Case Study

Understanding hydrological flow paths in conceptual catchment models using uncertainty and sensitivity analysis

Eva M. Mockler\textsuperscript{a, e}, Fiachra E. O’Loughlin\textsuperscript{b}, Michael Bruen\textsuperscript{a}

\textsuperscript{a} UCD Dooge Centre for Water Resources Research and UCD Earth Institute, University College Dublin, Dublin 4, Ireland
\textsuperscript{b} School of Geographical Sciences, University of Bristol, Bristol BS8 1SS, United Kingdom

A R T I C L E  I N F O
Article history:
Received 16 February 2015
Received in revised form 27 August 2015
Accepted 31 August 2015
Available online 3 September 2015

Keywords:
Conceptual catchment modeling
Hydrological processes
Groundwater
Sensitivity analysis
Parameter identification

A B S T R A C T
Increasing pressures on water quality due to intensification of agriculture have raised demands for environmental modeling to accurately simulate the movement of diffuse (nonpoint) nutrients in catchments. As hydrological flows drive the movement and attenuation of nutrients, individual hydrological processes in models should be adequately represented for water quality simulations to be meaningful. In particular, the relative contribution of groundwater and surface runoff to rivers is of interest, as increasing nitrate concentrations are linked to higher groundwater discharges. These requirements for hydrological modeling of groundwater contribution to rivers initiated this assessment of internal flow path partitioning in conceptual hydrological models.

In this study, a variance based sensitivity analysis method was used to investigate parameter sensitivities and flow partitioning of three conceptual hydrological models simulating 31 Irish catchments. We compared two established conceptual hydrological models (NAM and SMARG) and a new model (SMART), produced especially for water quality modeling. In addition to the criteria that assess streamflow simulations, a ratio of average groundwater contribution to total streamflow was calculated for all simulations over the 16 year study period. As observations time-series of groundwater contributions to streamflow are not available at catchment scale, the groundwater ratios were evaluated against average annual indices of base flow and deep groundwater flow for each catchment. The exploration of sensitivities of internal flow path partitioning was a specific focus to assist in evaluating model performances. Results highlight that model structure has a strong impact on simulated groundwater flow paths. Sensitivity to the internal pathways in the models are not reflected in the performance criteria results. This demonstrates that simulated groundwater contribution should be constrained by independent data to ensure results within realistic bounds if such models are to be used in the broader environmental sustainability decision making context.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

In the natural environment, hydrological flows exist as a continuum throughout the surface landscape and subsurface formations. Hydrological models attempt to capture the dominant processes in a catchment to predict river flows. For practical reasons, this flow continuum is simplified into discrete flow paths to facilitate conceptual understanding, model development and data analysis. The number of flow paths identified can depend on the catchment characteristics and the ultimate objective of the investigation, with two to four flow paths typically representing responses of flow processes reaching a river (e.g. SMARG [Kachroo, 1992; Khan, 1986; Tan and O’Connor, 1996], HBV [Bergrøn, 1995], NAM [Nielsen and Hansen, 1973] and PRMS [Leavesley et al., 1983, 1996]).

The merits of conceptual, parametrically parsimonious, hydrological models for investigating the dominant pathways and processes in catchments have been widely discussed (e.g. Refsgaard and Henriksen, 2004; Sivapalan, 2003). Model parameter identification is a fundamental challenge for hydrologists (Duan et al., 2006; Sivapalan, 2003). The presence of parameter interactions in conceptual rainfall–runoff (CRR) models can make a priori parameter prediction methods unreliable (Wagener and Wheater, 2006). Ideally, a model should be parametrically parsimonious while still capturing the dominant processes of the catchment with limited parameter interactions. Many hydrological models have been developed and used for decades for both research and operational hydrology. However, new model structures are still being developed to incorporate new conceptual understanding of specific catchment processes and places (Beven, 1999), and to facilitate the demands of new pressures on water resources, including nutrient enrichment (Futter et al., 2014).

http://dx.doi.org/10.1016/j.cageo.2015.08.015
0098-3004/© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
There is a growing body of literature investigating model structure uncertainty (Breuer et al., 2009; Clark et al., 2008; Gupta et al., 2012; Kavetski and Fenicia, 2011; Wagener et al., 2001). The focus is increasingly turning to the internal movement of water within these conceptual models to investigate if each of the simulated processes contributing to the total flows are realistic (e.g. Fenicia et al., 2011; Kokkenen and Jakeman, 2001). This hydrological partitioning is particularly important when coupling flow simulations with water quality, as the flow path can have a significant effect on solute transport and attenuation (Futter et al., 2014; Medici et al., 2012). Typically, particulate phosphorus is delivered via overland flow (Jordan et al., 2005). Nitrate is typically delivered to streams via subsurface pathways, with links between increasing nitrate concentrations and groundwater contributions (Tesoriero et al., 2009).

Sensitivity analysis (SA) is “the principal evaluation tool for characterizing the most and least important sources of uncertainty in environmental models” (U.S.EPA, 2009). The central role of sensitivity analysis for testing and implementing environmental models is widely noted (Refsgaard et al., 2007; U.S.EPA, 2009). Sensitivity analysis of parameters of water quality models has been undertaken using one-at-a-time sensitivity analysis (e.g. Morris, 1991) with Latin Hypercube Sampling, for example, for simulating dissolved oxygen with the ESWAT model (Vandenbergh et al., 2001) and nitrogen with the INCA-N model (Rankinen et al., 2013), or other Monte Carlo methods (e.g. McIntyre et al., 2005; Sánchez-Canales et al., 2015). More recently, variance based sensitivity methods have been employed for the parameters of SWAT (Nossent et al., 2011; Zhang et al., 2013).

Variance based sensitivity analysis (e.g. Sobol, 2001) is recommended as a superior method for which the computational effort is not prohibitive (Saltelli and Annoni, 2010; Tang et al., 2007; U.S.EPA, 2009). For non-linear conceptual hydrological models such as those investigated in this study, variance based methods are ideal to investigate the parameter sensitivities and interactions in the global parameter space (e.g. O’Loughlin et al., 2013; van Werkhoven et al., 2008, 2009; Zhan et al., 2013).

The aim of the study was to identify a suitable hydrological model that can represent the internal flow paths in Irish catchments. In this paper, a new model that was developed with a focus on sub-surface flow paths, SMART, is compared with two well-established conceptual models, NAM and SMARG. The parameters and internal flow paths the three models are compared using (i) an uncertainty analysis and (ii) a variance based sensitivity analysis method. The analysis is carried out on multiple metrics of the three models simulating a 16 year period in 31 Irish catchments.

2. Data

2.1. Catchment data

The majority of Ireland’s area (70,000 km²) has central, gently undulating lowlands of elevations generally less than 150 m above
sea level, with areas of higher elevations near coasts. Annual rainfall varies from in excess of 3000 mm in the western mountains to less than 800 mm along the east coast. Mean annual temperatures range between 9 °C and 10 °C.

The 31 study catchments (Fig. 1, Table 1) were selected on the basis of having good quality meteorological and hydrometric data available for the 16 year study period beginning from 1 January 1990. The chosen catchments cover over 35% of the country and represent a variety of meteorological and geological conditions across Ireland, with catchment areas ranging from 151 km² to 2460 km².

Meteorological data consisted of daily rainfall and potential evapotranspiration values obtained from the Irish meteorology office, Met Éireann. The catchment-area averaged rainfall was calculated using the Thiessen method, with each catchment using data from at least two precipitation stations and the largest catchment (Boyne) containing 13 stations. Annual average rainfall (AAR) ranges from 820 mm in the Ryewater to 1897 mm in the Flesk with an overall average of 1189 mm. Potential evapotranspiration was obtained from ten stations distributed over the study area, with data from the nearest station selected for each catchment and assumed to be spatially uniform.

Hydrometric data for each catchment consisted of daily mean flows originating from the Irish Office of Public Works (OPW) and the Irish Environmental Protection Agency (EPA). Periods within the 16 years of the study with missing flow data at the catchment outlet were not included in the analysis. Of the 31 catchments, four have missing flow data for over 25% of the study period, with the majority having less than 10% missing values (Table 1).

### Table 1

<table>
<thead>
<tr>
<th>ID</th>
<th>WATERBODY</th>
<th>Area (km²)</th>
<th>AAR (mm yr⁻¹)</th>
<th>No. rain gauges</th>
<th>Mean Q (m³ s⁻¹)</th>
<th>% Missing Q data</th>
<th>ReCo</th>
<th>BFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Anner</td>
<td>437</td>
<td>913</td>
<td>3</td>
<td>6.8</td>
<td>3.6</td>
<td>0.39</td>
<td>0.62</td>
</tr>
<tr>
<td>2</td>
<td>Aughrim</td>
<td>203</td>
<td>1423</td>
<td>3</td>
<td>5.7</td>
<td>31.2</td>
<td>0.19</td>
<td>0.79</td>
</tr>
<tr>
<td>3</td>
<td>Bandon</td>
<td>424</td>
<td>1576</td>
<td>2</td>
<td>14.9</td>
<td>6.7</td>
<td>0.15</td>
<td>0.53</td>
</tr>
<tr>
<td>4</td>
<td>Barrow</td>
<td>2419</td>
<td>865</td>
<td>11</td>
<td>33.2</td>
<td>0.8</td>
<td>0.32</td>
<td>0.67</td>
</tr>
<tr>
<td>5</td>
<td>Blackwater</td>
<td>2334</td>
<td>1255</td>
<td>12</td>
<td>59.7</td>
<td>4.6</td>
<td>0.24</td>
<td>0.62</td>
</tr>
<tr>
<td>6</td>
<td>Bonet</td>
<td>264</td>
<td>1670</td>
<td>2</td>
<td>10.8</td>
<td>20.4</td>
<td>0.38</td>
<td>0.35</td>
</tr>
<tr>
<td>7</td>
<td>Boyne</td>
<td>2460</td>
<td>903</td>
<td>13</td>
<td>37.8</td>
<td>0.6</td>
<td>0.29</td>
<td>0.68</td>
</tr>
<tr>
<td>8</td>
<td>Bride</td>
<td>334</td>
<td>1305</td>
<td>5</td>
<td>9.5</td>
<td>1.6</td>
<td>0.28</td>
<td>0.64</td>
</tr>
<tr>
<td>9</td>
<td>Camlin</td>
<td>253</td>
<td>884</td>
<td>4</td>
<td>3.9</td>
<td>8.9</td>
<td>0.22</td>
<td>0.54</td>
</tr>
<tr>
<td>10</td>
<td>Clare</td>
<td>700</td>
<td>1146</td>
<td>3</td>
<td>16</td>
<td>11.8</td>
<td>0.3</td>
<td>0.61</td>
</tr>
<tr>
<td>11</td>
<td>Clodagh</td>
<td>254</td>
<td>904</td>
<td>2</td>
<td>3.9</td>
<td>12.5</td>
<td>0.29</td>
<td>0.61</td>
</tr>
<tr>
<td>12</td>
<td>Dee</td>
<td>334</td>
<td>918</td>
<td>3</td>
<td>4.3</td>
<td>2.9</td>
<td>0.21</td>
<td>0.62</td>
</tr>
<tr>
<td>13</td>
<td>Deel Moy</td>
<td>151</td>
<td>1922</td>
<td>4</td>
<td>6.7</td>
<td>12.9</td>
<td>0.09</td>
<td>0.33</td>
</tr>
<tr>
<td>14</td>
<td>Deel Munster</td>
<td>439</td>
<td>1191</td>
<td>2</td>
<td>10.7</td>
<td>26</td>
<td>0.17</td>
<td>0.41</td>
</tr>
<tr>
<td>15</td>
<td>Erne</td>
<td>1492</td>
<td>1008</td>
<td>3</td>
<td>30.2</td>
<td>20.6</td>
<td>0.13</td>
<td>0.79</td>
</tr>
<tr>
<td>16</td>
<td>Feale</td>
<td>647</td>
<td>1532</td>
<td>3</td>
<td>22</td>
<td>14</td>
<td>0.15</td>
<td>0.31</td>
</tr>
<tr>
<td>17</td>
<td>Fergus</td>
<td>511</td>
<td>1135</td>
<td>2</td>
<td>10.4</td>
<td>2.3</td>
<td>0.51</td>
<td>0.7</td>
</tr>
<tr>
<td>18</td>
<td>Flesk</td>
<td>329</td>
<td>1897</td>
<td>2</td>
<td>14.4</td>
<td>6.9</td>
<td>0.13</td>
<td>0.39</td>
</tr>
<tr>
<td>19</td>
<td>Graney</td>
<td>280</td>
<td>1384</td>
<td>2</td>
<td>7.7</td>
<td>4.1</td>
<td>0.11</td>
<td>0.54</td>
</tr>
<tr>
<td>20</td>
<td>Little Brosna</td>
<td>479</td>
<td>962</td>
<td>3</td>
<td>8.3</td>
<td>23.6</td>
<td>0.3</td>
<td>0.58</td>
</tr>
<tr>
<td>21</td>
<td>Maigue</td>
<td>763</td>
<td>1018</td>
<td>4</td>
<td>13</td>
<td>10.5</td>
<td>0.27</td>
<td>0.52</td>
</tr>
<tr>
<td>22</td>
<td>Moy</td>
<td>1975</td>
<td>1313</td>
<td>9</td>
<td>58.8</td>
<td>8.4</td>
<td>0.21</td>
<td>0.78</td>
</tr>
<tr>
<td>23</td>
<td>Mulkear</td>
<td>648</td>
<td>1244</td>
<td>3</td>
<td>15.5</td>
<td>8.7</td>
<td>0.16</td>
<td>0.52</td>
</tr>
<tr>
<td>24</td>
<td>Nenagh</td>
<td>293</td>
<td>1041</td>
<td>2</td>
<td>6.4</td>
<td>28.5</td>
<td>0.2</td>
<td>0.58</td>
</tr>
<tr>
<td>25</td>
<td>Nore</td>
<td>2418</td>
<td>962</td>
<td>18</td>
<td>39.6</td>
<td>0.8</td>
<td>0.32</td>
<td>0.63</td>
</tr>
<tr>
<td>26</td>
<td>Rinn</td>
<td>281</td>
<td>1027</td>
<td>2</td>
<td>5.7</td>
<td>32.6</td>
<td>0.08</td>
<td>0.61</td>
</tr>
<tr>
<td>27</td>
<td>Robe</td>
<td>238</td>
<td>1220</td>
<td>4</td>
<td>6.2</td>
<td>5.5</td>
<td>0.37</td>
<td>0.56</td>
</tr>
<tr>
<td>28</td>
<td>Ryewater</td>
<td>210</td>
<td>820</td>
<td>7</td>
<td>2.4</td>
<td>6.8</td>
<td>0.18</td>
<td>0.51</td>
</tr>
<tr>
<td>29</td>
<td>Shournagh</td>
<td>208</td>
<td>1213</td>
<td>4</td>
<td>5.1</td>
<td>28.6</td>
<td>0.23</td>
<td>0.7</td>
</tr>
<tr>
<td>30</td>
<td>Suck</td>
<td>1207</td>
<td>1061</td>
<td>8</td>
<td>25.2</td>
<td>18.2</td>
<td>0.26</td>
<td>0.65</td>
</tr>
<tr>
<td>31</td>
<td>Suir</td>
<td>1583</td>
<td>1113</td>
<td>3</td>
<td>34</td>
<td>2.1</td>
<td>0.3</td>
<td>0.63</td>
</tr>
</tbody>
</table>

#### 2.2. Catchment groundwater flow indices

Groundwater is the part of the sub-surface water that is in the saturated zone, which typically flows through aquifers, although it can expand with increasing moisture conditions to include flow through the subsoils and soils. Two indices representing groundwater flow are used in this study to indicate proportion of groundwater contributing to streamflow in catchments:

1. The groundwater recharge coefficient (ReCo) is calculated from the Geological Survey of Ireland (GSI) groundwater recharge map (Hunter Williams et al., 2013), and does not incorporate any streamflow time-series. ReCo represents the deep groundwater resource in a catchment. The main hydro-geological properties used to generate the map were soil drainage properties, subsoil permeability and subsoil thickness. For example, groundwater flow is predicted as low in areas overlain by thick, low permeability clay, and where low permeability aquifers are not able to accept percolating waters. ReCo is calculated as the predicted annual groundwater recharge (mm) as a percentage of the annual effective rainfall (mm).

2. The Base Flow Index (BFI) is a measure of the proportion of streamflow that is drawn from natural storages in the catchment. It was calculated from streamflow time-series by the Office of Public Works (OPW) using the 5-day minima method (Institute of Hydrology, 1980). BFI is greater than the recharge coefficient as it can include flow through soils and subsoils.
3. Methods

For three hydrological models, uncertainty and sensitivity analyses were undertaken on three model outputs (outlined in Section 3.2) generated from simulating each of the 31 catchments over the 16 year study period (Fig. 2).

3.1. Conceptual rainfall runoff models

Two established models (Sections 3.1.1 and 3.1.2), and a newly developed model (Section 3.1.3) were selected for this study. Further background information on hydrological models and applications can be found in Singh and Frevert (2005) and Beven (2012).

3.1.1. NAM model

The ‘Nedbør-Afstrømnings-Model’ (NAM) model (Nielsen and Hansen, 1973) is an internationally established model, and has previously been used in Irish catchments for investigating the contributions of groundwater and surface water to streamflow (Mockler and Bruen, 2013; O’Brien et al., 2013; RPS, 2008).

The NAM model has two storage reservoirs for soil moisture accounting and reservoirs representing four hydrological pathways (Fig. 3a). Some small amendments were made to reduce the original 15-parameter NAM model to a more parsimonious 11-parameter structure. These included (i) omitting the snow component from the structure, as it is not relevant to the Irish study catchments, (ii) relating the two quick flow routing parameters of two linear reservoirs in series to one parameter (SUPERCK), and (iii) fixing the groundwater contribution factor equal to one (following the assumption that groundwater transfers between catchments are negligible at this scale). Eight of these parameters control the moisture content in storages representing the surface, soil and groundwater storages, and three parameters relate to the routing components (Table 2a).

3.1.2. SMARG model

The Soil Moisture Accounting and Routing with Groundwater component (SMARG) model was developed in NUI Galway (Kachroo, 1992; Khan, 1986; Tan and O’Connor, 1996). Its origins are in the layers model (O’Connell et al., 1970) and its water balance component is based on the ‘Layers Water Balance Model’ (Nash and Sutcliffe, 1970). The SMARG model has been widely applied in Irish catchments (Bastola et al., 2011; Goswami et al., 2005; O’Brien et al., 2013; RPS, 2008).

SMARG has a soil moisture accounting component that represents the catchment as a vertical stack of soil layers. This component keeps account of the rainfall, evaporation, runoff, and soil storage processes using six parameters (Fig. 3b, Table 2b). When there is rainfall in a time step, the excess rainfall is calculated as the depth of water that exceeds potential evapotranspiration. This depth of water is used to calculate surface runoff, which is the sum of (i) direct runoff, (ii) infiltration excess, and (iii) a portion of saturation excess. The remainder of the saturation excess contributes to the groundwater, as determined by the groundwater weighting parameter (G). The routing component uses linear reservoirs with three parameters (Table 2b) to simulate the attenuation effects of the catchment.

3.1.3. SMART model structure

The SMART model was developed to facilitate water quality modeling in Irish catchments, and was informed by the strengths of the NAM and SMARG models. The model has six soil layers of equal depth (Fig. 3c, Table 2c), similar to the SMARG, with six soil moisture accounting parameters. Drain flow is included as a separate flow path in the model, as this can be an important pathway for nutrients in agricultural catchments (e.g. Madison et al., 2014), and is related to soil moisture excess and the drain parameter (S), which varies between 0 and 1. Interflow is a combination of soil moisture excess and outflow from the soil layers, calculated using the soil outflow coefficient (D). Shallow and deep groundwater pathways are each calculated from individual outflow equations, also related to the outflow coefficient (D) parameter. Further details on the SMART model development are available in Mockler et al. (2014).

3.2. Uncertainty analysis and evaluation criteria

A parameter uncertainty analysis was undertaken for each hydrological models using Latin Hypercube sampling of the ranges
outlined in Table 2a–c, assuming uniform probability distribution functions (see Fig. 2). In addition to analysis of the full set of sampled parameter sets, 1000 behavioral parameter sets were identified for each model based on the streamflow simulation performance. Similar Monte Carlo methods are frequently used to sample possible variations in inputs and parameters using assumed probability distribution functions e.g. the GLUE methodology (Beven and Binley, 1992).

In this study, we used two performance criteria to assess the adequacy of the simulation of total streamflow against the observed streamflow. The first is based on the Nash Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), a widely used goodness of fit measure based on the error variance defined as

$$\text{NSE} = 1 - \frac{\sum_{t=1}^{n}(Q_{o,t} - Q_{m,t})^2}{\sum_{t=1}^{n}(Q_{o,t} - \bar{Q}_o)^2},$$

where $Q_{o,t}$ is the observed flow for time-step $t$, $Q_{m,t}$ is the modeled flow at time-step $t$, $\bar{Q}_o$ is the mean observed flow and $n$ is the length of the time series. A bounded version of the Nash–Sutcliffe criterion (Mathevet et al., 2006) was calculated as

$$C_\text{IM} = \text{NSE}/(2 - \text{NSE}).$$

The $C_{2M}$ criterion varies between $-1$ and $1$ and is less optimistic for positive values compared to NSE, thereby generating a less skewed distribution.

The second criteria used is the mean residual error criterion (MR), which calculates the difference between simulated and observed flows in the overall water balance as

$$\text{MR} = \frac{1}{n} \sum_{t=1}^{n} (Q_{o,t} - Q_{m,t}),$$

where $Q_{o,t}$ is the observed flow for time-step $t$ and $Q_{m,t}$ is the modeled flow at time-step $t$. MR evaluates the overall water balance, whereas the NSE focuses on the correlation of the time series.

In addition to the criteria that assess streamflow simulations, a ratio of average groundwater contribution ($\text{GW}_{\text{avg}}$) to total streamflow was calculated for all simulations over the study period, as

$$\text{GW}_{\text{avg}} = \frac{1}{n} \frac{\sum_{t=1}^{n} (\text{GW}_{m,t})}{\sum_{t=1}^{n} (Q_{m,t})},$$

where $\text{GW}_{m,t}$ is the modeled groundwater flow at time-step $t$. 

**Fig. 3.** Schematic representation of the NAM (a), SMARG (b) and SMART (c) models, with internal flow paths identified as either quick flow (blue) or groundwater (green). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
Sobol’s method (Sobol, 1993) is a global sensitivity analysis (GSA) which decomposes the output variance into relative contributions from input parameters and interactions (e.g. Shin et al., 2013; Tang et al., 2007; van Werkhoven et al., 2009). A sensitivity index (SI) representing the importance of the driving variable \( X_i \) to the output \( Y \) can be defined as (Saltelli and Annoni, 2010)

\[
SI_i = \frac{Var(Y|X_i)}{Var(Y)}
\]

where \( Var(\cdot) \) and \( E(\cdot) \) are variance and expectation functions respectively. This is a measure of the first-order sensitivity indices (FSI) of each parameter on the model output, often referred to as the main effect (Saltelli et al., 2008). The total-order sensitivity indices (TSI) represents the total effect of a parameter including interactions, and can be defined as (Saltelli and Annoni, 2010)

\[
TSI_i = \frac{E(Var(Y|X_{\setminus i}))}{Var(Y)}
\]

where \( X_{\setminus i} \) denotes the matrix of all factors but \( X_i \). For this study, Saltelli’s scheme (e.g. Saltelli et al., 2008, 2010) was used to compute FSI and TSI with \( n(k + 2) \) Monte Carlo simulations, where \( k \) is the number of parameters and \( n \) is the initial sample size used (9000 in this study). This results in 99,000, 108,000 and 117,000 simulations for the 9, 10 and 11 parameter model, respectively. For each model, parameter sets were generated using Latin Hypercube sampling with uniform distributions following the ranges detailed in Table 2(a–c). Saltelli’s scheme was computed using the SAFE Toolbox (Pianosi et al., 2015) as a framework for assessment of the robustness and convergence of the sensitivity indices.

O’Loughlin et al. (2013) evaluated the parameters of the SMARG model with variance-based sensitivity analysis, and showed that sensitivities vary with time-step, flow regime and evaluation metric. In this study, all models use a daily time-step and are evaluated over the full range of flow regimes with the 16 year study period.

For the purpose of GSA, the GWavg ratio was treated as a model output, with SI and TSI calculated in the same manner as the C2M and MR criteria. As observations time-series of groundwater contributions to streamflow are not available at catchment scale, the groundwater ratios were evaluated against average annual indices of base flow (BFI) and deep groundwater flow (ReCo) for each catchment (described in Section 2.2).

The sensitivity indices were compared to AAR to see if parameters have a different importance based on hydrological regimes, as identified by AAR. The Spearman rank correlation coefficient \( (r) \) was preferred over Person’s \( R^2 \) to assess the relationships between sensitivity indices and AAR as variables may be non-normally distributed and a linear relationship between the variables was not assumed.

4. Results and discussion

4.1. Uncertainty analysis and hydrological model performance

4.1.1. Hydrograph simulations

Results from the simulations using the Latin Hypercube sampling parameter sets show that for each catchment, each model had some simulations that performed well at simulating streamflow, as evaluated by C2M (Fig. 4) and MR. The mean C2M results were 0.44, 0.18, 0.51 for the NAM, SMARG and SMART model, respectively, with results varying between catchments. The selection of parameter ranges (Table 2) and assumption of uniform distributions for the 9, 10 and 11 parameter model, respectively. For each model, parameter sets were generated using Latin Hypercube sampling with uniform distributions following the ranges detailed in Table 2(a–c). Saltelli’s scheme was computed using the SAFE Toolbox (Pianosi et al., 2015) as a framework for assessment of the robustness and convergence of the sensitivity indices.

O’Loughlin et al. (2013) evaluated the parameters of the SMARG model with variance-based sensitivity analysis, and showed that sensitivities vary with time-step, flow regime and evaluation metric. In this study, all models use a daily time-step and are evaluated over the full range of flow regimes with the 16 year study period.

For the purpose of GSA, the GWavg ratio was treated as a model output, with SI and TSI calculated in the same manner as the C2M and MR criteria. As observations time-series of groundwater contributions to streamflow are not available at catchment scale, the groundwater ratios were evaluated against average annual indices of base flow (BFI) and deep groundwater flow (ReCo) for each catchment (described in Section 2.2).

The sensitivity indices were compared to AAR to see if parameters have a different importance based on hydrological regimes, as identified by AAR. The Spearman rank correlation coefficient \( (r) \) was preferred over Person’s \( R^2 \) to assess the relationships between sensitivity indices and AAR as variables may be non-normally distributed and a linear relationship between the variables was not assumed.

3.3. Variance based global sensitivity analysis

Sobol’s method (Sobol, 1993) is a global sensitivity analysis (GSA) which decomposes the output variance into relative contributions from input parameters and interactions (e.g. Shin et al., 2013; Tang et al., 2007; van Werkhoven et al., 2009). A sensitivity index (SI) representing the importance of the driving variable \( i \) to the output \( Y \) can be defined as (Saltelli and Annoni, 2010)

\[
SI_i = \frac{Var(Y|X_i)}{Var(Y)}
\]

where \( Var(\cdot) \) and \( E(\cdot) \) are variance and expectation functions respectively. This is a measure of the first-order sensitivity indices (FSI) of each parameter on the model output, often referred to as the main effect (Saltelli et al., 2008). The total-order sensitivity indices (TSI) represents the total effect of a parameter including interactions, and can be defined as (Saltelli and Annoni, 2010)

\[
TSI_i = \frac{E(Var(Y|X_{\setminus i}))}{Var(Y)}
\]
4.1.2. Assessment of groundwater simulations

We further examined the 1000 behavioral simulations for each hydrological model, with a focus on groundwater contribution to streamflow. It is noteworthy that the behavioral sets were not selected using an objective function that optimizes groundwater simulations, such as the NSE with log values (Krause et al., 2005). Rather, this study aimed to assess the groundwater contribution of simulations that would be suitable for a range of low to high flows, as is required in catchment simulations for water quality (Futter et al., 2014; Medici et al., 2012). Moreover, the simulations were not constrained by the groundwater indices (BFI or ReCo) that were used in this assessment, and instead were used as independent evaluators.

Fig. 5 shows the distribution of GWavg by hydrological model for the 1000 behavioral sets, and for all model simulations using Latin Hypercube sampling. For each model, the distribution of GWavg for the total number of sampling sets and the 1000 behavioral sets are broadly similar. These results highlight that the majority of NAM model simulations have a lower contribution of groundwater than is indicated by both the ReCo, which represents the deep groundwater, and the BFI, which represents total baseflow contributions. The distributions of GWavg for the SMARG and SMART models are more closely aligned with the ReCo and BFI values.

The internal flow partitioning for each catchment was notably different across the models. To demonstrate this, the percentage of groundwater contributing to simulated streamflow was compared with catchment groundwater flow indices. Correlations between GWavg results and the BFI and ReCo (defined in Section 2.2) indicate whether the internal hydrological processes of the models are aligned with the understanding of processes from catchment characteristics. Of the three models, the SMART model had the strongest correlations with ReCo and BFI values across the catchments (Table 3). This indicates that the processes of SMART that produce quick in-stream responses and groundwater flow are more representative of what is expected from catchment characteristics.

The range of simulated GWavg for each catchment produced by the 1,000 simulations indicated the degree of uncertainty in attributing flow to quick flow or groundwater. The SMART had the widest prediction ranges, which tended to increase with increasing BFI (Fig. 8) i.e. greater uncertainty in groundwater dominated catchments. The SMART model produced tighter ranges of GWavg estimates (Table 3, Fig. 6), indicating that the SMART model has less uncertainty simulating internal processes, without providing any additional groundwater or baseflow information.

There is a growing body of literature highlighting the importance of assessing model structure adequacy (Breuer et al., 2009; Clark et al., 2008; Gupta et al., 2012; Wagener et al., 2001). In this study, the three conceptual models have different representations of surface runoff (or quick flow) and groundwater contributions to streamflow. All three models assume that the surface water and groundwater of the study catchments aligned, and that there are no transfers into or out of the catchment. This assumption may not be true, particularly for catchments with extensive subsurface paleochannels crossing catchment boundaries, or conduit karst aquifers i.e. Clare, Fergus, Robe and Suck catchments. The NAM and SMART models have more detailed internal flow partitioning compared to the SMARG model, and therefore are less flexible to adapt to different hydrological conditions. In particular, the internal flow paths of catchments with conduit karst aquifer bedrock may need to be interpreted, where a...
simulated quick flow path is actually representing the groundwater conduit flow.

4.2. Global sensitivity analysis

Assessing the convergence and uncertainty bounds (from bootstrap sampling) of the sensitivity indices identified the base sample size for Latin Hypercube sampling. Although convergence of SI and TSI was achieved for all models for the $C_{2M}$ and MR model output with a base sample of 5000, the SI and TSI values for $GW_{avg}$ output required an increased base sample size (9000) to achieve convergence. The base sample size of 9000 was selected which produced relatively tight uncertainty bounds (average TSI confidence intervals between 0.04 and 0.12) for the streamflow

Table 3
Summary of $C_{2M}$, MR and $GW_{avg}$ results for the 31 catchments and correlation of $GW_{avg}$ with catchment recharge coefficients (ReCo) and base flow indices (BFI) for 1000 behavioral simulations for 3 models.

<table>
<thead>
<tr>
<th>Model</th>
<th>$C_{2M}$ median (min, max)</th>
<th>MR median (min, max)</th>
<th>$GW_{avg}$ median (min, max)</th>
<th>$GW_{avg}$ corr with ReCo</th>
<th>$GW_{avg}$ corr with BFI</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAM</td>
<td>0.6 (0.34,0.84)</td>
<td>0.01 (−0.19,0.2)</td>
<td>0.12 (0.087)</td>
<td>0.18 (p=0.32)</td>
<td>0 (p=0.98)</td>
</tr>
<tr>
<td>SMARG</td>
<td>0.59 (0.21,0.9)</td>
<td>0 (−0.32,0.19)</td>
<td>0.33 (0.099)</td>
<td>0.37 (p=0.04)</td>
<td>0.66 (p=0.01)</td>
</tr>
<tr>
<td>SMART</td>
<td>0.65 (0.39,0.85)</td>
<td>0 (−0.17,0.08)</td>
<td>0.42 (0.07)</td>
<td>0.45 (p=0.01)</td>
<td>0.87 (p=0.01)</td>
</tr>
</tbody>
</table>

Fig. 5. Distribution of the fraction of groundwater contributing to total flow for 31 catchments from NAM (blue), SMARG (yellow) and SMART (red) for all model simulations (‘all’; dashed line) and behavioral sets (‘best’; solid line) using Latin Hypercube sampling of standard parameter ranges with uniform distributions. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 6. Simulated groundwater contribution ($GW_{avg}$) against BFI for each study catchment for the NAM (blue), SMARG (orange) and SMART (red) models. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

E.M. Mockler et al. / Computers & Geosciences 90 (2016) 66–77 73
performance criteria (C2M and MR), as indicated by the standard deviation results from bootstrap sampling (Table 4). This is in line with average TSI confidence intervals values reported in similar studies e.g. Tang et al. (2007) (values between 0.0 and 0.12 for daily time-step) and (Nossent et al., 2011) (values between 0.1 and 0.14). The uncertainty ranges for GWavg indices were wider (twice the standard deviation for C2M, see Table 4), therefore the rank of sensitive parameters are uncertain. Nonetheless, the sensitive parameters were still clearly identifiable from the results (Fig. 7) and this was deemed sufficient for a general parameter assessment to support the uncertainty analysis.

In the following sections, long-term average Sobol’s sensitivity results from the 16 year study period are presented for each model in turn. Fig. 7 shows the TSI for all model parameters across the 31 study catchments in color-coded grids with blue indicating low values (< 0.1) and orange indicating high values (> 0.8). Some TSI had slightly negative values, particularly with C2M model output, which were attributed to numerical errors in the Saltelli method. As these occur for sensitivity indices when the analytical sensitivity indices are close to zero, only unimportant factors are affected (Saltelli et al., 2008). Changes in the sampling ranges of sensitive parameters can impact sensitivity results (Shin et al., 2013). In order to ensure comparable results across catchments in this study, broad parameter ranges that include suitable values for all study catchments were selected.

van Werkhoven et al. (2008) used Sobol’s sensitivity analysis to assess the parameters of a conceptual rainfall–runoff model across a hydroclimatic gradient with a narrower range of AAR compared to those of this study, but a wider range of annual potential evapotranspiration. Similar to results from that study, Fig. 7 shows that, across all of the models, less parameters have notable SI values for the long term average model output (MR), compared to the peak-fitting evaluation criteria (C2M). Similar to findings of Zhan et al. (2013), routing parameters of the NAM (SUPERCK), SMARG (SRK) and SMART (SK) models are sensitive when evaluated by NSE based output (C2M results; Fig. 7).

4.2.1. NAM sensitivity results

The quick flow routing parameter (SUPERCK) is the most sensitive parameter when evaluated with C2M. Correlations between TSI and catchment AAR (r = 0.78, p < 0.001; Fig. 8), suggest that the SUPERCK parameter is more identifiable in wetter catchments. Contrasting trends are evident between first-order and higher-order indices for SUPERCK, indicating that when the parameter is identifiable, parameter interactions are reduced.

The highest TSIs for NAM’s GWavg output are the upper layer storage capacity (UMAX) and the interflo news coefficient (CQIF). CQIF has prominent sensitivities for the groundwater flow processes, even though this parameter is not directly in the process equations. This is because all of the internal flow paths in the NAM model, including groundwater, are proportional to the relative volume in the lower zone store, which is related to the interflo news parameters (CQIF and CLIF). Therefore, the interflo news parameters are sensitive with respect to many of the internal processes in NAM (Fig. 7). This soil moisture accounting mechanism does not represent the natural draining mechanisms in catchment, as the lower zone store can only be depleted by evapotranspiration.

4.2.2. SMARG sensitivity results

The quick flow routing parameter of the Nash gamma function (SRK) is the most sensitive when evaluated by the C2M and MR criteria. When evaluated by MR, the SRK sensitivities are due to the curtailment of the unit pulse response function when inappropriate parameter values are selected during optimization (Goswami and O’Connor, 2010). The sensitivities only relate to high flows (as seen in O’Loughlin et al., 2013), and are present across all catchments, regardless of appropriate impulse response function memory length. TSI values for SRK have positive correlations with AAR for MR output (r = 0.89, p < 0.001; Fig. 8), indicating that the routing lag parameter is more sensitive in wetter catchments. An opposing trend is seen for TSI values for the evapotranspiration coefficient (T) which is an adjustment factor for potential evapotranspiration input data, as the T parameter is more sensitive in drier catchments. The maximum infiltration capacity parameter (Y) and evaporation decay parameter (C) are not sensitive for any of the three criteria (also seen in O’Loughlin et al., 2013). The soil layer depth parameter (Z) also shows very low SI values across all the study catchments. This is due to the structure of equations of the SMARG model, and may result in difficulty identifying parameter values, particularly for temperate climate conditions such as Ireland.

The ‘direct runoff’ coefficient (H) and groundwater weighting parameter (G) are prominent in the groundwater flow evaluation (GWavg), as these parameters determine the internal split of flows in the model. They are not notable in the evaluation by model performance (C2M and MR) as these are calculated on total flows.

4.2.3. SMART sensitivity results

For the SMART model, the catchment rainfall correction coefficient (T) consistently shows high TSI values for the MR output as this parameter adjusts the precipitation input to match the observed flows. The impact of this parameter increases as the volume of precipitation increases, with TSI values positively correlated with AAR (r = 0.87, p < 0.01; Fig. 8). Trends in values of TSI are less defined for C2M evaluation, compared to MR, with moderate values for the soil moisture outflow parameter (S) and soil moisture capacity parameter (Z). As subsurface flow partitioning is determined from drainage calculated from the soil moisture storage layers, the S and Z have high TSI values across all catchment when evaluated by the GWavg output.

The SMART model was developed with a focus on sub-surface flow paths, driven by challenges of simulating diffuse nutrient impacts on water quality. As the model development was
informed from components of the NAM and SMARG models, a comparison of model structures is of interest. The SMART model structure was designed on the concept of the SMARG soil moisture layers, with the following notable changes:

1. The maximum infiltration capacity parameter ($Y$) of the SMARG model was not included in the SMART model as it is a difficult parameter to estimate at catchment scale. Instead, the infiltration excess process was conceptually combined with ‘direct’ runoff.

2. The evaporation coefficient ($C$) and soil moisture capacity parameter ($Z$) of the SMARG model showed very low first order sensitivities (both with an average TSI of 0 for all evaluation

---

Fig. 7. Total-order sensitivity indices (TSI) evaluated by mean residual (MR), bounded NSE ($C_{2M}$) and average groundwater fraction (GWavg) model output for the NAM (top row), SMARG (middle row) and SMART (bottom row).
criteria), indicating poorly identifiable parameters. The SMART model structure incorporates the soil moisture layer concept of the SMARG model, but the structural changes included redefining the soil layers and outflows. This resulted in increases in TSI for the Z parameter from an average of 0 in the SMARG model to 0.18 in the SMART (Fig. 7: SMART results). As evaporation is calculated from the soil layers, there was also an increase in average catchment TSI for the C parameter from 0 in the SMARG model to 0.12 (Fig. 7: MR SMART results).

3. Linear reservoirs were selected for routing quick flow, similar to the NAM model, in place of the Nash cascade routing component of the SMARG model. This resulted in reduced parameter interactions and the conservation of volume in the routing component of the SMART model.

5. Conclusions

For coupled water quantity and water quality modeling, a hydrological model is required that can capture both the total flows and the groundwater contributions to streamflow. A comparison of results from Monte Carlo simulations of 31 study catchments for the NAM, SMART and SMARG models highlighted that the relative contribution of groundwater depends on both the model structure and the catchment characteristics. Results showed that the new SMART model was superior to the two established models, NAM and SMARG, at representing both the total streamflow and the internal flow paths of the 31 Irish study catchments.

The SMART model development was influenced by the favorable aspects of the SMARG and NAM structures to enhance model parameter identification while maintaining a structure that can properly identify overland, interflow, upper and lower groundwater as discrete flow paths contributing to streamflow. Results from Sobol’s sensitivity method confirmed that the SMART model development reduced the number of poorly identifiable parameters, compared to the SMARG model.

Internal flow partitioning varies greatly between models and, to varying degrees, between behavioral parameter sets for each model. This study illustrated this by comparing the simulated annual groundwater contributions to streamflow with additional independent information, in the form of groundwater flow indices. For studies interested coupling water quality and hydrological simulations, it is recommended that an appropriate model structure is selected and, where available, additional information on plausible groundwater flow contributions is incorporated into model calibration.

Acknowledgments

The data for this paper are available from Met Eireann, OPW, and the Irish EPA. This work was supported by the Environmental Protection Agency Research Programme through the STRIVE Pathways Project (2007-WQ-CD-1-S1) and the CatchmentTools Project (2013-W-FS-14). Dr Fiachra O’Loughlin is funded by the Leverhulme Trust (RPG-409). The authors would like to thank B.Elssasser for providing data for this research. We are particularly grateful to the anonymous reviewers who considerably helped in improving the manuscript.

References


Goswami, M., O’Connor, K.M., 2010. A “monster” that made the SMAR conceptual model development was in Pathways Project (2007-WQ-CD-1-S1) and the CatchmentTools Project (2013-W-FS-14). Dr Fiachra O’Loughlin is funded by the Leverhulme Trust (RPG-409). The authors would like to thank B.Elssasser for providing data for this research. We are particularly grateful to the anonymous reviewers who considerably helped in improving the manuscript.


