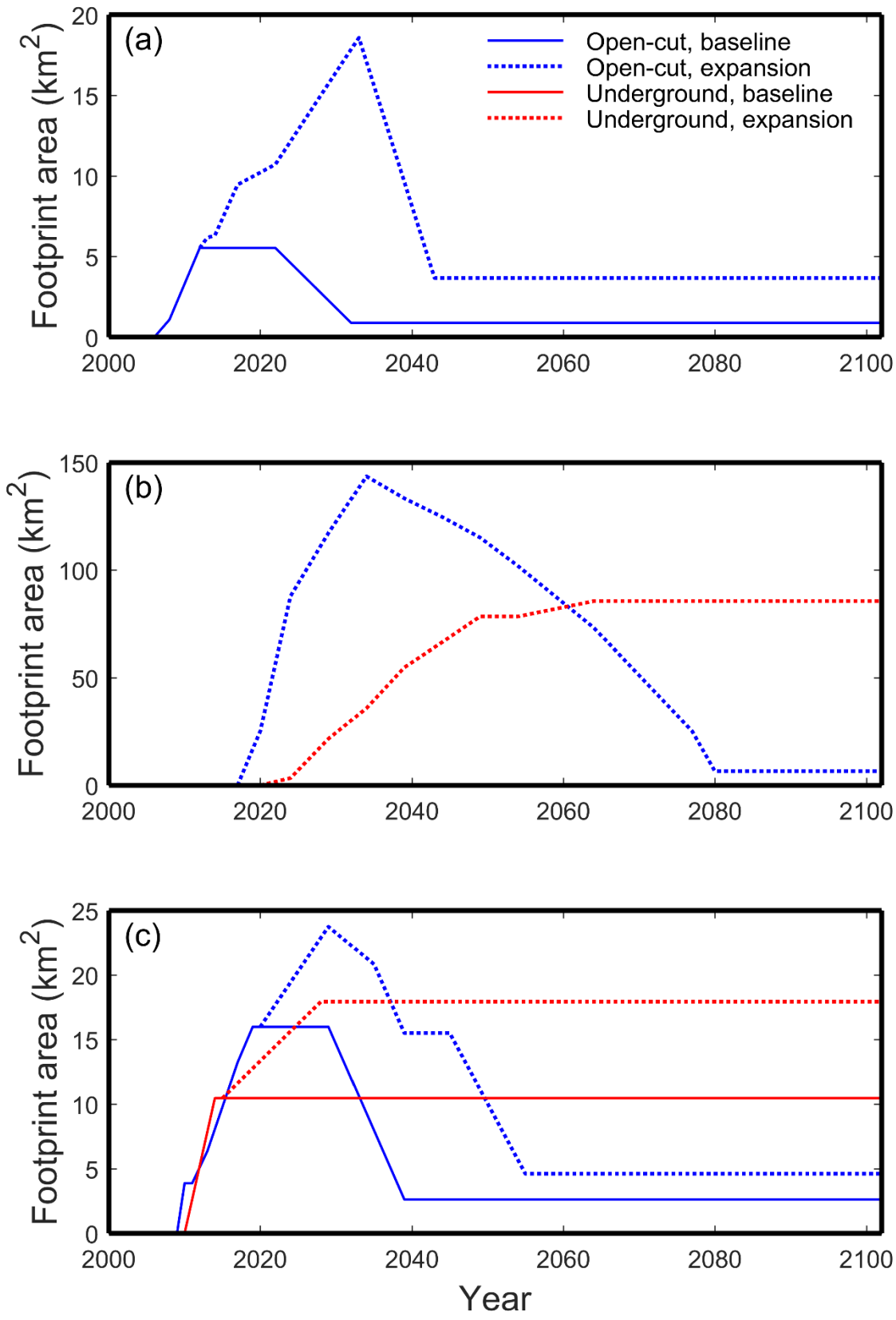


Figure 20. Time series of observed (1900–2012) and projected (2013–2102) annual precipitation averaged over the Richmond River modelling domain (blue line); the red line is a centrally weighted moving average.



1003

1004 **Figure 21. Assumed temporal variation of footprint area for three individual coal**

1005 **mines: (a) an open-cut mine in the Namoi River basin with baseline and expansion**

scenarios; (b) a proposed open-cut and underground mine in the Suttor River basin; and (c) an open-cut and underground mine in the Hunter River basin with both baseline and expansion scenarios.

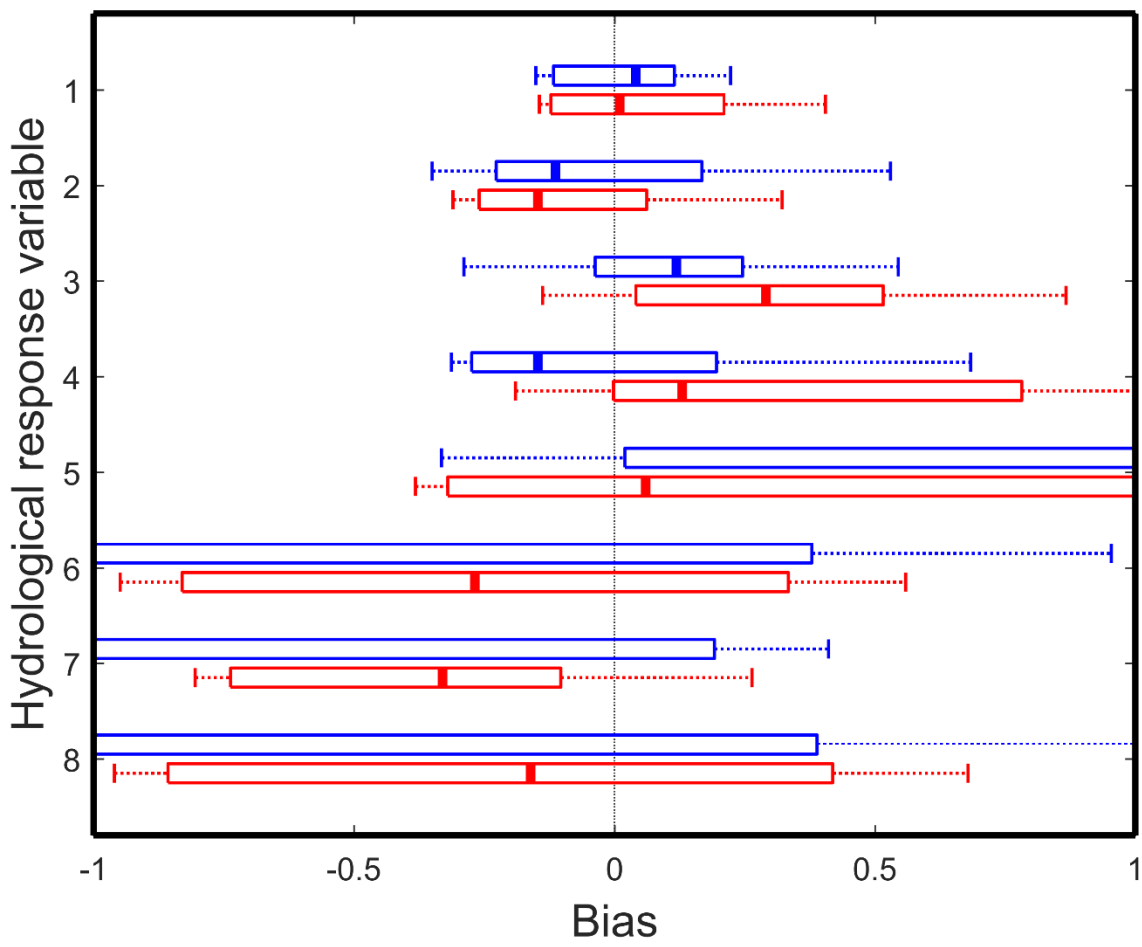


Figure 22. Biases of eight hydrological response variables in 16 calibration catchments for the streamflow model AWRA-L in the Gloucester-Karuah modelling domain. Boxplots show 10th, 25th, 50th, 75th and 90th percentiles for the high flow calibration (blue) and the low flow calibration (red).

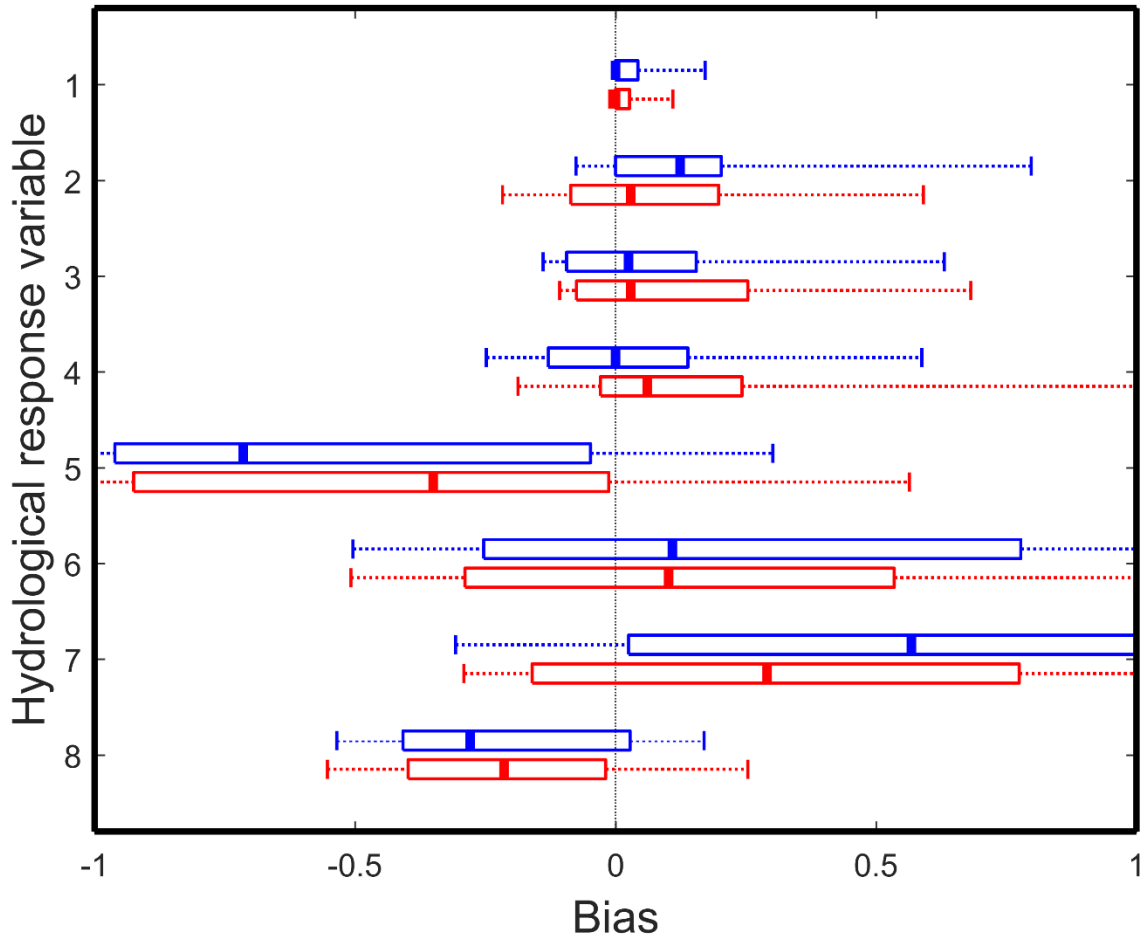


Figure 23. Biases of eight hydrological response variables in 23 calibration catchments for the river model AWRA-R in the Namoi modelling domain. Boxplots show 10th, 25th, 50th, 75th and 90th percentiles for the high flow calibration (blue) and the low flow calibration (red).

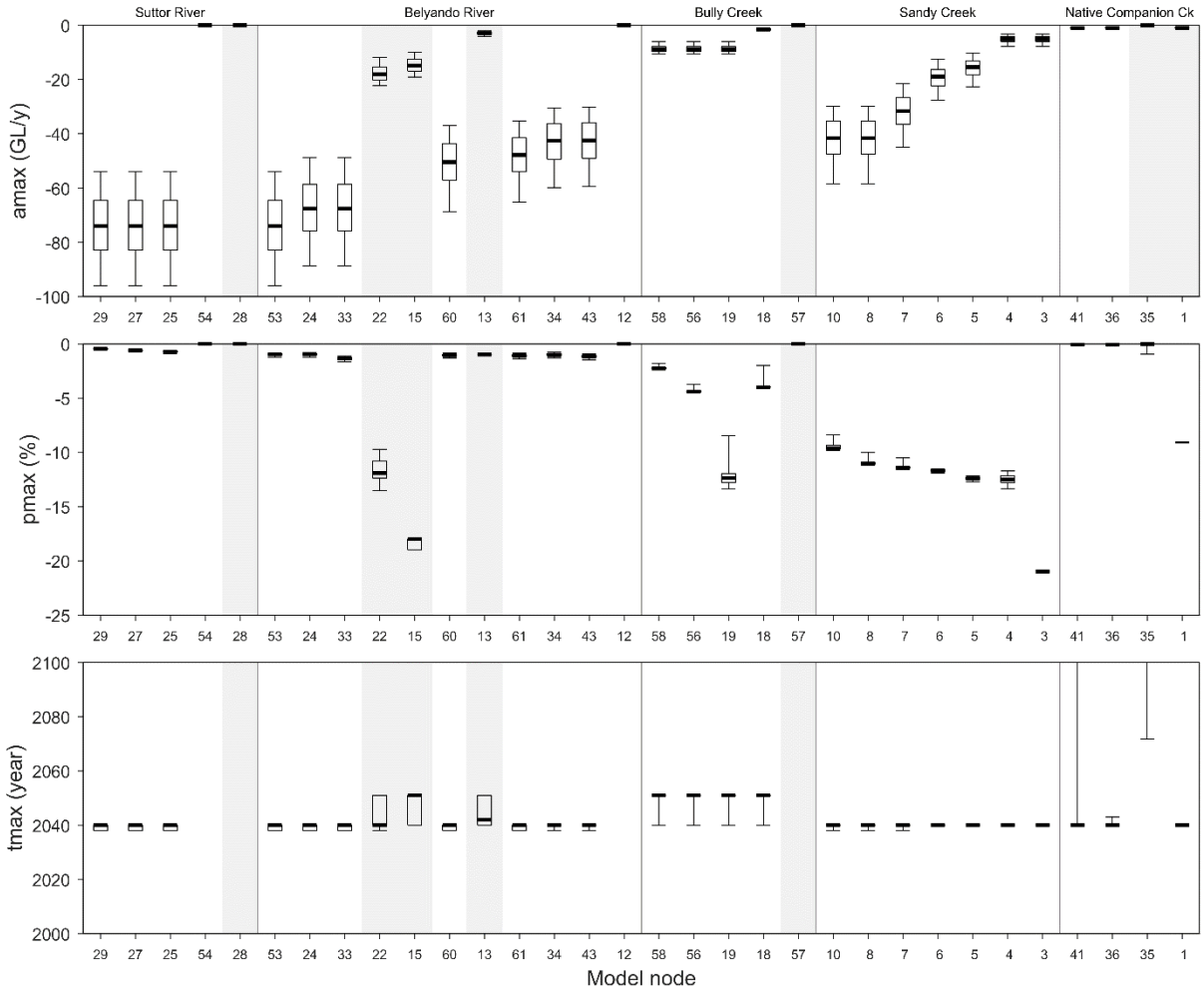


Figure 24. Maximum absolute change (amax), maximum percentage change (pmax) and time of maximum change (tmax) in hydrological response variable 1 (annual flow) at selected model nodes within the Suttor River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

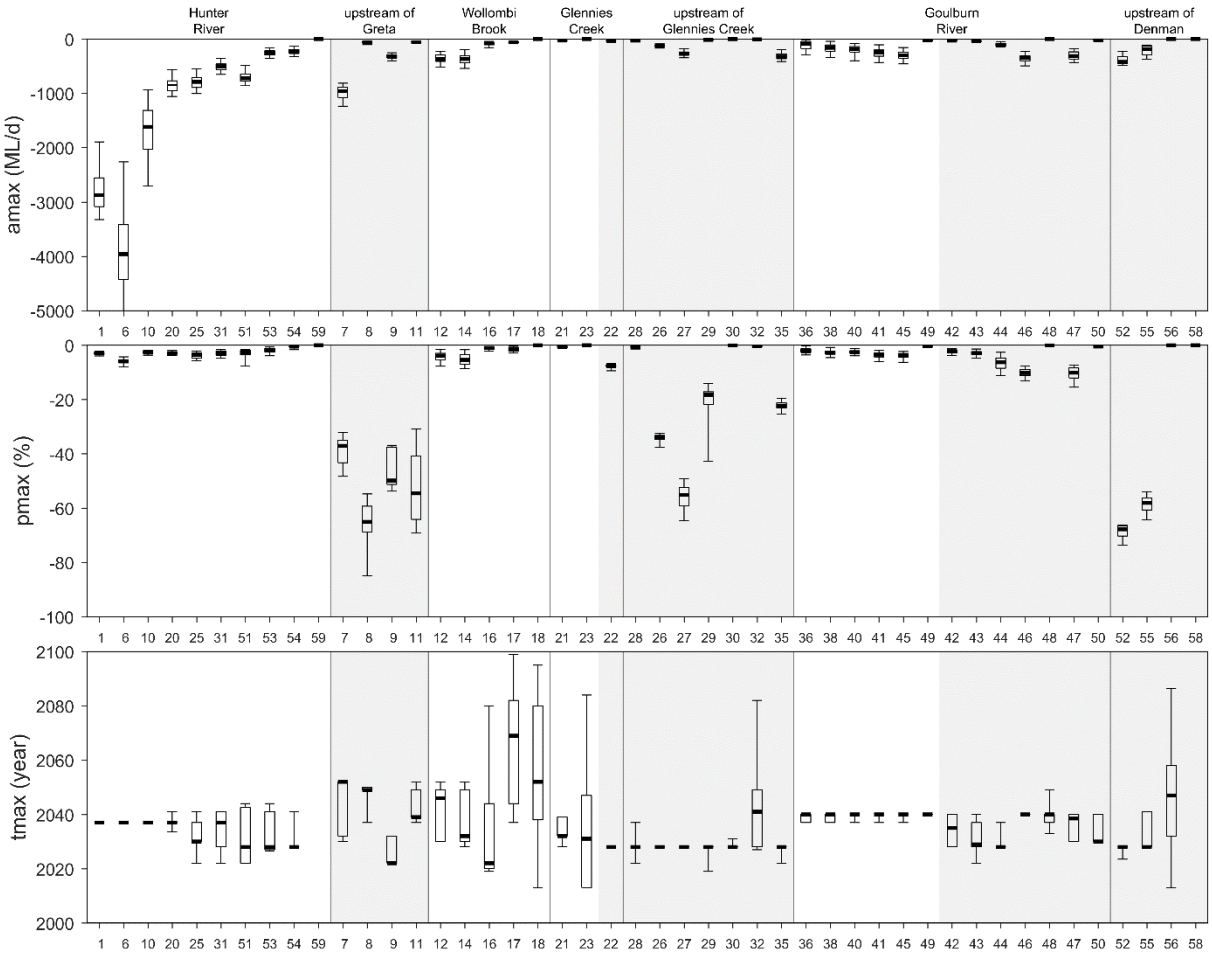


Figure 25. Maximum absolute change (amax), maximum percentage change (pmax) and time of maximum change (tmax) in hydrological response variable 2 (99th percentile flow) at selected model nodes within the Hunter River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

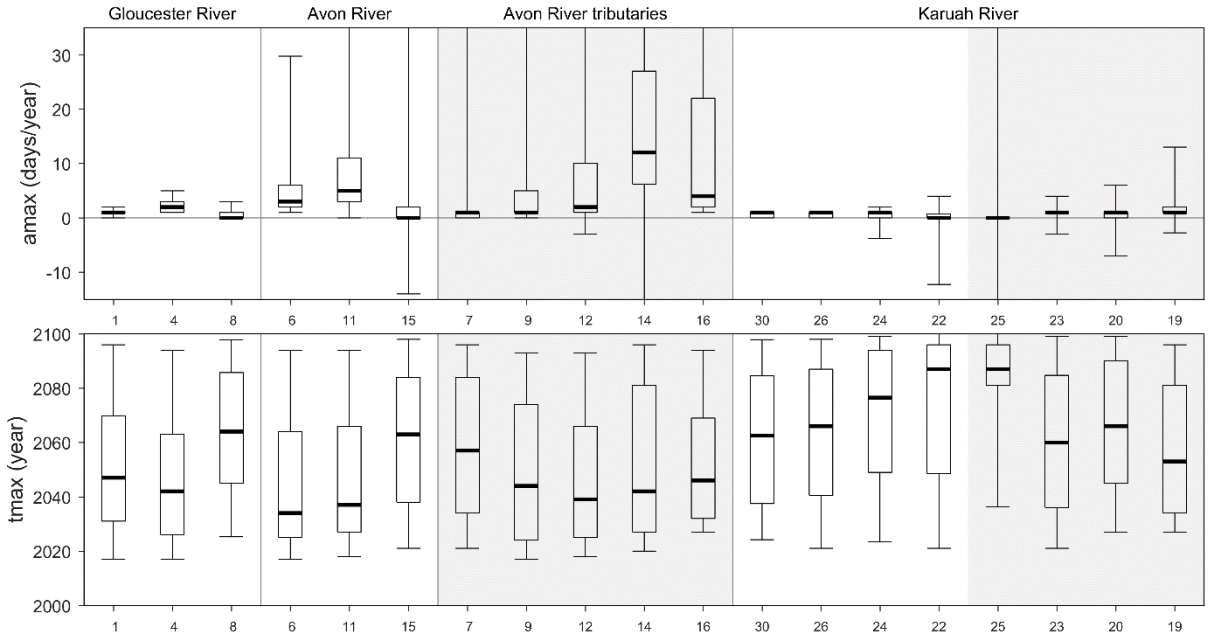


Figure 26. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 6 (number of low flow days) at 19 model nodes within the Gloucester-Karuah modelling domain. The streamflow network accumulates towards the left of the figure to nodes 1 and 30. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

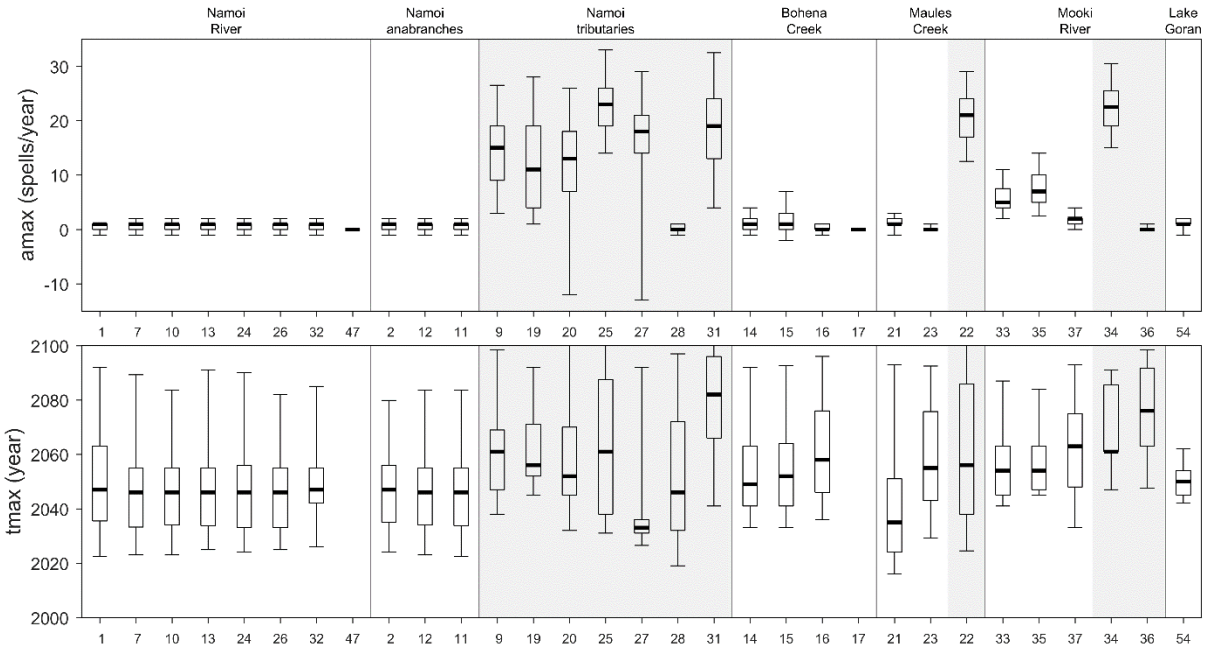


Figure 27. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 7 (number of low flow spells) at selected model nodes within the Namoi River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

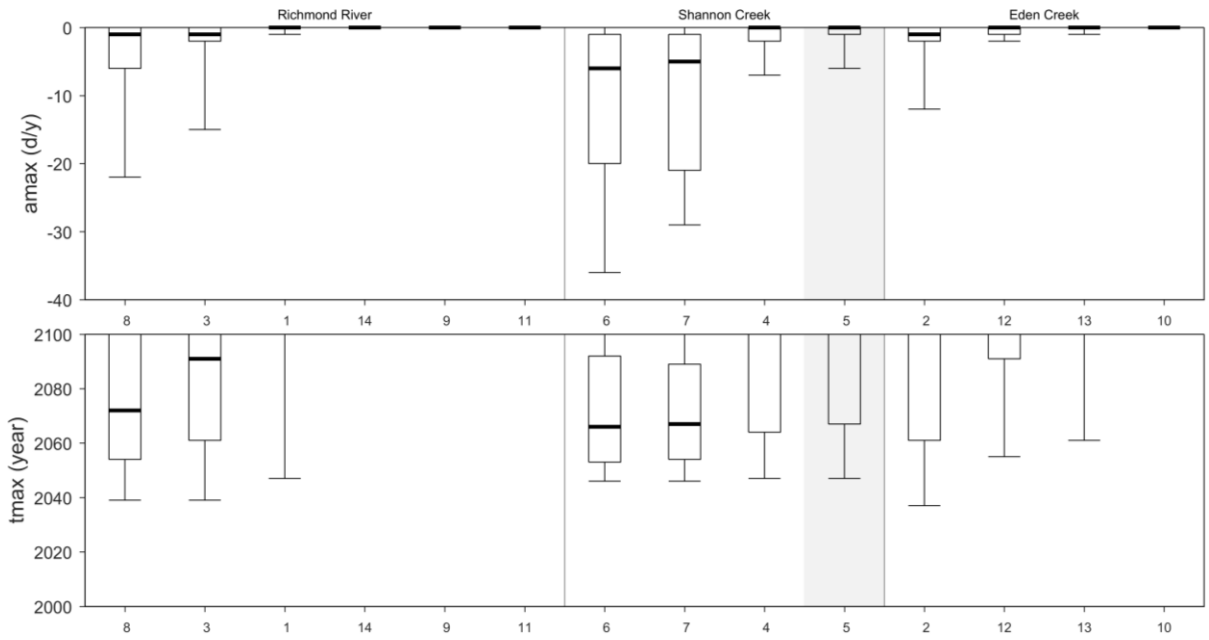


Figure 28. Maximum absolute change (amax) and time of maximum change (tmax) in hydrological response variable 8 (length of longest low flow spells) at 14 model nodes within the Richmond River catchment. The streamflow network accumulates towards the left of the figure. Shading indicates nodes that are on tributaries of the named watercourses. Boxplots show the 5th, 25th, 50th, 75th and 95th percentiles.

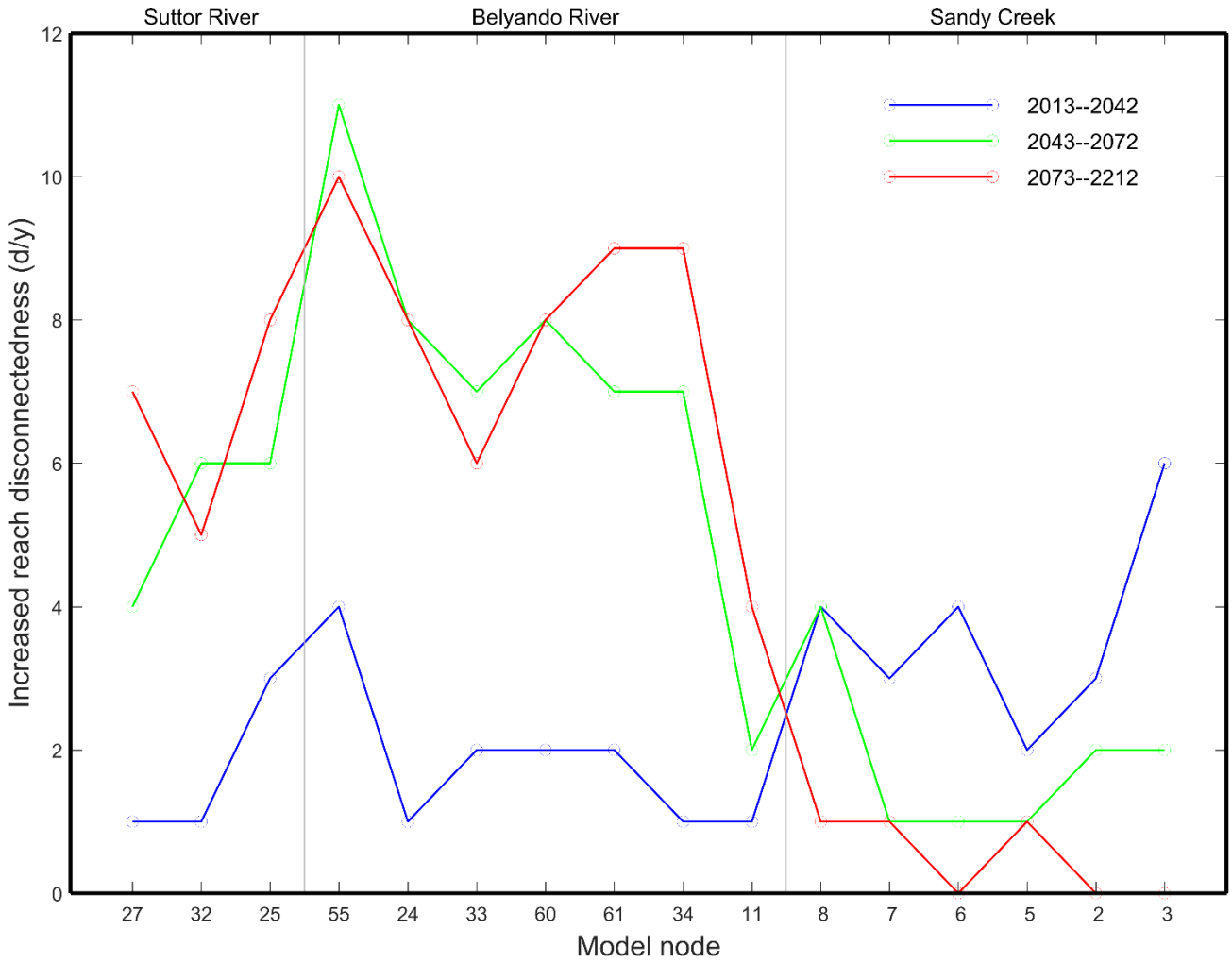


Figure 29. Average increase in hydrological response variable 12 (stream disconnectedness) over three 30-year time periods, expressed in average number of days per year, for a contiguous transect of model nodes in the Suttor River (nodes 27–25), the Belyando River (nodes 55–11) and Sandy Creek (nodes 8–3). The streamflow network accumulates towards the left of the figure.

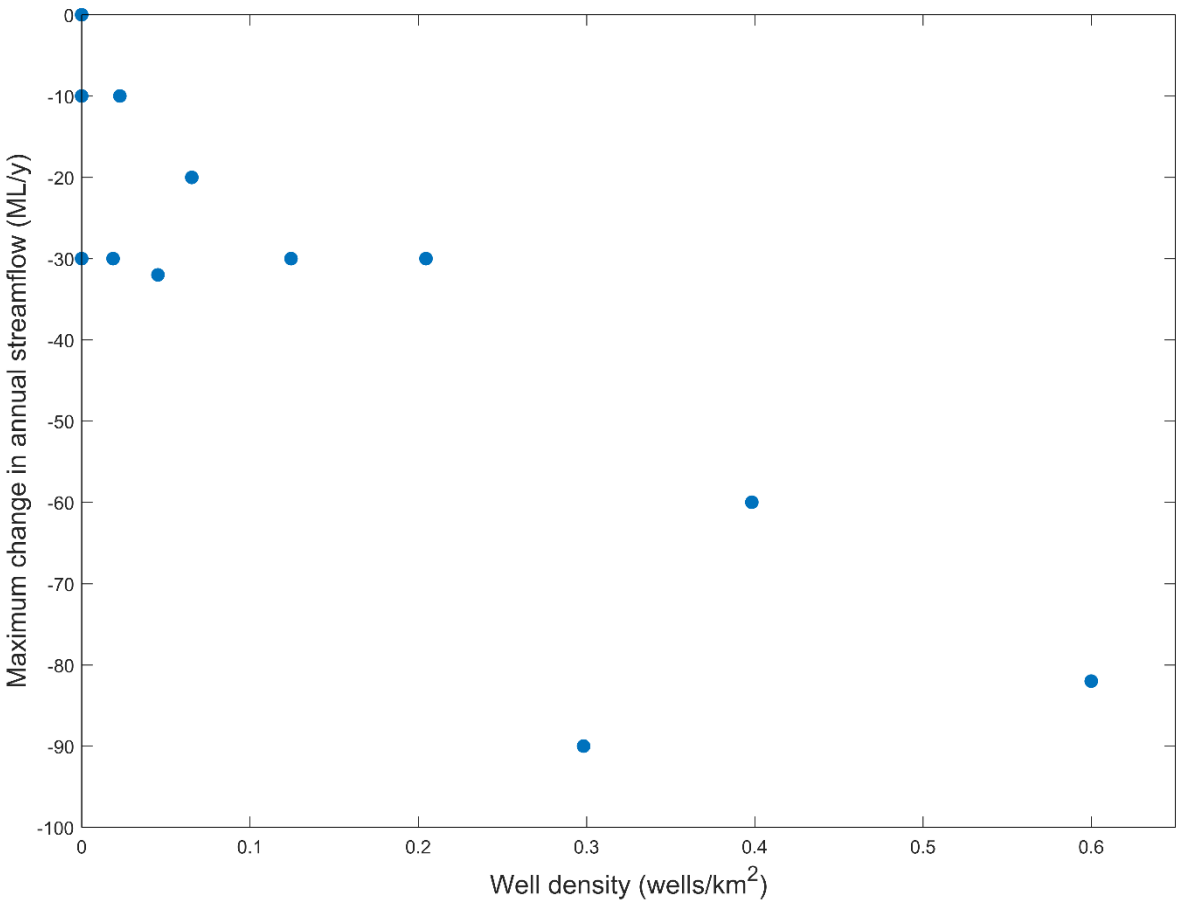


Figure 30. 95th percentile of maximum change in annual flow at model nodes in the Richmond River catchment as a function of the density of coal seam gas wells within 2 km of the node subcatchment.

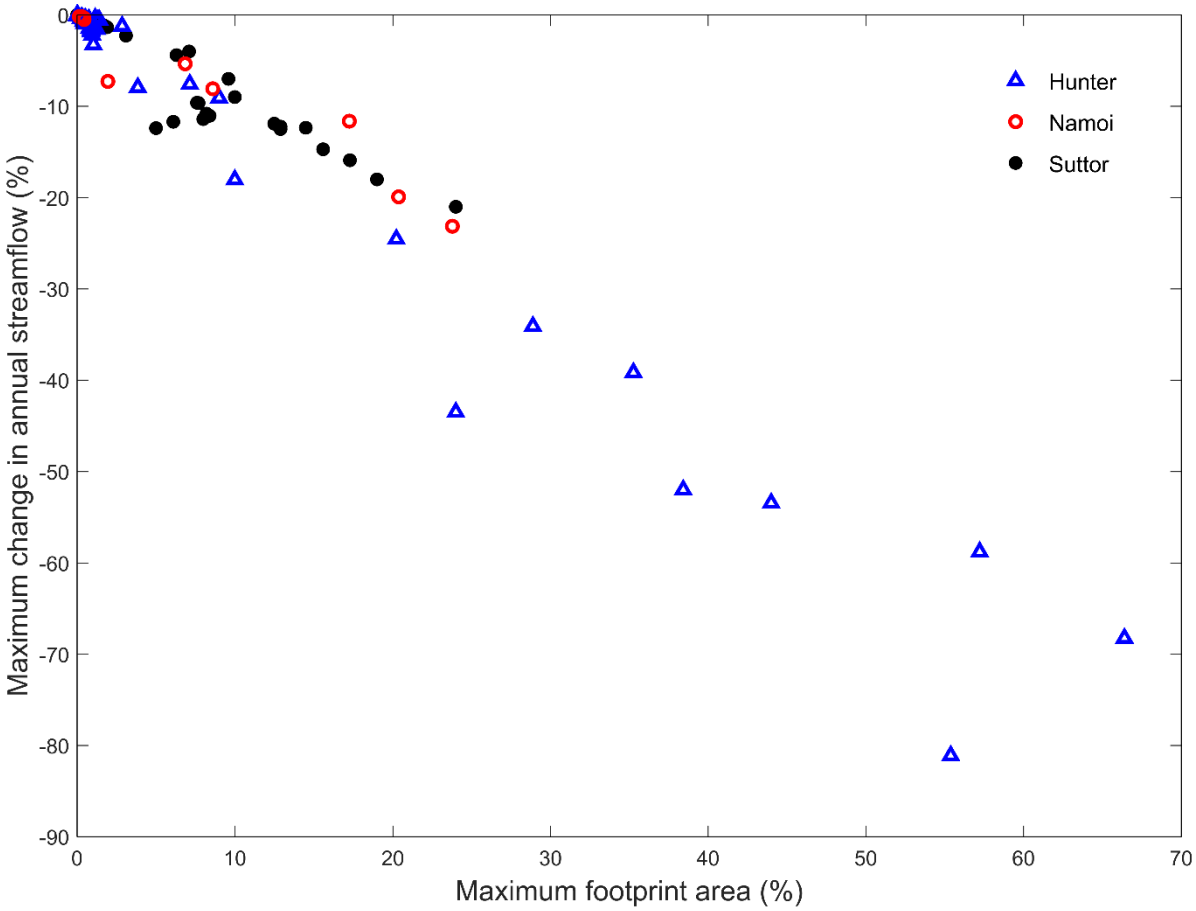


Figure 31. Predicted maximum change in annual streamflow as a function of the maximum proportion of mine footprint in a catchment for model nodes in the Hunter, Namoi and Suttor river basins. For the Namoi River basin, only model nodes with catchments further than 20 km from a proposed coal seam gas field are included.

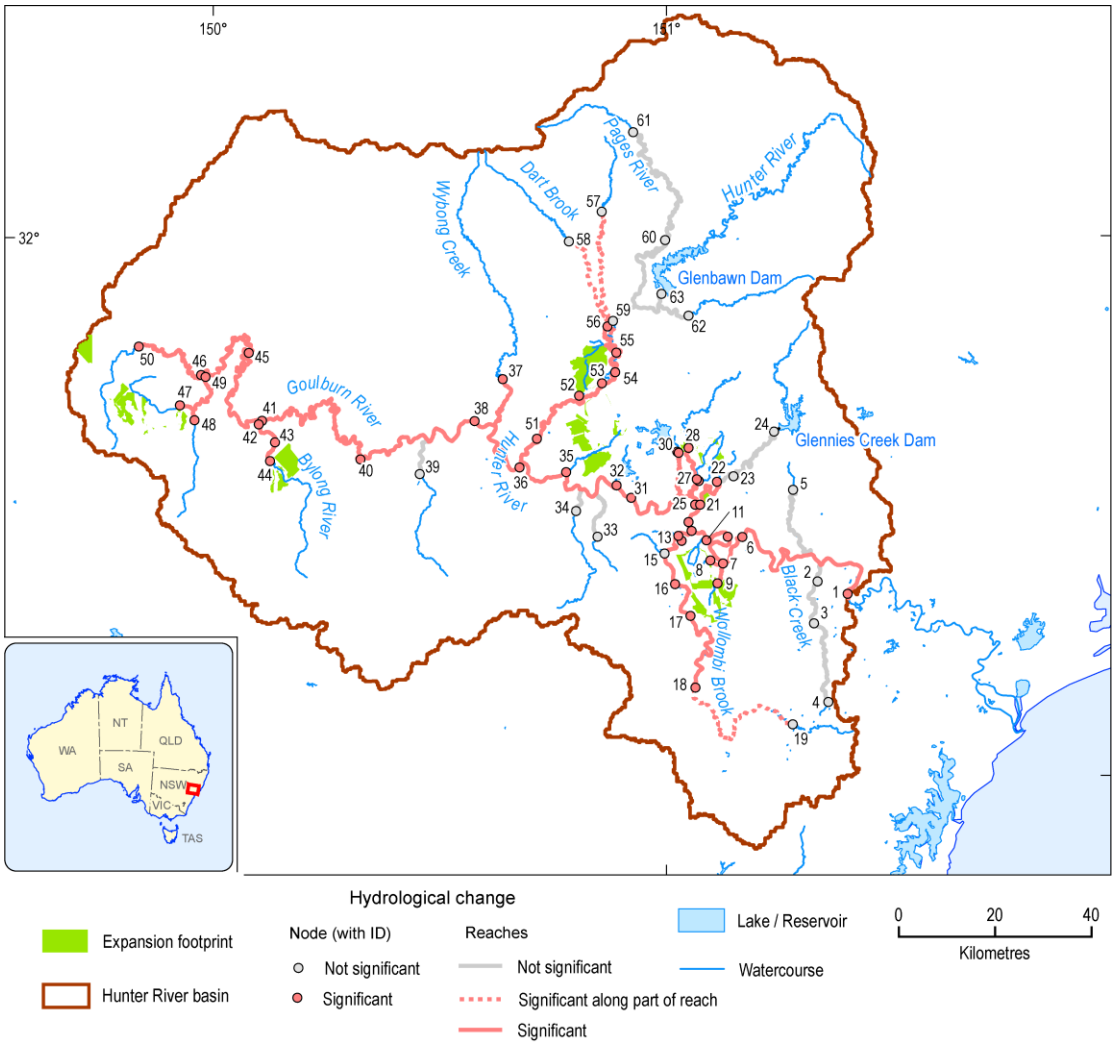


Figure 32. Reaches in the Hunter River catchment that are potentially susceptible to changes in streamflow characteristics under the expansion scenario.

Assumption or model choice	Data	Resources	Technical	Effect on predictions
Selection of calibration catchments	Medium	Low	Low	Low
High-flow and low-flow objective function	Low	Low	High	Low
Selection of goodness-of-fit function for each hydrological response variable	Low	Low	Low	Low
Selection of acceptance threshold for uncertainty analysis	Medium	High	Medium	Medium
Interaction with the groundwater model	Medium	Medium	High	Low
Implementation of the expansion scenario	High	Low	Low	High

Table 2. Qualitative uncertainty analysis for the surface water modelling in the Namoi modelling domain.

1077