Understanding Freshwater Quality Problems in a Changing World
Proceedings of H04, IAHS-IAPSO-IASPEI Assembly, Gothenburg, Sweden, July 2013 (IAHS Publ. 361, 2013).

Parameterizing dynamic water quality models in ungauged basins: issues and solutions

EVA MOCKLER & MICHAEL BRUEN
UCD Dooge Centre for Water Resources Research, School of Civil, Structural and Environmental Engineering, University College Dublin, Ireland
eva.mockler@ucd.ie

Abstract The redundancy and ill-conditioned nature of model identification and parameter estimation make it difficult to identify a hydrological model that can correctly split the flow dynamics and associated contaminant transport and transformation between each of its conceptual pathways. The Pathways project, funded by the Irish EPA, combines insights from conceptual catchment modelling and detailed fieldwork investigations to inform an integrated water management tool. Within this project, a user-friendly GIS application is being developed for environmental managers interested in water quality modelling using the Pathways Computational Engine (PACE) model. This variable structure water quality model can be used to investigate hydrological and contaminant processes at sub-catchment scale and to incorporate expert knowledge relating to flow pathways and contaminant transport along these pathways. In this paper, parameter identification issues related to the hydrological model are discussed with an investigation of groundwater parameters for catchments underlain by locally important aquifers.

Key words hydrological processes; catchment modelling; flow path contributions; parameter identification

INTRODUCTION

Predictions of water quality in ungauged basins

Making predictions in ungauged or poorly gauged basins is a fundamental challenge for hydrologists, and has been the driver for international collaborations including the IAHS Predictions in Ungauged Basins (PUB) decade and the Model Parameter Estimation Experiment (MOPEX) initiative (Duan *et al*., 2006). A cornerstone of these projects was investigating parameter transferability and enhanced techniques for the *a priori* estimation of parameters. Parameter estimation in hydrological catchment models has been widely studied for decades, and yet values are still often selected from best fit to streamflow data on a site specific basis. Calibrating to catchment outlet time-series alone can produce a large set of equally “good” parameter sets and results in a wide variety of different flow path contributions to the outlet flows. This may be adequate if water quantity alone is the quantity of interest. However, for water quality modelling at catchment scale, the hydrological pathways influence the mobilisation, transport and attenuation of most contaminants and so getting accurate flow path contributions is vital.

Regionalisation methods typically involve parameter transfer to ungauged basins through calibration of a set of gauged catchments and relating optimised parameter values to catchment characteristics (Vogel, 2005). Over-parameterized models can result in difficulty in identifying clear relationships, due to multiple optimum parameter sets (Götzinger & Bárdossy, 2007). For predictions in ungauged basins, a robust model structure and parameter set must be preferred over the “best” performing model in the study site. In this study, the focus is on the hydrological predictions with a view to reducing the number of model parameters while maintaining a realistic partitioning of flow pathways contributing to streamflow, as this is required for the contaminant modelling.

There are many existing models for investigating fluxes of nutrients and contaminants in catchments, including SWAT (Soil and Water Assessment Tool; Arnold & Fohrer, 2005), OpenFLUID (Fabre *et al*., 2010), HYPE (HYdrological Predictions for the Environment; Lindstrom *et al*., 2010), INCA-P (Integrated Nitrogen in Catchments; Dean *et al*., 2009) and MONERIS (MOdelling Nutrient Emissions in RIver Systems; Behrendt *et al*., 2002). Due to the complexities in modelling water quality species, these models have a relatively large number of parameters relative to water quantity models, and have greater uncertainties if applied to ungauged basins.

Water quality modelling in Irish catchments

The Irish Environmental Protection Agency (EPA) initiated the Pathways Project in 2007, to develop the Pathways Catchment Management Toolbox (CMT) for environmental managers interested in water quality modelling and its environmental consequences. The CMT is a GIS-based application linked with a flexible hydrological model for investigating the movement of pollution through our river basins and identifying the potential sources of contaminants. The Pathways Computational Engine (PACE) model is a semi-distributed flexible conceptual model suitable for catchment-scale modelling that is being developed in parallel. The CMT acts as a user-friendly interface to the PACE model to aid environmental managers to identify critical source areas for selected contaminants and evaluate alternative strategies for land use and their impacts on aquatic ecosystems and target areas for enforcement of regulations. In Ireland, this is complicated by the extreme heterogeneity in geology, soils and land use (Archbold *et al*., 2010).

The prevalent issues for Irish rivers are eutrophication and organic pollution due to nutrients and oxygenation conditions. Indicators for this organic enrichment are phosphate, nitrate, total ammonia and BOD. Currently, 10% of rivers are classified under the EU Water Framework Directive as poor status, or moderately polluted, with the cause generally attributed to organic pollution (McGarrigle *et al*., 2010). It is therefore vital that the PACE model can simulate the proportion of flow through surface and sub-surface hydrological flow paths and the transport and attenuation of nutrients associated with these.

CONSTRAINING HYDROLOGICAL PARAMETER SETS

By querying available hydrological data, modellers can constrain the plausible conceptual model and so reduce the preferred parameter space. Diagnostic tests for model structures discussed by McMillan *et al*. (2011), illustrate how precipitation, soil moisture and flow data can be used to determine a parsimonious model structure. For example, recession analysis was used to investigate the required number of linear or non-linear stores to represent seasonal streamflow dynamics. Such data interpretation can be used to refine conceptual models of catchments in conjunction with other resources, such as tracer studies when available. The impacts of uncertainties on model predictions can be reduced by limiting the number of parameters to be calibrated (Huang & Liang, 2006).

Global sensitivity analysis is the study of the relationships between information flowing into and out of a computational model (Saltelli *et al*., 2000), and can be used to reduce the dimensions of useful hydrological parameter sets. The variation of model output can be apportioned to different sources of model input variation including input data, variables and model parameters; however, the results can be sensitive to the modelling time-step (O'Loughlin *et al*., 2012). Determining and fixing non-sensitive variables in the model can ensure the model is parsimonious and allows relationships between the remaining parameters and catchment characteristics to be more clearly identified.

Uniform random sampling (URS) is used as a tool for complete exploration of the parameter space, to identify the presence of multiple local optima and to graphically investigate parameter uncertainty (Wagener & Wheater, 2006). The URS procedure samples a large number of parameter sets from ranges constrained by reasonable upper and lower bounds to explore the feasible parameter space of model structures.

In this study, parameters of a conceptual rainfall–runoff model are investigated using URS to assess the impacts of fixing non-sensitive parameters, and to determine the suitability of the model structure for water quality modelling in Irish catchments.

Conceptual rainfall–runoff model

A number of different conceptual catchment models were investigated as potential candidates for inclusion in the computational engine of the Catchment Management Tool. One of these models structures is a lumped conceptual rainfall–runoff model with a structure somewhat similar to that of the NAM model (Nielsen & Hansen, 1973), with two storage reservoirs for soil moisture

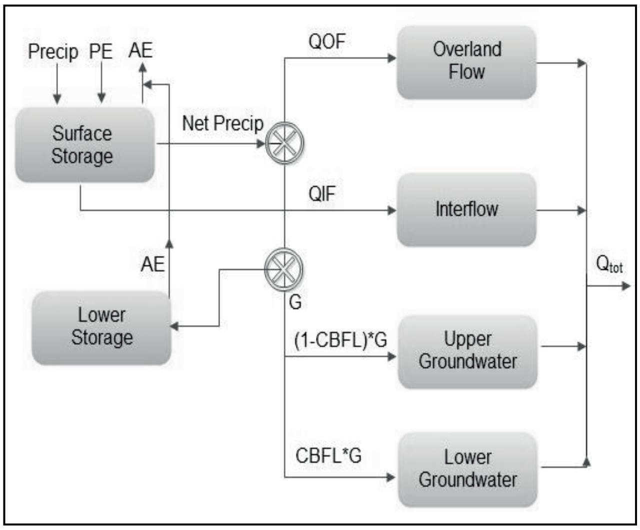

Fig. 1 Model schematic showing conceptual flow paths contributing to stream flow (Q_{tot}).

Table 1 Model parameters.

Parameter	Description	Range
UMAX	Upper layer storage capacity (mm)	1.0–50.0
LMAX	Lower layer storage capacity (mm)	20.0–500.0
CQOF	Runoff coefficient in overland flow equation	0.01–1.0
CLOF	Threshold coefficient in overland flow equation	0.0–0.8
CQIF	Drainage coefficient in interflow equation (1/hour)	0.00001–0.05
CLG	Threshold coefficient in recharge equation	0.0–0.95
CBFL	Part of recharge going to lower groundwater	0.0–0.95
SUPERCK	Time constant of linear reservoir for routing OF and IF (hours)	6.0–100.0
CKBFU	Time constant of upper groundwater storage (hours)	1000–4000
CKBFL	Time constant of lower groundwater storage (hours)	2000–8000

accounting and reservoirs to represent four hydrological pathways, as illustrated in Fig. 1. The 10 parameters, including two storage limits, four temporal parameters and four coefficients are listed in Table 1. The snow component is not relevant to this study and therefore not included. The relative volume in the lower storage is used in the calculation of the proportion of flow contributing to each of the flow paths, and is only depleted by evapotranspiration.

STUDY CATCHMENTS

Irish catchments with locally productive aquifers

Ireland has an area of approximately 84 000 km^2, with the majority of the country in the gently undulating lowlands of the central plain with elevations generally less than 150 m above sea level. The maritime climate is shaped by the westerly atmospheric circulation of the middle latitudes and the warm North Atlantic Drift, with mild, moist winters and cool cloudy summers, and westerly to southwesterly prevailing winds.

Surface water catchments in Ireland are often underlain by mixed aquifers with heterogeneous groundwater flow regimes, with flow in most aquifers through faults, joints and fractures. Poorly productive aquifers have low transmissivities and a greater degree of fluctuation in the water table than productive aquifers, with flow path lengths of generally less than about 300 m (RPS, 2008). Study catchments were selected, as illustrated in Fig. 2, with predominantly locally productive bedrock aquifers that are moderately productive only in local zones. The catchments have a broad range of annual average rainfall (AAR) and mean streamflow (Table 2). Each catchment was simulated using the above model structure at a daily time step for a 16 year period (1990–2005).

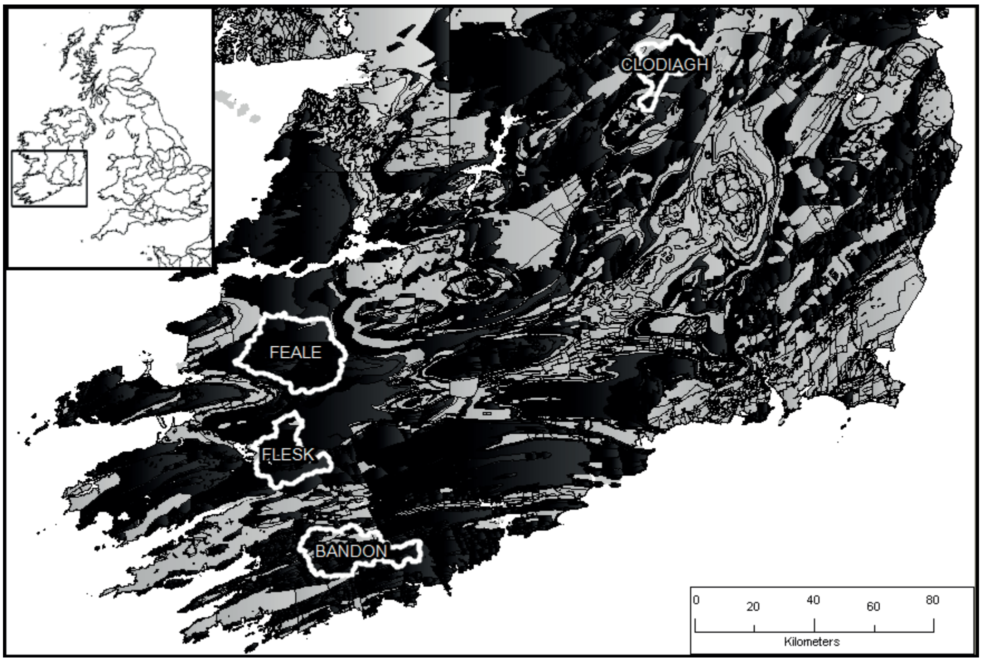

Fig. 2 Irish locally productive bedrock aquifers shown in dark shading (all other categories in light shading), with the Clodiagh, Feale, Flesk and Bandon catchment boundaries.

Table 2 Catchment characteristics including regional karstified diffuse (Rkd), locally important aquifer (Ll), poorly productive aquifer (Pl), annual average rainfall (AAR), mean streamflow (Mean Q) and the ratio of the 5th and 95th stream flow percentile (Q95:Q5).

Catchment	Area (km^2)	Rkd (%)	Ll (%)	Pl (%)	AAR (mm/year)	Mean Q (m^3/s)	Q95:Q5
Bandon	424	0.0	77.2	22.8	1576	14.9	29.5
Clodiagh	254	12.9	82.6	0.5	904	3.9	18.9
Feale	647	1.3	98.8	0.0	1532	22.0	49.9
Flesk	329	7.1	91.2	0.0	1897	14.4	21.3

METHODOLOGY

Monte Carlo simulations

A broad parameter space, as detailed in Table 1, was sampled 50 000 times and evaluated for mean residual criteria and the Nash-Sutcliffe Efficiency (NSE), defined as:

$$NSE = 1 - \frac{\sum_{t=1}^{n}\left(Q_{o,t} - Q_{m,t}\right)^2}{\sum_{t=1}^{n}\left(Q_{o,t} - \overline{Q}_o\right)^2} \tag{1}$$

where $Q_{o,t}$ is the observed flow for time-step t, $Q_{m,t}$ is the modelled flow at time-step t, and n is the length of the time series.

Parameters were then graphically identified with values that did not taper to a narrow range with increasing NSE performance (Fig. 3), illustrating parameter interactions most notably in the groundwater flow paths. In order to reduce the parameter set, four of these were given a fixed value set as the average of the median values of the top 500 performing parameter sets, as evaluated by NSE in the four catchments. The chosen parameters CLG (0.39), CBFL (0.45), CKBFU (2326) and CKBFL (4850) were fixed and Monte Carlo simulations were re-run and results were re-evaluated and compared with the original simulations.

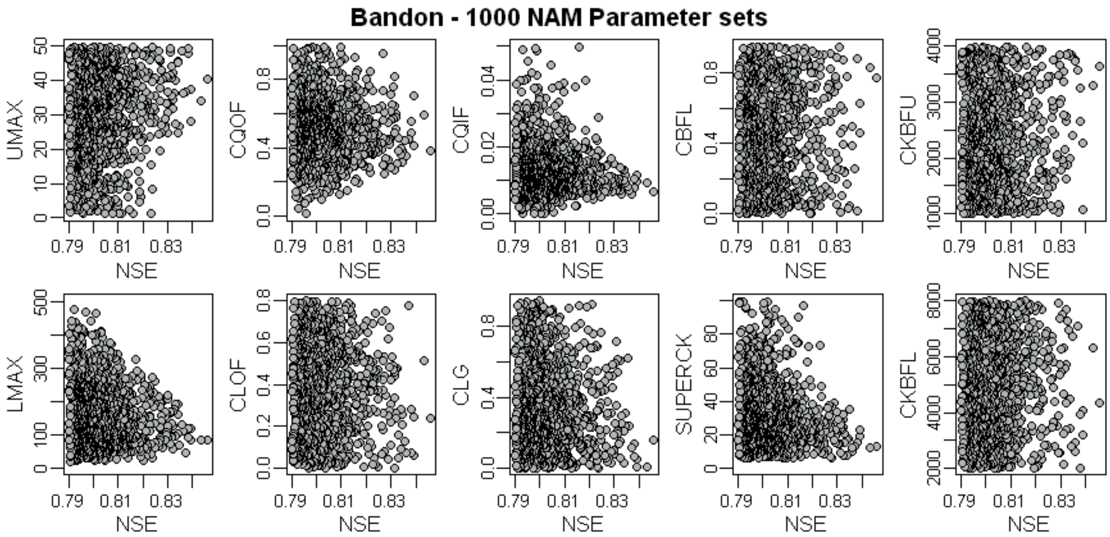

Fig. 3 Parameter sets for 1000 simulations plotted against NSE for Bandon catchment.

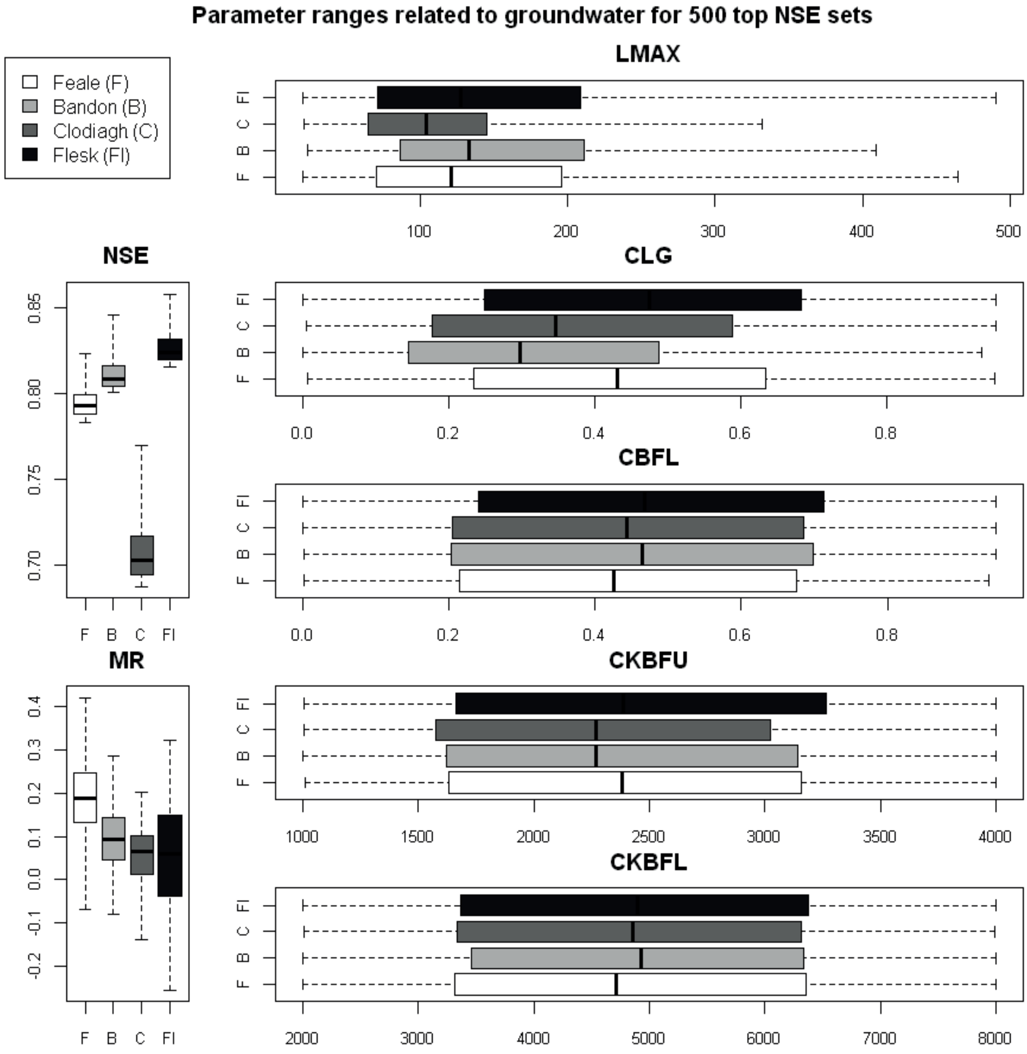

Fig. 4 Parameter ranges for LMAX, CLG, CBFL, CKBFU and CKBFL for Feale (F), Bandon (B), Clodiagh (C) and Flesk (Fl) catchments, with NSE and mean residual (MR) values for the top 500 parameter sets.

RESULTS

Parameter values

Figure 4 illustrates that the four catchments' parameters related to groundwater have similar distributions, particularly CBFL, CKBFU and CKBFL, which relate to the timing of flows and reflect the similar aquifer properties of the study catchments. These parameters, along with CLG, the threshold coefficient in the recharge equation, were chosen to be replaced with the average median as the fixed value.

Parameter sets for the top 500 NSE performing model simulations for the Bandon catchment are shown in Fig. 5. Some parameters, including LMAX, CQOF, SUPERCK and CKIF, show a trend towards an optimum value with increasing values of NSE, unlike others, including CBFL, CKBFU and CKBFL, which have a broad scatter (Fig. 3). These density plots show a slight reduction in parameter value range from before fixing the groundwater parameters (right columns) to after fixing values (left columns), most notably for SUPERCK and CQIF. The NSE values for the original simulations are plotted against the NSE results after fixing the four groundwater parameters for each catchment (Fig. 6). General performance improvements in the simulations are attributed to the reduced number of parameters searching the available range. The results are satisfactory, except for a reduction in the best performing parameter sets in the Clodiagh catchment.

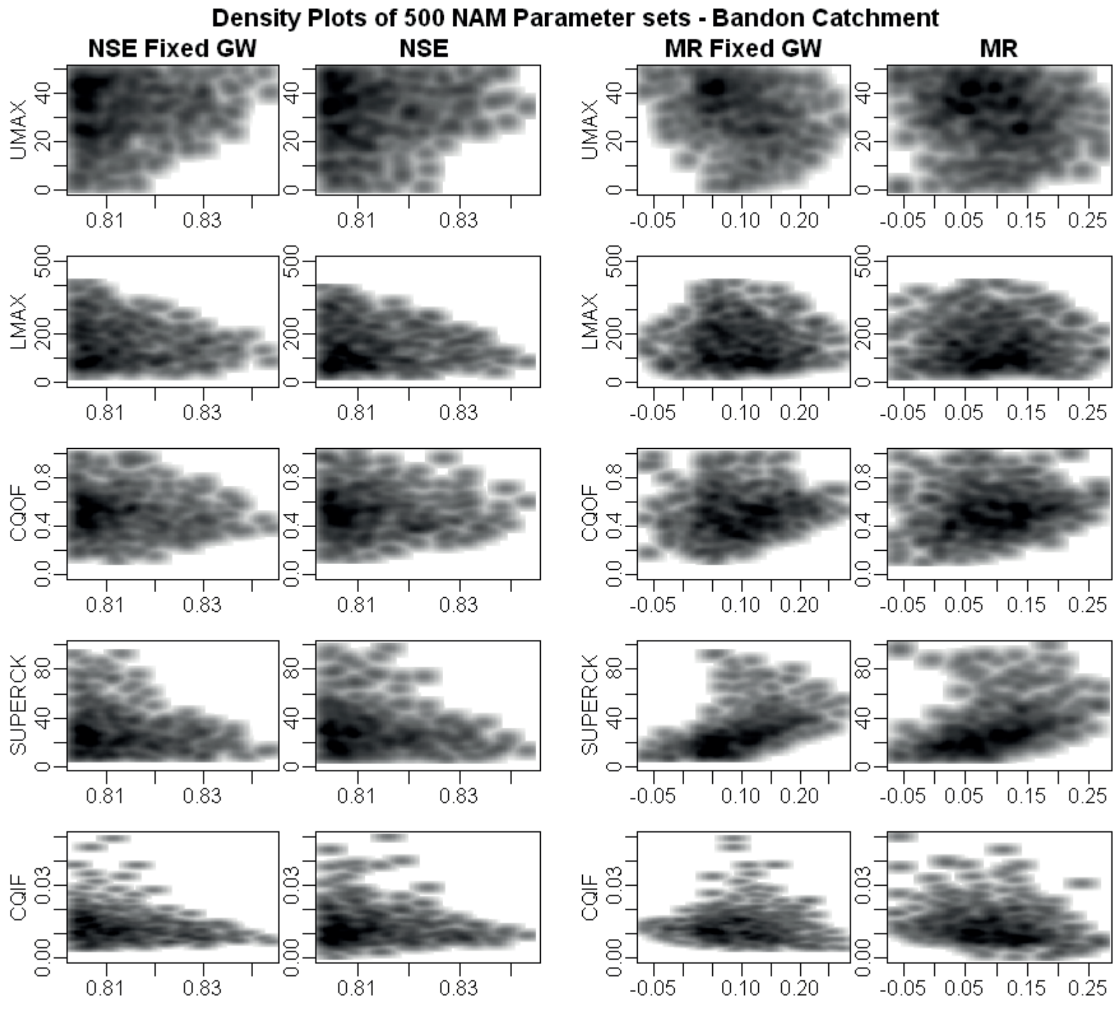

Fig. 5 Parameter values for UMAX, LMAX, CQOF, SUPERCK and CQIF after fixing groundwater parameters (left columns) and before (right columns), plotted against NSE and mean residual (MR) values for top 500 parameter sets for Bandon catchment.

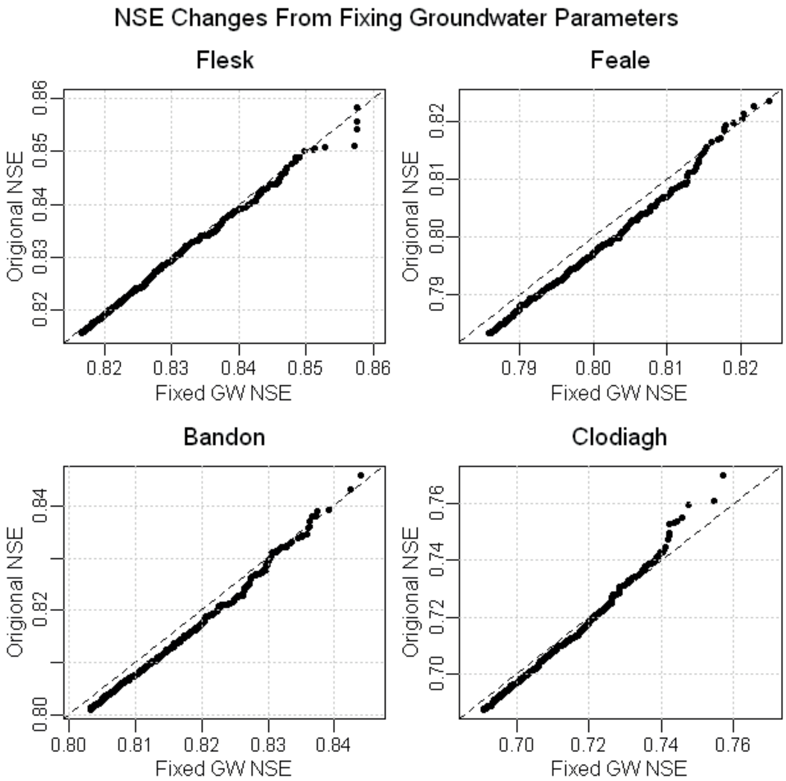

Fig. 6 NSE values for top 500 parameter sets of original simulations *vs* simulations with four fixed parameters.

DISCUSSION

Process understanding

The model structure discussed here represents a conceptual understanding of hydrological processes in these locally productive aquifers, with four hydrological flow paths. In order to reduce the number of parameters while maintaining the structure of the model for the purpose of water quality modelling, groundwater parameters are given a fixed value related to aquifer type. In this manner, the conceptual model of these catchments is maintained while reducing the parameter set, so facilitating the water quality equations which are driven by the hydrological flow paths.

While some parameters show a clear trend towards converging to a value as the NSE value increases, others have values spanning the permissible range within the top 1% of simulations that produce "good" model performance (Fig. 3). Figure 5 illustrates a slight reduction in the parameter value ranges for the Bandon catchment after the number of parameters is reduced from 10 to 6. The broad scatters of some parameters, prevalent in those related to groundwater, illustrate the equifinality problem (Beven & Binley, 1992), and suggest that the 10 parameter model structure may not be suitable for predictions in ungauged basins.

Moving forward

The large variations in parameter values with the top 1% of simulations indicate that, even while performing well, the model structure facilitates excessive parameter interactions that may impact negatively on the success of a regionalisation study. Further conceptual model structures are being investigated and as research progresses and further field data are collected, this conceptual model may be revised, altering the model structure and parameters in the search for a more parsimonious model. Similar parameter identification techniques for model structures associated with other catchment types will be required to facilitate the parameterization of the PACE model within the Pathways CMT.

CONCLUSIONS

Parsimonious models with the ability to capture the dominant hydrological processes are required for successful water quality modelling at catchment scale (Dooge, 2005). If intended for widespread use, such models must be transferrable to ungauged basins and a key requirement is the reduction of the dimension of the set of parameters to be identified. This paper shows how Monte Carlo simulation can assist with this and demonstrates the results for a number of catchments in Ireland. Future possibilities include using new field data, particularly high resolution time series of the chemical composition of hydrological flow paths and river reaches, to give greater insight to catchment processes, which can be used to inform model structure and further refine parameter sets. In this way, conceptual parsimonious models can enhance our understanding of the current state of the system and allow modellers to understand the effect of future changes due to land use and climate change.

Acknowledgements Support for this work was provided by the Pathways Project, funded by the Irish Environmental Protection Agency (2007-W-CD-1-S1).

REFERENCES

Archbold, M., Bruen, M., Deakin, J., Doody, D., Flynn, R., Kelly-Quinn, M., Misstear, B. & Ofterdinger, U. (2010) *Contaminant Movement and Attenuation along Pathways from the Land Surface to Aquatic Receptors – A Review*. EPA. Available from: http://www.epa.ie/downloads/pubs/research/water/name,30471,en.html.

Arnold, J.G. & Fohrer, N. (2005) SWAT2000: current capabilities and research opportunities in applied watershed modelling. *Hydrol. Processes* 19(3), 563–572.

Behrendt, H., Kornmilch, M., Opitz, D., Schmoll, O. & Scholz, G. (2002) Estimation of the nutrient inputs into river systems - experiences from German rivers. *Regional Environmental Change* 3(1-3), 107–117.

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

Bruen, M. (2009) *Hydrology and the Water Framework Directive in Ireland*. Biology & Environment: Proceedings of the Royal Irish Academy, The Royal Irish Academy.

Dean, S., Freer, J., Beven, K., Wade, A.J. & Butterfield, D. (2009) Uncertainty assessment of a process-based integrated catchment model of phosphorus. *Stochastic Environmental Research and Risk Assessment* 23(7), 991–1010.

Dooge, J. (2005) Bringing it all together. *Hydrology and Earth System Sciences Discussions* 9(1/2), 3–14.

Duan, Q., Schaake, J., Andreassian, V., Franks, S., Goteti, G., Gupta, H., Gusev, Y., Habets, F., Hall, A. & Hay, L. (2006) Model Parameter Estimation Experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *J. Hydrol.* 320(1), 3–17.

Fabre, J., Louchart, X., Colin, F., Dagès, C., Moussa, R., Rabotin, M., Raclot, D., Lagacherie, P. & Voltz, M. (2010) OpenFluid: a software environment for modelling fluxes in landscapes. *LANDMOD2010. Montpellier, France*.

Gotzinger, J. & Bardossy, A. (2007) Comparison of four regionalisation methods for a distributed hydrological model. *J. Hydrol.* 333(2-4), 374–384.

Huang, M.Y. & Liang, X. (2006) On the assessment of the impact of reducing parameters and identification of parameter uncertainties for a hydrologic model with applications to ungauged basins. *J. Hydrol.* 320(1-2), 37–61.

Lindstrom, G., Pers, C., Rosberg, J., Stromqvist, J. & Arheimer, B. (2010) Development and testing of the HYPE (Hydrological Predictions for the Environment) water quality model for different spatial scales. *Hydrol. Research* 41(3-4), 295–319.

McGarrigle, M., Lucey, L., & O'Cinneide, M. (2010) Water Quality in Ireland 2007–2009. *Environmental Protection Agency, Ireland*, ISBN: 978-1-84095-387-9.

McMillan, H.K., Clark, M.P., Bowden, W.B., Duncan, M. & Woods, R.A. (2011) Hydrological field data from a modeller's perspective: Part 1. Diagnostic tests for model structure. *Hydrol. Processes* 25(4), 511–522.

Moussa, R., Voltz, M. & Andrieux, P. (2002) Effects of the spatial organization of agricultural management on the hydrological behaviour of a farmed catchment during flood events. *Hydrol. Processes* 16(2), 393–412.

Nielsen, S. & Hansen, E. (1973) Numerical simulation of the rainfall-runoff process on a daily basis. *Nordic Hydrol.* 4, 171–190.

O'Loughlin, F., Bruen, M. & Wagener, T. (2013) Parameter sensitivity of a watershed-scale flood forecasting model as a function of modelling time-step. *Hydrology Research* 44(2), 334–350.

RPS (2008) Further characterisation study: An integrated approach to quantifying groundwater and surface water contributions of stream flow. Southwestern River Basin District, Ireland.

Saltelli, A., Chan, K. & Scott, E.M. (2000) *Sensitivity Analysis*. Wiley New York.

Vogel, R. (2005) Regional calibration of watershed models. In: *Watershed Models* (ed. by Singh, V. & Frevert, D.), 47–71, CRC Press, Boca Raton, Florida, USA.

Wagener, T. & Wheater, H.S. (2006) Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *J. Hydrol.* 320(1–2), 132–154.