A comparison of SWAT, HSPF and SHETRAN/GOPC for modelling phosphorus export from three catchments in Ireland

Ahmed Nasr\textsuperscript{a}, Michael Bruen\textsuperscript{a,}\textsuperscript{*}, Philip Jordan\textsuperscript{b}, Richard Moles\textsuperscript{c}, Gerard Kiely\textsuperscript{d}, Paul Byrne\textsuperscript{c}

\textsuperscript{a}Centre for Water Resources Research, University College Dublin, Earlsfort Terrace, Dublin 2, Ireland. Tel.: +353-1-7167378; Fax: +353-1-7167399; E-mail address: Michael.Bruen@ucd.ie

\textsuperscript{b}School of Environmental Sciences, University of Ulster, Coleraine, NI

\textsuperscript{c}Department of Chemical and Environmental Science, University of Limerick

\textsuperscript{d}Department of Civil and Environmental Engineering, University College Cork
Abstract
Recent extensive water quality surveys in Ireland revealed that diffuse phosphorus (P) pollution originating from agricultural land and transported by runoff and subsurface flows is the primary cause of the deterioration of surface water quality. P transport from land to water can be described by mathematical models that vary in modelling approach, complexity and scale (plot, field and catchment). Here, three mathematical models (SWAT, HSPF and SHETRAN/GOPC) of diffuse P pollution have been tested in three Irish catchments to explore their suitability in Irish conditions for future use in implementing the European Water Framework Directive. After calibrating the models, their daily flows and total phosphorus (TP) exports are compared and assessed. The HSPF model was the best at simulating the mean daily discharge while SWAT gave the best calibration results for daily TP loads. Annual TP exports for the three models and for two empirical models were compared with measured data. No single model is consistently better in estimating the annual TP export for all three catchments.

Keywords: Phosphorus; SWAT; HSPF; SHETRAN; GOPC

1. Introduction
The introduction of the Water Framework Directive in Europe (EEC, 2000) required Member States to review water quality problems in all their water bodies. In Ireland, riverine and lake eutrophication due to diffuse pollution has been identified as a major problem (Earle, 2003) and phosphorus (P) is the limiting nutrient controlling
eutrophication in inland waters (McGarrigle et al., 2002). Therefore an effective way to
tackle eutrophication is to control P inputs, both from point and diffuse sources.

Formerly, phosphorous from point sources was the major cause of serious pollution
incidents in most Irish rivers (McGarrigle et al., 2002). However, in response to the
Urban Wastewater Directive (EEC, 1991) many wastewater treatment plants in Ireland
were upgraded to include a tertiary process resulting in a large reduction in pollution
from point sources. Now, in many catchments most nutrients entering rivers are from
diffuse sources and therefore, this study modelled this influence, concentrating on P
transport in three Irish catchments. The catchments were chosen on the basis of
availability of the data required by the models and because they have different climate,
land use and soil types. The modelled variable is total phosphorus (TP) load because of
its direct relationship with impacts on receiving waters (Hilton et al., 2006).

According to the DPSIR conceptual framework (Drivers, Pressures, State, Impact and
Response) (Irvine et al., 2005) that will guide the selection of modelling techniques in
Ireland, it is likely that the most useful models will be of the physically-based or
mechanistic types. Three widely used, physically-based, models were selected to cover a
range of variation in (i) the complexity of their representation of the physical, chemical,
and bio-chemical processes involved in P mobilisation and transport, (ii) the degree of
complexity in spatial disaggregation of the catchment, and (iii) the normal simulation
time step. The models are: Soil Water and Analysis Tools (SWAT) (Arnold et al., 1998);
Hydrological Simulation Program – FORTRAN (HSPF) (Bicknell et al., 1997); and
Système Hydrologique Européen TRANsport (SHETRAN) (Ewen et al., 2000) coupled with the Grid Oriented Phosphorus Component (GOPC) (Nasr et al., 2005). The differences between the three models are discussed here first and then their application to the study catchments. Finally, their flow and TP load simulations are described and assessed and the TP loads compared with empirical models.

2. Differences between the SWAT, HSPF, and SHETRAN/GOPC models

2.1. Processes representation

The models chosen range from semi-empirical to fully physically-based in how they represent the relevant hydrological, chemical and bio-chemical processes transforming the P compounds both in the soil and during its transport by water. The SWAT model uses semi-empirical equations to represent most of these processes. HSPF models the catchment response by changes in water, sediment, and chemical amounts in a series of vertical storages. The fluxes between the various storages and exchanges with the river reaches are modelled with equations that have parameters determined by measurement and/or calibration. In contrast, SHETRAN/GOPC is an example of a fully physically-based model which relies wholly on relationships derived from physical and chemical laws. In order of increasing hydrological complexity, the models are ranked SWAT, HSPF, SHETRAN and in order of increasing complexity in representing P processes, HSPF, SWAT, GOPC.

2.2. Spatial representation
The three models have different procedures for representing spatial variation within the
catchment.

- SWAT divides the catchment into a number of sub-catchments, each of which has
  a number of Hydrologic Response Units (HRUs) with uniform land use and soil
types (without reference to their actual spatial position within the sub-catchment).

- HSPF divides the catchment on the basis of land use alone. Each land use can
  consist of pervious and impervious parts.

- In SHETRAN/GOPC, the catchment is divided into a horizontal orthogonal grid
  network and in the vertical direction by a column of horizontal layers at each grid
  square. Each grid element can have different land use and hydraulic properties.

  The channel system is represented on the boundaries of the grid squares.

2.3. Temporal resolution

SWAT operates only at a daily time step. Both HSPF and SHETRAN/GOPC can
simulate at any time step from one minute up to one day. In all cases, the input time
series should always be available at intervals equal to or less than the simulation time
step.

3. Study catchments

The Clarianna catchment (23 km$^2$) is located in County Tipperary in an area which is one
of the most intensively farmed catchments within the lower Shannon region. The Dripsey
catchment (15 km$^2$) is located near the town of Donoughmore in the south of Ireland and
ultimately drains into Inniscarra lake, a freshwater lake that in recent years has experienced signs of eutrophication (Scanlon et al., 2004).

The Oona Water catchment (96 km$^2$) is located in County Tyrone and ultimately drains into Lough Neagh which is a water source for Belfast.

4. Data

The model comparisons are based on simulations of daily time series of discharge and TP load at each catchment outlet. Data used in these simulations are summarised in Table 1. Each of the three models has been calibrated for the period from 1/12/2000 to 29/7/2001 in the Clarianna catchment, and from 1/1/2002 to 31/12/2002 in the Dripsey and Oona Water catchments. To allow HSPF and SHETRAN show their best performances, the available time step resolution of the input data has been also employed as a time step of simulation. As SWAT’s time step must be one day, daily input time series for it were derived from the available high resolution input data.

The only input of P was assumed to be direct application of fertiliser and animal slurries on the land. In the Clarianna and Dripsey catchments, the total annual P load applied on the soil was taken as 15 kg P ha$^{-1}$ in line with the National (Teagasc) recommendations (Teagasc, 1998). For the Oona Water catchment, P inputs of 18.9 million kg P to the soil were assumed based on a P balance in Northern Ireland (Jordon, 2003). In each catchment, the estimated value of the total annual P load has been distributed evenly over the twelve months of the year.
The observed TP concentrations at the outlet of the three catchments (Jordan et al., 2005) are summarised in Table (1) and daily TP loads were calculated from these and flow data. The modelled TP is the sum of the dissolved and sediment-attached P load estimates.

5. Model calibration and parameter estimation

5.1. Approach used in the calibration

Manual calibration has been used in the vast majority of reported applications of the three models (e.g. Jha et al., 2002; Wang et al., 1999; Bathurst, 1986) although some very limited attempts at automatic calibration have been made (e.g. Eckhardt and Arnold, 2001; Doherty and Johnston, 2003). Despite the considerable effort that has to be made to implement automatic calibration in these studies the results obtained were still within the range of the manually calibrated models.

To avoid the complexity and computational demands that arise when using automatic calibration (which would not allow these models to be used in many practical situations) a simple manual strategy was employed to calibrate the parameters of the three models to produce reasonable estimates for both discharges and TP loads. The three models were first calibrated to produce reasonable simulations of the discharge. Then the parameters of the best flow calibration were used without any further change during the P calibration, following the three step strategy proposed by Gupta et al. (2003; p. 11).

5.2. Initial estimates of model parameters (Level Zero estimates)
The main objective for the Level Zero estimates is to populate a default data set of parameters for the three models for each of the three test catchments. The SWAT and HSPF models can automatically produce default values for all model parameters by linking each soil and land use type with corresponding internal tables of default parameters. For the third model, the SHETRAN/GOPC combination, the initial values of the SHETRAN parameters were all taken from guidance given in its user’s manual.

5.3. Improving the parameter estimates (Level One estimates)

The effective parameters for flow simulation in HSPF were based on USEPA (2000). For its phosphorous modelling, the method used assumes a first order kinetics equation to represent each of the soil P processes. Each equation contains two parameters, the kinetics rate, which is calibrated manually, and the temperature coefficient, which is maintained at a specified value (USEPA, 2000). For the SWAT model the user’s manual guidance was followed to improve the parameter values for flow and nutrient simulation.

The SHETRAN model was developed to be used without calibration due to the physical nature of its parameters which are intended to be obtained from direct measurements. However in most previous applications of this model (e.g. Anderton et al., 2002), some parameters have been calibrated and this has been done here.

To illustrate the complexity of each model in terms of parameter calibration, Tables (2) and (3) give the number of the effective parameters and the methods by which they have been used in the water and P simulations, and also cite the source for the methods. We
refer the reader to the user’s manual of each model for definitions and descriptions of the
parameters and these are not repeated here. However, the number of parameters which
can be adjusted for the three models is reported in Tables (2) and (3).

5.4. Adjustment of the parameters (Level Two estimates)

Manual improvement of effective parameters for complex models is difficult to carry out
in a reliable and consistent manner due to (i) large numbers of parameters, (ii)
equifinality, (iii) parameter sensitivity, and (iv) uncertainty (Gupta et al., 2003). In this
study, however, the manual calibration is to provide results corresponding to practical
situations when these models are used by typical users with only a general knowledge of
sophisticated calibration techniques. Thus the models are compared in terms of their
likely performances in reality rather than their potential best performances if unlimited
calibration resources were available.

A systematic approach to manual calibration has been followed in this study where for
each model and in each catchment, one parameter was changed at a time and the resulting
shapes of the hydrograph in the flow calibration or load graphs in the P calibration were
visually compared with the observed. In addition, the Nash-Sutcliffe index (Nash and
Sutcliffe, 1970) was also calculated for each model run. In the flow calibration, first the
focus was on visually matching the peaks, then on matching the flow recessions and
finally on the low flow values. This requires sequential adjustment to the particular
parameters in the models that influence each of the three components of the hydrograph.
The P calibration followed a similar procedure where the focus was first on the high
values, then on the low values, and finally on the overall shape of the simulated time series.

6. Comparison criteria

The three models are compared both on their daily and annual results. Two criteria are used to assess the models in simulating the daily discharges and TP loads. Firstly, for each catchment, the flow hydrograph was plotted together with the rainfall hyetograph so that the flow simulation and its consistency with rainfall can be observed and the daily TP results were superimposed on the graph. This allows a direct visual appreciation of the influence of the hydrological modelling on the transport of P. Importantly, it also reveals many systematic aspects of the model performance such as tendencies to over or under estimate, which are seasonal or related to flow or rainfall, and tendencies to match high peaks but not low ones. These tendencies would be difficult to detect with a single numerical index.

The second criterion is based on two numerical measures: the Nash-Sutcliffe index ($R^2$); and the fraction of the mean of the squares of the errors due to bias ($B\%\text{MSSE}$). These are calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (x_i - \hat{x}_i)^2}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$$ (1)
where \( x_i \) is the observed value, \( \hat{x}_i \) is the value estimated by a model, \( \bar{x} \) is the mean of the observed values, \( \bar{\hat{x}} \) is the mean of the estimated values, and \( n \) is the number of data values.

The \( R^2 \) index describes the ability of the model to explain the variability in the data, while the B\%MSSE index describes the model’s ability to match the central tendency (mean) of the data. A good model is one that has a good visual match to the observed data, with a value of \( R^2 \) close to one, and a small value of B\%MSSE.

7. Comparison with simple empirical models

The three physically-based models are compared with two empirical models specifically developed to estimate annual TP export. The first, (DM) is derived from an equation developed by Daly et al. (2006) specifically for use in Irish conditions and the second model (ECM) is an export coefficient model (Johnes, 1996) used in the UK.

8. Results

8.1. Discharge performance

The hydrographs of observed and estimated discharge (Figs 1.a) (2.a) (3.a) show that, in general, none of the models is able to replicate the entire shape of the hydrographs
throughout the simulation period. However, HSPF is the best at matching the discharge hydrographs and SWAT performed better than SHETRAN. The noticeable weakness in SHETRAN is its failure to adequately model the flow peaks and recessions. Most of its estimated peaks are either very much higher or lower than the observations and also the estimated flow recession is not as flat as the observed.

The numerical criteria for the flow simulations from each model are compared in Table (4). The best results for $R^2$ are 0.95, 0.74, and 0.91 and for the HSPF model in the Clarianna, Dripsey, and Oona respectively. The $R^2$ for SWAT (0.91) is better than for SHETRAN (0.74) in the Dripsey. However, SHETRAN has an $R^2$ of 0.8 which is better than the 0.73 for SWAT in the Oona. The B%MSSE values for all three models in Clarianna are all low and this means that the bias in estimating the mean of the observed flow was not the major source of error. In the Dripsey and Oona, the B%MSSE value for HSPF is high compared to SWAT and SHETRAN with the latter giving the best results for B%MSSE for these two catchments.

8.2. Daily TP results

Figs ((1.b), (2.b), and (3.b)) show the TP results. SWAT performed quite well in all three catchments. The GOPC TP load performance is second, despite the problems with the flow estimates. It is quite surprising that HSPF was the worst for TP load in the three catchments although it was best for discharges. The main reason for this could be its smaller scope for calibration due to the fewer number of P parameters, as show in Table
(3), and this indicates the limited capability of the first order kinetics equation used in the P simulation by HSPF.

The numerical criteria for TP loads are summarised in Table (4). In the Clarianna, the value of $R^2$ for SWAT (0.59) is the best while the corresponding value for GOPC (0.40) is the worst of the three models. HSPF has the highest B%MSSE while the other two models have low values, indicating that their bias is less significant. In the Dripsey, the best $R^2$ value of 0.51 is obtained from the GOPC model while the HSPF model value was the worst (0.22). All three models have low B%MSSE values indicating an overall acceptable bias performance in this catchment. In the Oona, SWAT has the best $R^2$ (0.56) while its B%MSSE value is second best to the HSPF value. The GOPC has the worst $R^2$ (0.23) and B%MSSE values.

8.3. Annual TP export results

Table (5) shows the observed and modelled values of annual TP export from each catchment. In the Clarianna, SWAT and GOPC are the best of the five models. Both empirical models greatly overestimated the TP export while HSPF underestimates. In the Dripsey, all models except DM are comparable with the observations, with HSPF being the best. In the Oona, SWAT is the only model that overestimated the TP export although the result is still acceptable. HSPF and DM also give acceptable results, although lower than the observed, whereas the GOPC and EM significantly underestimate.

The results from the three physically-based models are better than the results from the two empirical models. This is not surprising because the empirical models are fitted to a
larger number of different catchments and are an average of the responses over a larger area and have much less catchment-specific information.

9. Conclusions

In the three catchments, the HSPF model was best in simulating the mean daily discharges. Discharge results from SWAT and SHETRAN were acceptable despite occasional deficiencies. Nevertheless, the best simulation for daily TP loads in the study catchments was by SWAT. The TP performance of the SHETRAN/GOPC combination was good although hampered by its discharge simulation. In terms of TP export, no single model was best for all three catchments, however the three physically-based models gave annual TP export estimates closer to the measurements than the two empirical models. In the short term, we recommend using SWAT for TP load estimation.

Acknowledgements

This project was supported by the National Development Programme (NDP) through the RTDI programme and co-funded by the Irish EPA and Teagasc (2000-LS-2.2.2-M2). In the Oona Water, we acknowledge the use of hydrology data from the Rivers Agency (Department of Agriculture and Rural Development, NI) and infrastructure from the NERC funded CHASM project (NER/H/S/1999/00164).

References


observations in a plantation of Corsican pine. Agriculture Meteorology 9, 367-384.


<table>
<thead>
<tr>
<th>Type of data</th>
<th>Clarianna</th>
<th>Dripsey</th>
<th>Oona Water</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Elevation Model</td>
<td>Resolution 20 m X 20 m</td>
<td>20 m X 20 m</td>
<td>50 m X 50 m</td>
</tr>
<tr>
<td>Land use</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil</td>
<td>Source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weather data (rainfall,</td>
<td>Method Observed</td>
<td>Observed</td>
<td>Observed</td>
</tr>
<tr>
<td>temperature, solar</td>
<td>Time step 15 minutes</td>
<td>30 minutes</td>
<td>5 minutes</td>
</tr>
<tr>
<td>radiation, relative humidity,</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wind speed)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flow discharge</td>
<td>Method Observed</td>
<td>Observed</td>
<td>Observed</td>
</tr>
<tr>
<td>Time step 15 minutes</td>
<td>15 minutes</td>
<td>15 minutes</td>
<td>1 hour</td>
</tr>
<tr>
<td>Soil phosphorus application</td>
<td>Method Estimated</td>
<td>Estimated</td>
<td>Estimated</td>
</tr>
<tr>
<td>regime</td>
<td>Frequency</td>
<td>Annual amount</td>
<td>Annual amount</td>
</tr>
<tr>
<td>Total phosphorus at catchment</td>
<td>Method Observed</td>
<td>Observed</td>
<td>Observed</td>
</tr>
<tr>
<td>outlet</td>
<td>Time step Concentrations of</td>
<td>flow proportional</td>
<td>Concentrations of</td>
</tr>
<tr>
<td></td>
<td></td>
<td>flow proportional</td>
<td>discrete samples</td>
</tr>
<tr>
<td></td>
<td></td>
<td>daily loads</td>
<td></td>
</tr>
</tbody>
</table>
Table (2) Summary of the methods used and parameters required to simulate processes of the water component of SWAT, HSPF, and SHETRAN models

<table>
<thead>
<tr>
<th>Process</th>
<th>SWAT Method</th>
<th>HSPF Method</th>
<th>SHETRAN Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interception</td>
<td>Water balance</td>
<td>Water balance</td>
<td>Rutter et al. (1971/72)</td>
</tr>
<tr>
<td>Potential evapotranspiration (E_p)</td>
<td>Monteith (1965)/ Priestley and Taylor (1972)/ Hargreaves et al. (1985)</td>
<td>User defined E_p</td>
<td>User defined E_p/ Monteith (1965)</td>
</tr>
<tr>
<td>Actual evaporation (E_a)</td>
<td>Ritchie (1972)</td>
<td>Accounting procedure</td>
<td>Feddes et al. (1976)</td>
</tr>
<tr>
<td>Runoff</td>
<td>SCS (1972)</td>
<td>Empirical relations</td>
<td>Approximation of St. Venant equation</td>
</tr>
<tr>
<td>Infiltration</td>
<td>Water balance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inter/Return flow</td>
<td>Kinematic model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseflow</td>
<td>Linear reservoir model</td>
<td>Linear reservoir</td>
<td></td>
</tr>
<tr>
<td>Percolation to groundwater</td>
<td>Water balance</td>
<td>Empirical relation</td>
<td></td>
</tr>
<tr>
<td>River flow routing</td>
<td>Variable storage/ Muskingum</td>
<td>Water balance</td>
<td>Approximation of St. Venant equation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>18</td>
<td>15</td>
<td>15</td>
</tr>
</tbody>
</table>


Table (3) Summary of the methods used and parameters required to simulate processes of the phosphorus component of SWAT, HSPF, and GOPC models

<table>
<thead>
<tr>
<th>Process</th>
<th>SWAT</th>
<th>HSPF</th>
<th>GOPC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Method</td>
<td>No. of parameters</td>
<td>Method</td>
</tr>
<tr>
<td>Soil P processes</td>
<td>Mass balance</td>
<td>6</td>
<td>First order kinetics</td>
</tr>
<tr>
<td>(adsorption/desorption, mineralisation/immobilisation, plant uptake)</td>
<td></td>
<td></td>
<td>equation</td>
</tr>
<tr>
<td>Leaching of P</td>
<td>Empirical relation</td>
<td>1</td>
<td>Mass Balance</td>
</tr>
<tr>
<td>Dissolution of P</td>
<td>Empirical relation</td>
<td>1</td>
<td>Mass Balance</td>
</tr>
<tr>
<td>in runoff water</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overland transport of dissolved P</td>
<td>Linear reservoir</td>
<td>-</td>
<td>Not simulated by this</td>
</tr>
<tr>
<td>dissolved P</td>
<td>model</td>
<td></td>
<td>model</td>
</tr>
<tr>
<td>Subsurface transport of dissolved P</td>
<td>User defined</td>
<td>1</td>
<td>Mass Balance</td>
</tr>
<tr>
<td>dissolved P</td>
<td>concentration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transport of attached or particulate P</td>
<td>Empirical relation</td>
<td>1</td>
<td>Mass balance</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total = 10</td>
<td></td>
<td>Total = 5</td>
</tr>
</tbody>
</table>
Table (4) Values of $R^2$ and B\%MSSE for the flow and phosphorus in the three study catchments

<table>
<thead>
<tr>
<th>Flow</th>
<th>Observed</th>
<th>Clarianna Estimated</th>
<th>Observed</th>
<th>Dripsey Estimated</th>
<th>Observed</th>
<th>Oona Water Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SWAT</td>
<td>HSPF</td>
<td>SHETRAN</td>
<td></td>
<td>SWAT</td>
</tr>
<tr>
<td>No. of days</td>
<td></td>
<td>241</td>
<td>241</td>
<td>241</td>
<td>365</td>
<td>365</td>
</tr>
<tr>
<td>$R^2$</td>
<td></td>
<td>-</td>
<td>0.91</td>
<td>0.95</td>
<td>0.75</td>
<td>-</td>
</tr>
<tr>
<td>B%MSSE</td>
<td></td>
<td>-</td>
<td>0.01</td>
<td>0.03</td>
<td>0.07</td>
<td>-</td>
</tr>
</tbody>
</table>

Table (5) Values of total phosphorus export (kg P/annum) in the three study catchments

<table>
<thead>
<tr>
<th>Catchment</th>
<th>Observed</th>
<th>Estimates from each model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SWAT</td>
</tr>
<tr>
<td>Clarianna</td>
<td>289</td>
<td>231</td>
</tr>
<tr>
<td>Dripsey</td>
<td>1719</td>
<td>1371</td>
</tr>
<tr>
<td>Oona</td>
<td>27496</td>
<td>33285</td>
</tr>
</tbody>
</table>
Figure (1) Clarianna catchment: Results of the three models (a) Flow, and (b) Total phosphorus load
Figure (2) Dripsey catchment: Results of the three models (a) Flow, and (b) Total phosphorus load
Figure (3) Oona Water catchment: Results of the three models (a) Flow, and (b) Total phosphorus load.