04 - Modelling Quantified Microbial Source Specific Pollution from Domestic Wastewater Treatment Systems during High Flows

Aisling Corkery,[1] John O'Sullivan,[1] Louise Deering,[1] Katalin Demeter,[1] Elisenda Ballesté,[2] Bat Materson,[1] Wim Meijer,[1] Gregory O'Hare[1]

[1]*University College Dublin, Dublin, Ireland*
[2]*University of Barcelona, Barcelona, Spain*

Abstract
The ability to model the transport of source specific faecal bacteria contamination in river networks can equip water resource managers with information of the different pathogens that are present. Such information can be particularly useful in catchment management plans for rivers from which potable water is extracted or where these rivers discharge to coastal zones where bathing or aquaculture is prominent and where the evaluation of human health risks is of primary importance.

This paper presents an application and performance assessment of the commercially available MIKE11 Hydrodynamic model for evaluating the fate of faecal bacteria of human origin, Human Gene Marker HF183f, from Domestic Wastewater Treatment Systems within the Dargle catchment. The Dargle is a spate river and the upper catchment is characterised by steep slopes that incorporates peat bogs and land used for forestry and agricultural purposes. Residential dwellings within the area are predominantly single detached units that rely on septic tanks for wastewater treatment. In the context of faecal bacteria of human origin, malfunctioning systems are of concern, particularly in terms of surface ponding, leakage to groundwater and direct discharge to surface waters. The MIKE11 model was calibrated in a two stages process. Firstly, the model was calibrated for prediction of discharge and microbial water quality parameters, namely E. coli and Intestinal Enterococci (IE), using data from a real-time sensor network within the catchment that comprised rain gauges, weather stations and water level recorders, data from which was used to determine flow records from stage-discharge ratings. E. coli and IE concentrations were determined from high resolution sampling during storm events. Following this, water quality samples taken during storm event sampling were used to identify and quantify the human gene marker HF183f using quantitative polymer chain reaction (qPCR) techniques. Results from the qPCR analysis were used to further calibrate the model at sub-catchment level for the transport of microbial bacteria derived from human origins. Using non-compliance statistics from the EPA National Inspection Plan, domestic sources have been calculated based on the percentage number of malfunctioning septic tank units and average daily faecal gene marker concentrations per household. The study highlights issues with how the fate of the human gene marker is modelled using MIKE11, particularly in terms of advection-dispersion inputs and the requirement to associate microbial concentrations with total runoff when modelling surface and groundwater pathways.

1. INTRODUCTION

Septic tanks or domestic waste water treatment systems (DWWTS) are commonly used to treat wastewater from residential dwellings in rural areas or small towns and villages that are not connected to a public wastewater scheme for treatment and disposal. Where the location, construction, operation and maintenance of these systems are inadequate, risks to water quality and human health may exist (EPA, 2013). This risk can be increased during high rainfall events, particularly in small rural catchments such as the Dargle River, Co. Wicklow. The Dargle is a spate river and the upper catchment is characterised by steep slopes that incorporates peat bogs together with lands used for forestry and agricultural purposes (Bruen *et al.*, 2001). Residential dwellings within the area are predominantly single, detached units that rely on DWTSs (CSO, 2012). In the context of faecal bacteria of human origin, malfunctioning systems are of concern, particularly in terms of surface

ponding, leakage to groundwater and direct discharge to surface waters (EPA, 2015). Recent inspections carried out by Wicklow County Council on behalf of the Environmental Protection Agency (EPA), suggest that approximately 58% of DWWTSs in the area are failing and approximately 42% are failing in such a way that they are causing a risk to both human health and the environment. System failures creating such risks are caused by a number of factors namely leaking to groundwater, discharging directly to surface waters and surface ponding, and occur in approximately 10%, 19% and 13% of DWWTSs in the area respectively (EPA, 2015). This paper presents an initial exploration and assessment of the MIKE11 hydrodynamic model for evaluating the fate of faecal indicator bacteria, namely *E. coli* and Intestinal Enterococci within the Dargle catchment in Co. Wicklow. Following this, it will also evaluate the performance of MIKE11 in modelling the human gene marker, HF183f, from Domestic Wastewater Treatment Systems at sub-catchment level (DWTSs).

2. METHODS

2.1 Study Area

The River Dargle catchment, including its tributary catchments totals just in excess of 133km^2. The river has its source in the Wicklow Mountains from where it flows into the sea at Bray Harbour. The river is characterised by a steep gradient, particularly in its upland catchment. The upper catchment is characterised by bog, forestry and agricultural lands, mainly arable, sheep, dry stock and dairy farming. The lower catchment has some agricultural land together with an urbanised coastal town making up the land use (Bruen *et al*., 2001). Residential dwellings within the upper catchment are predominantly single detached units that rely on septic tanks for wastewater treatment. In the lower catchment residential areas are generally connected to the public wastewater treatment system, but some older houses still use septic tanks (CSO, 2012). Two sub-catchments of differing characteristics located in the upper and lower Dargle catchments, were chosen for the study. The first of these, located in the upper catchment was the Glencullen catchment and the second, in the lower Dargle catchment was the Kilmacanogue catchment respectively – both rivers as shown in Figure 1 form part of the tributary network of the Dargle river. Over the course of the study, from the beginning of January 2012 to the end September 2012, the Glencullen River had an average base flow of 0.154m^3/s and a high water flow of 26.4m^3/s, while the Kilmacanogue River had an average base flow of 0.087m^3/s and a high water flow of 2.16m^3/s.

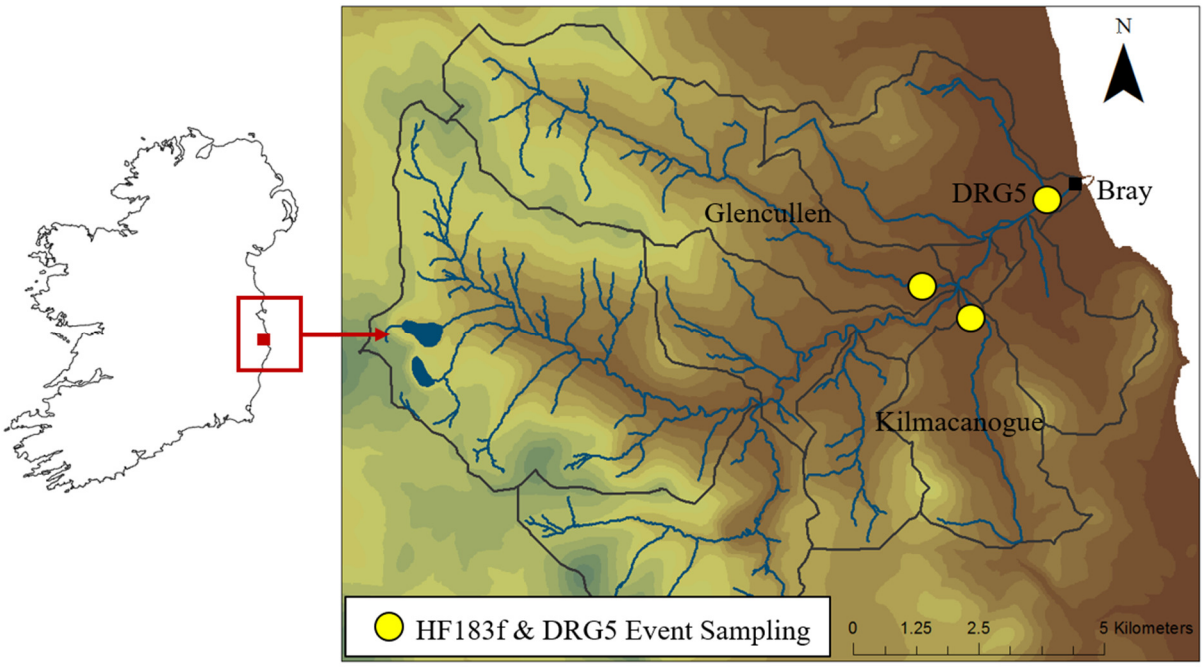

***Figure 1**: Dargle catchment, showing the modelled sub-catchments, DRG5 and the coastal town of Bray.*

2.2 Microbial Source Tracking and qPCR

Microbial Source Tracking (MST) is a molecular technique used to investigate the presence of faecal pollution and to discern its type of origin (human, animal etc.). DNA based methods employed are end point Polymerase Chain Reaction (PCR) (used for detection) and quantitative PCR (qPCR) (used for quantification). The presence of human faecal matter in the Dargle River, was analysed using these molecular methods. The Mike11 water quality model was then used to model and simulate the results from these data sets.

2.3 Data Collection

For the set-up and calibration of the catchment model, a sensor network within the Dargle catchment provided a continuous stream of rainfall and river water level data from each sub-catchment. Water levels were related to river discharge through rating relationships established at each of the river gauge locations during the study. Water quality data consisted of water samples taken periodically throughout 10 no. storm events at various locations within the catchment, including the Glencullen and Kilmacanogue sub-catchments and the Dargle River just upstream of the estuary (Drg5 in Figure 1). The storm events occurred during the period June – September 2012, coinciding with the summer bathing season. The catchment model formed part of a larger integrated catchment-coastal real-time forecasting bathing water quality models for the Bray coastal area, hence the summer bias with regard to sampling events, (see Bedri *et al*., 2014). These water samples were tested for Escherichia coli (*E. coli*) and Intestinal Enterococci (IE) at a catchment scale at Drg5 and sub-catchment water samples were tested for *E. coli*, IE and the human gene marker HF183f. Test results from the 10 no. sampling events were divided, with five event results being used in the calibration and validation of the catchment water quality model respectively.

2.4 Model Setup

The catchment modelling was performed in two-steps. Firstly, the NAM model simulated the rainfall-runoff processes at the sub-catchment level and MIKE11 routed this flow through the river network. Following this, MIKE11 simulated the in-stream processes (advection, dispersion and decay) of the transported water quality parameters (*E. coli*, IE and HF183f) in the river network. A 5 m Digital Elevation Model (DEM) of the Dargle catchment was used to delimit the sub-catchments and create the stream network. This, in addition to weighted areal rainfall, evapotranspiration data, and river flow data for a 9-month period from 1st January to the 30th September 2012 were required for the NAM and MIKE 11 model set-up and calibration. Weighted areal rainfall was computed by the NAM model from recorded rainfall within the sub-catchments using the Thiessen method, and daily potential evapotranspiration data was acquired from the nearest Met Éireann (Irish Meteorological Service) synoptic weather station. Model simulations identified two sensitive parameters; the maximum storage capacity in the surface storage zone (U_{max}), and the time constant for routing overland flow (CK_{12}). Values of other model parameters were derived from the literature (Doulgeris *et al*., 2012; Jennings *et al*., 2008, 2007).

Sub-catchments in the NAM model were linked to the MIKE11 model via river branch connections in the River Network Editor and cross-sections for each river branch were extracted from the DEM at longitudinal channel intervals of 200m. A Manning's *n* resistance of 0.04 was specified for channels in the river network. The majority of river channels in the network are mountainous streams with steep banks, with trees and brush along the banks submerged during high flows. From Chow (1959), a Manning's *n* of 0.04 was considered to be the most appropriate resistance value to represent these conditions. Upstream flow and downstream water levels were specified as boundary conditions and the global wave approximation was used for the routing of flow in the river network.

Initial water quality simulations, were based on diffuse loadings of *E. coli* and IE in the sub-catchments calculated from average daily *E. coli* and IE concentrations of cattle, sheep and horse faeces, together with animal stocking densities derived from agricultural census data that were consistent with the land-use within the catchment. *E. coli* and IE loads from wastewater treatment plants within the catchment were included as point source discharges to the model. The water quality model was calibrated by adjusting constant decay rates for *E. coli* and IE and comparing concentrations of simulated *E. coli* and IE to the measurements collected during the storm events. The decay rates (*k*) tested for *E. coli* and IE were 1.31, 1.15 and 0.57/day (T_{90} = 42, 48 and 96 hours) and 0.51, 0.32 and 0.25 (T_{90} = 4.5, 7 and 9-days) respectively. These values were consistent with ranges in similar studies that are reported in scientific literature (e.g. Bartram and Rees, 2000; Anderson *et al.* 2005). For the dispersion of *E. coli* and IE, the recommended ranges for small tributaries and upland reaches (1 to 5m^2/s) and lower reaches (5 to 20m^2/s) were varied along all river sections within the catchment to assess the sensitivity of the E. coli and IE response to changing dispersion. From these tests, a dispersion coefficient of 1m^2/s was assigned to tributaries and upland reaches and for the lowland reaches, a coefficient of 5m^2/s was used.

For microbial source tracking simulations water quality samples taken during one of the storm events was used to identify and quantify the human gene marker HF183f in two sub-catchments, namely the Glencullen and Kilmacanogue, using qPCR techniques. Results from the qPCR analysis were used to further calibrate the model at sub-catchment level for the transport of microbial bacteria of human origin. Using non-compliance statistics from the EPA National Inspection Plan and CSO census data, concentrations of domestic sources of the human marker were added to surface water runoff and groundwater pathways for each sub-catchment. These concentrations were calculated based on the average daily faecal gene marker concentrations (gcn/100ml) per household and the percentage number of DWTSs causing surface ponding and leaking to groundwater respectively as reported in the EPA National Inspection Plan. Furthermore, direct discharges to surface waters were calculated in a similar manner and added as a point source to each sub-catchment. Dispersion coefficients used in the previous calibration of *E. coli* and IE were retained for the modelling of HF183f, while decay rates (*k*) tested were 4.61, 2.3, 1.54, 1.32 and 1.15/day (T_{90} = 12, 24, 36, 42 and 48-hours). These ranges were in line with the findings from persistence studies on human markers reported in the literature (e.g. Tambalo *et al*., 2012 and Liang *et al*., 2012). Furthermore, HF183f persistence studies carried out over the course of this study gave T_{90} values of approximately 42-hours.

3. RESULTS AND DISCUSSION

3.1 River Flow

The two most sensitive parameters in the catchment model were U_{max}, the maximum storage capacity in the surface storage zone, and CK_{12}, the time constant for routing overland flow. The default auto-calibration range for U_{max} is 10 – 20mm, representing the storage capacity of most catchments. Steep catchments like the Dargle catchment however, have less storage capacity and consequently, it was considered appropriate to reduce this range in the auto-calibration process to input values of between 0.1 and 15mm and this gave U_{max} values that improved the fit to observed runoff and water balance in most sub-catchments. The auto-calibration range for U_{max} values adopted is consistent with ranges in published literature (see for example Jennings *et al*., 2008).

The Nash-Sutcliffe model efficiency coefficient (NSE) and the percentage error in computed Water Balance (WBL) were used to assess the fit between observed and simulated hydrographs. To illustrate the effects of reducing the U_{max} range in the auto-calibration process, Figure 2 compares the water balance and hydrographs respectively at the outfall of the Glencullen sub-catchment (a significant contributor in generating flows in the River Dargle) obtained using the reduced and the default U_{max} ranges. The NSE of 0.74 for the simulated hydrograph produced using the reduced U_{max} range is shown to be improved from the value of 0.71 using the default settings. The fit between simulated and observed runoff hydrographs was further improved in the auto-calibration of the CK_{12} parameter. While default values for CK_{12} between 10 to 50-hours are used for most catchments, significantly lower values are recommended for flashy or spate catchments. The Dargle and its tributary catchments are fast-responding and inputs of CK_{12} values between 3 and 11 hours were used to describe the time constant for routing the overland flow (DHI, 2013). Using the optimised CK_{12} value obtained for the Glencullen sub-catchment in the auto-calibration process produced further improvements of the NSE coefficient to 0.81 in the comparison of observed and simulated hydrographs Glencullen sub-catchment (Figure 3). Applying this two-stage calibration procedure to the other sub-catchments in the Dargle river system resulted in NSE coefficients that varied from 0.77 to 0.85 for the remaining sub-catchments of the river system.

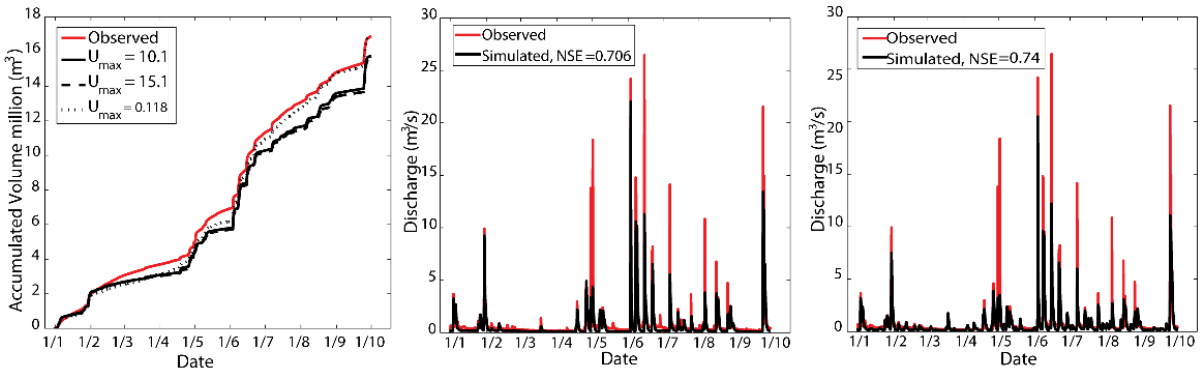

***Figure 2**: Water balance and hydrographs of the Glencullen sub-catchment obtained using the reduced and default U_{max} ranges.*

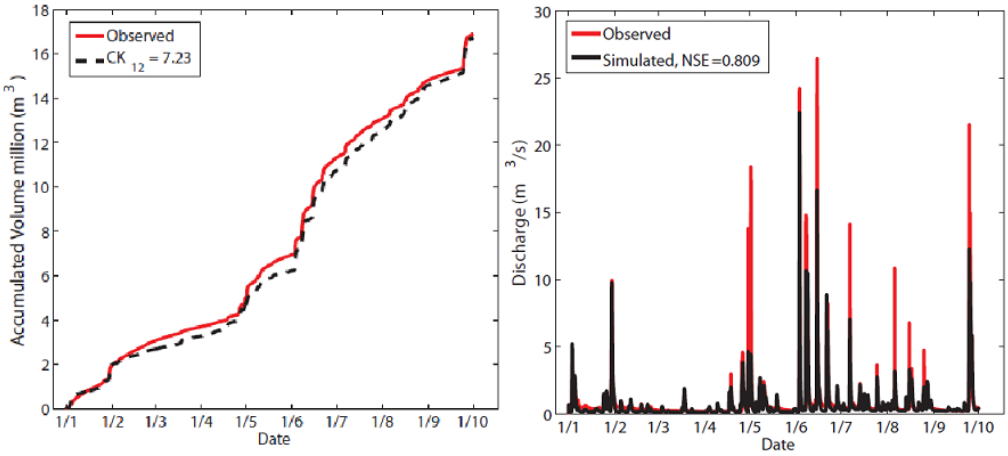

***Figure 3**: Water balance and hydrographs of the Glencullen sub-catchment obtained using optimised CK_{12} value.*

3.2 Catchment Water Quality

The best fit between simulated and observed *E. coli* and IE during peak flows was achieved using decay constants (k) of 1.31 d-1 (T90 = 42 hours) and 0.32 d-1 (T90 = 7-days) respectively. These decay rates are consistent with values used in similar studies that are cited in literature (see for example Bartram and Rees, 2000; Anderson *et al*., 2005) and confirm that die-off rates for *E. coli* in river systems are significantly lower than for IE. The performance of the catchment model in predicting faecal indicator bacteria concentrations was assessed by determining the Root Mean Square of Errors standard deviation ratio (RSR) (see Table 6 in Bennett *et al*., 2013 and developed by Moriasi *et al*, 2007).

$$RSR = \frac{\sqrt{\sum^{n}i=1(O-P)^2}}{\sqrt{\sum^{n}i=1(O-\bar{O})^2} \quad (1)$$

Where P and O are the predicted and observed values for a simulated variable (*E. coli* or IE) respectively, and $\bar{O}$ is the mean of the observed values. Performance was evaluated for simulated and observed *E. coli* and IE for four storm events at location Drg5 (Figure 1). The RSR values over the four events of 2.73 and 1.13 for *E. coli* and IE respectively, indicate that predictions of IE are more accurate than those for *E. coli*. However, RSR is greater than 0.70 which some modellers deem unsatisfactory. Comparisons of simulated and measured concentrations of *E. coli* and IE for the two storm events in Figure 4 show that measured concentrations of *E. coli* and IE rise and fall with the flow hydrograph. Simulated *E. coli* and IE concentrations are also shown to rise with the flow hydrograph and reach concentrations similar to measured values. However, on the receding limb of the flow hydrographs as shown for the two storms (Figure 4), simulated concentrations are shown to return to baseline values more slowly than for measured values. Differences between the simulated and observed concentrations in this regard suggest that issues exist in the manner the 'first flush' of *E. coli* and IE to the river system and subsequent dilution effects are considered in the modelling process. A study by Haydon and Deletic (2006) highlights similar overestimations in simulated bacteria concentrations when modelling the surface transport of pathogens to river systems in receding floods.

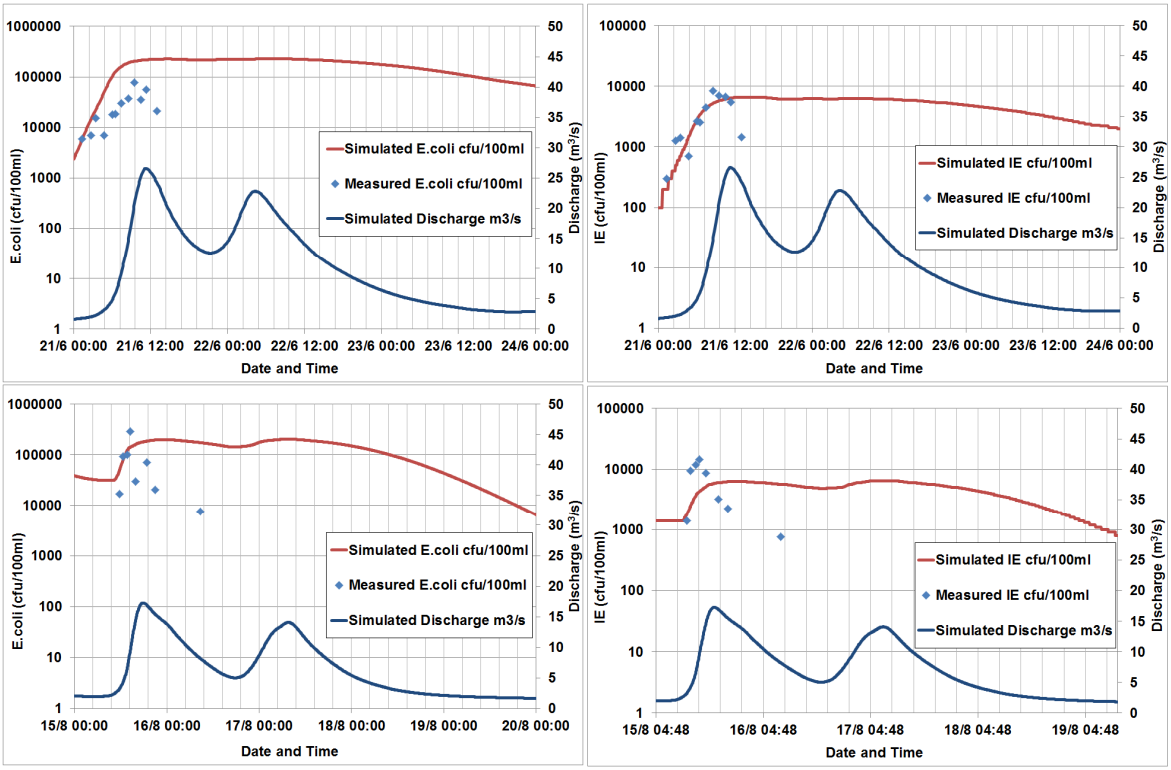

Figure 4: *Comparison of simulated and measured concentrations of E. coli and IE for two storm events.*

3.3 Sub-Catchment Microbial Source Tracking

The best fit to peak concentrations between simulated and observed values of the human marker HF183f during high flows was achieved using a decay constant (k) of 1.32 (T_{90} = 42 hours). Furthermore, and following from Gill *et al*. (2012), it was assumed that the human marker attenuation rate was similar to that of *E. coli* and that concentrations in the percolation area were 3.75% of those inside the DWWTSs. The performance of the catchment model in predicting human marker concentrations was also assessed using the Root Mean Square of Errors - standard deviation ratio (RSR).

The RSR values for the Glencullen and Kilmacanogue sub-catchments were 1.01 and 1.36 respectively. Again, it should be noted that RSR is greater than 0.70 which may be deem unsatisfactory. Figure 5 compares simulated and measured human marker concentrations for both sub-catchments during the same storm event. It can be seen that simulated human marker concentrations in the Glencullen River reach the initial rise in measured values but fail to reach peak concentrations. On the receding limb of the flow hydrographs, simulated concentrations return to baseline values more slowly than for measured values. For example, simulated concentrations in the Glencullen River took 3 days to reach baseline values in the order of $6x10^4$ for the modelled event (for visual clarity of measured versus simulated results this was not shown in Figure 5) Differences between simulated and observed concentrations suggest that issues exist with how dilution effects are considered in the modelling process. Again it is noted that Haydon and Deletic (2006) observed similar overestimations in simulated bacteria concentrations when modelling transport of pathogens to river systems in receding floods. Furthermore, when specifying distributed sources in the advection-dispersion module of MIKE11, microbial concentrations for surface ponding and leakage to groundwater could only be associated with total surface runoff and total groundwater runoff respectively, leading to elevated simulated concentrations in the receding limb and elevated background levels. It is possible that a load based model for microbial inputs would yield better results. Higher peaks in measured, relative to simulated concentrations, particularly in the Kilmacanogue sub-catchment, may be an indication of other sources of human marker, such as sewer misconnections or a greater percentage of

malfunctioning DWWTSs than was assumed (EPA non-compliance percentages for Co. Wicklow are only based on 31 DWWTS inspections).

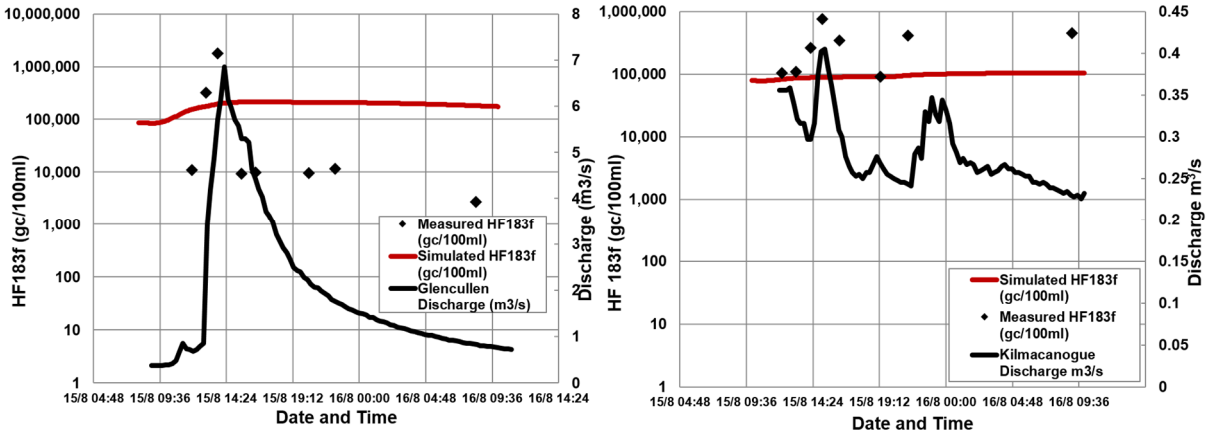

Figure 5: *Simulated v Measured HF183f in the Glencullen and Kilmacanogue River.*

4. CONCLUSSION

This study aimed to assess the performance of the commercially available software MIKE11 in modelling transport of the human gene marker, HF183f, from DWTSs to a river system during high flows, focusing on surface ponding, leakage to groundwater and direct discharge to surface waters as the main pathways. Such a model has the potential to be useful in catchment management plans for rivers from which potable water is extracted or for rivers that discharge to coastal zones where evaluation of human health risks is important.

Model calibration and validation showed that the hydrological and hydrodynamic components of the model have produced a reasonable fit to measurements (e.g. Nash-Sutcliffe efficiency coefficients of between 0.77 and 0.85 for the computed runoff in the sub-catchments of the Dargle river system. Water quality simulations at a catchment scale showed that the model has captured well the measured E. coli and IE at the peaks of the hydrograph. However, simulated E. coli and IE concentrations take longer to return to baseline levels, suggesting that the model does not accurately capture the 'first flush' and subsequent dilution effects as runoff continues to the river system on the receding limb of flow hydrographs.

Results from qPCR analyses were used to calibrate the model for human sources of contamination. The model gave reasonable results when evaluated using RSR, and simulated concentrations reach the initial rise in measured concentrations in the Glencullen River. However peak concentrations were never reached and simulated concentrations took much longer to return to baseline values on the receding flood. This highlights issues with how the fate of human gene marker is modelled using MIKE11, particularly in terms of advection-dispersion inputs and the requirement to associate microbial concentrations with total runoff with respect to surface water and groundwater. It is suggested that a load based modelled may be more appropriate for modelling transport of the human gene marker and other microbial contaminants.

Acknowledgments

The authors would like to thank the Interreg Ireland-Wales 2007-2013 EU Structural Funds Programme for the funding of the project.

References

Anderson, K.L., Whitlock, J.E., Harwood, V.J. (2005) Persistence and Differential Survival of Faecal Indicator Bacteria in Subtropical Waters and Sediments, Applied and Environmental Microbiology, 71, 6, 3041 - 3048.

Bartram, J., Rees, G. (2000) Monitoring Bathing Waters – A Practical Guide to the Design and Implementation of Assessments and Monitoring Programmes', E&FN Spon Press, London and New York. ISBN 0-419-24390-1.

Bedri, Z., Corkery, A., O'Sullivan, J.J., Alvarez, M.X., Erichsen, A.C., Deering, L.A., Demeter, K., O'Hare, G.M.P., Meijer, W.G., Masterson, B. (2014). An integrated catchment-coastal modelling system for real-time water quality forecasts. Environmental, Modelling and Software, 61, 458-476

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili Libelli, S., Newman, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V. (2013). Characterising Performance of Environmental Models. Environmental Modelling and Software, 40(0) 1-20.

Bruen, M.P., Chawla, R., Crowther, J., Francis, C.A., Kay, D., Masterson, B.F., O'Conner, P.E., Parmentier, B., Stokes, J., Thorp, M.B., Watkins, J., Wyer, M.D. (2001). Achieving EU Standards in Recreational Waters, Maritime Ireland/Wales INTERREG Report No.6, The Marine Institute. ISSN 1393-9025.

Central Statistics Office (2012). Profile 4 – The Roof Over Our Heads. The Stationary Office, ISBN 978-1-4064-2660-1.

Chow V.T. (1959) Open-Channel Hydraulics, McGraw-Hill, USA, ISBN 07-010776-9.

DHI, (2013) MIKE11, A Modelling System for Rivers and Channels, Reference Manual', DHI Group Horshølme, Denmark.

Doulgeris, C., Georgiou, P., Papadimos, D., Papamichail, D. (2012) Ecosystem Approach to Water Resources Management Using the MIKE 11 Modeling System in the Strymonas River and Lake Kerkini', Journal of Environmental Management, 94, 1, 132-143.

Environmental Protection Agency (2013). A Risk Based Methodology to Assist in the Regulation of Domestic Waste Water Treatment Systems. Environmental Protection Agency, ISBN 978-1-84095-481-4.

Environmental Protection Agency (2015). National Inspection Plan, Domestic Waste Water Treatment Systems, a Review of the Period 1st July 20130-30th June 2014. Environmental Protection Agency, ISBN 978-1-84095-481-4.

Haydon, S., Deletic, A. (2006). Development of a Coupled Pathogen-Hydrologic Catchment Model. Journal of Hydrology, 328 (3–4) 467-480.

Gill, L.W., O'Luanaigh N., O'Suilleabhain C., Misstear, B.D.R., Johnston P.M, (2012) Attenuation of on-site effluent contaminants through percolation areas of different Irish subsoils. EPA International Symposium on Domestic Wastewater Treatment & Disposal Systems.

Jennings, S., Glasgow, G., Fay, K., Hunter, T. (2008) Further Characterisation Study: An Integrated Approach to Quantifying Groundwater and Surface Water Contributions of Stream Flow, RPS Consulting Engineers Ltd, Belfast, UK.

Jennings, S, Elsaesser, B, Baker, G, Bree, T, Daly, D, Fitzpatrick, J, Glasgow, G & Hunter-Williams, T 2007, An Integrated Approach to Quantifying Groundwater and Surface Water Contributions of Stream Flow, Paper presented at National Hydrology Seminar, Tullamore, Ireland, 08/11/2007, pp. 72-81

Liang, Z., He, Z., Zhou, X., Powell, C.A., Yang, Y., Roberts, M.G., Stoffella, P.J. (2012) High Diversity and Differential Persistence of Fecal Bacteroidales Population Spiked into Freshwater Microcosm, Water Research, 46, 247-257.

Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingnew, R.L., Harmel, R.D., Veith, T.L. (2007) Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. Am. Soc. Agric. Biol. Eng. 50, 885-900

Tambalo, D.D., Fremaux, B., Boa, T., Yost, C.K. (2012) Persistence of Host-Associated Bacteroidales Gene Markers and their Quantitative Detection in an Urban and Agricultural Mixed Prairie Watershed, Water Research, 46, 2891-2904.