Published paper

Stovin, V.R. and Guymer, I. (2013) *A practical model to describe temporal variations in total suspended solids concentrations in highway runoff*. Acta Geophysica, 61 (3). pp. 706-731. ISSN 1895-6572
http://dx.doi.org/10.2478/s11600-013-0101-9

A Practical Model to Describe Temporal Variations in Total Suspended Solids Concentrations in Highway Runoff

Virginia R. Stovin[1*] and Ian Guymer[2]

[1]Department of Civil and Structural Engineering, University of Sheffield, Mappin Street, Sheffield, S1 3JD, UK, tel. +44(0)114 222 5051, email v.stovin@sheffield.ac.uk

[2]School of Engineering, University of Warwick, Coventry, CV4 7AL, UK.

Abstract

Techniques to predict temporal variations in concentrations and loads of suspended solids from highway runoff are required to estimate impacts on receiving water ecology and to inform the design of interception/treatment devices. A recent UK study included the collection of rainfall, highway runoff rates and sediment load and quality data from six different sites where motorway runoff drained directly into a receiving watercourse. This data set is used to critically evaluate a previously-published model (Kim *et al*, 2005) aimed at predicting temporal variations in runoff quality. The comparisons, based on discrete samples collected during 21 storm events, suggest that a simplification of the model, requiring just two parameters, provides a robust estimate of temporal variations in TSS. Generic parameter values are provided, and the model's application is illustrated. The model captures first flush effects well, but the identified generic parameters fail to fully-predict the variation in absolute TSS values that are observed in practice.

1 INTRODUCTION

1.1 Background

An integrated hydrological and biological research programme was jointly funded by the UK Highways Agency (HA) and the Environment Agency (EA) to provide authoritative advice on the circumstances in which highway runoff is likely to have a significant ecological effect on receiving waters, aimed at ensuring that the Highways Agency will meet the requirements of the EU Water Framework Directive. The research programme comprised 3 projects. Two of the projects aimed to develop ecologically-based receiving water standards for: i) soluble pollutants (via Runoff Specific Thresholds); and

ii) sediment bound pollutants, including an estimate of the likelihood for sediment deposition within the receiving water (Guymer *et al*, 2010).

The third project was intended to identify key pollutants and develop a predictive model for pollutant concentrations and loads in highway runoff. Predictive relationships for both soluble and sediment-derived Event Mean Concentrations (EMCs) were derived using multiple linear regression analysis (Crabtree *et al.,* 2009). The factors which were found to have a significant influence on pollutant concentrations were: climatic region, Annual Average Daily Traffic (AADT), month, maximum hourly rainfall intensity and antecedent dry weather period. However, the best degree of explanation, which was for dissolved copper, was only 38%. For some determinands, such as total cadmium and PAHs, multiple linear regression could not be used, so stochastic, Monte Carlo, simulation models were employed instead.

This research programme resulted in the development of a Highways Agency Water Risk Assessment Tool (HAWRAT) which is now the focus of the revised HD45/09 Guidance (HMSO, 2009). The spreadsheet-based tool employs a statistical approach to predicts EMCs throughout a 10 year rainfall time series for a local site, and compares the predicted concentration statistics with the derived Runoff Specific Thresholds to advise on the level of ecological impact. However, it does not attempt to describe the temporal variations in concentration or load during individual events, or relate this to the receiving water hydrological response.

As part of the second UKHA/EA project, which focused on sediment impacts, detailed studies were performed in six rivers in the UK that receive untreated highway drainage. At least 10 storm events were sampled at each site, and the amount of particulate material discharged during each event and its associated metal and PAH contaminants measured (Gaskell *et. al.*, 2004). Importantly, the sampling programme for this project included the collection of discrete – rather than composite – samples of highway runoff. More than 97% of particulate material discharged during storm events was found to be less than 63 μm in size and particle-associated contaminants were detected in all storm samples at all sites. In-situ deployments of invertebrates were performed on four occasions to assess the potential bioaccumulation of particle-associated contaminants in highway drainage. The results (Gaskell *et. al.*, 2007) showed that stream organisms were detrimentally impacted when sediments in the highway runoff deposited on the stream bed close to the outfall. This implies that the timing of sediment discharges needs to be considered alongside the hydrological response in the receiving water course in order to assess the risk of ecological impacts. In addition, the impacts of contaminants in runoff on the biota of receiving watercourses depend on the magnitude, duration and frequency of exposure (e.g. US EPA, 2000). Sediment in highway runoff has been shown to correlate strongly with individual pollutant loads (Sansalone *et al.*, 1998; Zanders, 2004). Luker and Montague (1994) suggest that up to 85% of pollutants are to be found as, adsorbed on, or absorbed by sedimentary particles.

The UKHA/EA sediment impacts project focused on assessing the quantity and quality of (representative) highway-derived Total Suspended Solids (TSS) and their impacts on stream ecology for a range of different receiving waters. It did not specifically set out to generate data that would enable TSS concentrations and loads to be predicted for unmonitored outfalls, although it is clear that this predictive capability is critical for the development of robust impact assessment tools. This paper therefore takes the opportunity to explore the previously-acquired data and seeks to identify a modelling approach that is capable of identifying temporal trends in TSS.

1.2 Previous research on highway runoff TSS

Many researchers have tried to develop useful models for predicting the quantity and quality of runoff from highways. These have ranged in scope from site-specific regression-based studies focusing on the prediction of EMCs from storm event parameters, to more complex, physically-based, models aimed at generating temporal variations in TSS through a more complete understanding of the influence of catchment characteristics and rainfall-runoff processes. However, the complexity of the underlying processes, and the unique characteristics of different locations and different rainfall events, means that there is no single widely-accepted robust, universal, modelling approach. Indeed, in a recently reported study measuring and predicting pollutant runoff from roads and parking lots in Korea, Maniquiz *et al.* (2010) presented results from over 40 events. For TSS they reported mean EMCs as 76 ± 95 mg/l and mean loads as 1.56 ± 2.42 kg, illustrating the large variability in values. This data was evaluated using multiple linear regression as a function of rainfall variables: total rain, antecedent dry days, rainfall duration and average rainfall intensity. Pearson correlation coefficients for EMCs were all negative, whilst for event loads the antecedent dry days counter-intuitively exhibited negative correlations. They conclude that with "*the high uncertainties*…*water quality sampling or long* *term monitoring is needed to gather more data that can be used for the development of estimation models*".

The following paragraphs cite some representative examples to provide a brief overview of the scope and limitations of the different types of research that has been undertaken.

Irish *et al.* (1998) developed a regression model for predicting total storm loads of constituents from highway run-off based on storm-water data collected from an expressway in Austin, Texas, USA. TSS was positively correlated with storm event characteristics (discharge, rainfall intensity and the antecedent dry period), whilst the intensity of the preceding storm showed a negative correlation. For a small event, preceded by a high-intensity event, the multiple regression relationship predicts a negative TSS load. This highlights one of the limitations of regression analysis, especially when predictions are made outside the range of the original variables. Overall the regression equations developed were able to describe over 90% of the observed loads in highway storm water runoff, but it should be noted that these are limited to total storm loads.

Opher and Friedler (2009) used data driven techniques (genetic algorithms) to develop and calibrate a predictive model for EMC of highway runoff pollutants. The models were trained and verified using 68 runoff events monitored in 92 highway sites in California between 1998 and 2004 and it is reported that the correlation between predicted and measured values of both training and verification data was mostly higher than previously-reported values. However, the approach is limited to a single, lumped, EMC prediction, and does not provide a basis for predicting temporal variations in TSS concentrations.

The model proposed by Massoudieh *et al.* (2008) considered both mobile and attached sediments were considered, together with the build-up during dry periods. Genetic algorithms were employed to calibrate the best-fit model parameters from field observations. However, in some cases the predicted and measured temporal variations did not closely match and no explanation was offered. They concluded that the technique lacked generalization, requiring site data for calibration and so could not be used in a predictive capacity.

Aryal *et al.* (2005) present data from long term monitoring in a highway drainage system in Switzerland. Suspended solids samples were taken at intervals corresponding to 0.12 mm rainfall in the 8.4 ha drainage area. Additionally, the drainage network, comprising 67 manholes and 280 gully pots, was simulated using general-purpose, deterministic drainage modelling software, InfoWorks-CS. A summary of the runoff simulation, comparing measured and predicted total outflow and peak discharge, was given for 11 storms, with 6 of the storm predictions categorized as good. The suspended solids concentrations were predicted based on an initial amount on the surface (maximum initial deposit of 12.5 kg/ha), sediment erosion and wash-off. In all predictions of a single discharge event the suspended solids load was significantly underestimated and this was attributed to the inadequate consideration of pipe sediment conditions. Long term simulations that permitted the consideration of initial surface and pipe conditions gave good agreement to measurements.

For the UKHA/EA data, information describing the components of the drainage system between carriageway and outfall was unavailable, and it may be argued that the effort involved in collecting the relevant data and generating a detailed hydraulic model for each individual outfall may be unjustifiably high. As a result several authors have considered simpler, semi-empirical, approaches which aim to predict runoff quality directly from either the rainfall or a measured or modelled runoff profile.

Kim *et al*. (2005) used data from a 3 year study to develop a new four parameter runoff quality model to describe temporal concentration variations for a suite of parameters. The general form of the equation, written in normalized time, i.e. as a proportion of storm duration, is:

$$c(t) = \delta + v(t)\{ \gamma^* + \beta^* \mathrm{Exp}[-\alpha v(t)]\} \qquad \text{(Eq. 1)}$$

where c(t) is the pollutant concentration, v(t) is the normalized cumulative volume (between 0 and 1) and the parameters α and γ* are related to total runoff, β* to rainfall, runoff coefficient and storm duration and δ is an initial concentration related to antecedent dry periods. Comparisons were shown between predictions and measurements for a few storms and it was concluded that individual storm event calibrations of the four parameter model could be used to fit to approximately 70% of the events. Event-specific calibrations were less good for TSS (R^2 = 0.84) than for some of the other pollutants that they considered. To use Equation 1 for predictions, the model parameters were related to storm characteristics. For TSS, Kim *et al*. (2005) correlated the parameters with the causal variables available within the database and obtained:

$$\alpha = 0.007(\mathrm{Trun}) + 3.83 \qquad \text{(Eq. 2)}$$

$$\beta^* = -1475\log_n(\mathrm{ARV}) - 9539 \qquad \text{(Eq. 3)}$$

$$\gamma^* = -83.74\log_n(\mathrm{Trun}) + 489.1 \qquad \text{(Eq. 4)}$$

$$\delta = 240.8\log_n(ADD) - 164.8 \quad \text{(Eq. 5)}$$

where Trun is the total runoff volume (m^3), ARV the average runoff velocity (m/hr) and ADD the antecedent dry days (days). The average runoff velocity (ARV) is defined as total rainfall (i.e. rainfall depth x catchment area) divided by catchment area and storm duration; it is therefore equivalent to mean rainfall intensity for the storm event.

The Kim *et al.* (2005) model was developed from a well-established US highway runoff database, its underlying structure appears to have been developed from a good understanding of the key controlling physical processes, and the authors provided formulae that enable parameter values to be estimated given only the catchment area and the storm event characteristics. For these reasons, there appeared to be a benefit in evaluating its ability to predict the observed TSS concentrations throughout a storm event in the context of the UKHA/EA sediment project data set.

2 METHODOLOGY

2.1 Field Data Collection

Six motorway/trunk road catchments in England were identified for data collection. The sites cover a range of geological, climatic, traffic flow, water chemistry and sediment characteristics (Gaskell *et al*, 2004). Surface runoff was carried from each highway catchment via a combination of gullies and/or filter drains. One of the major limitations of the UKHA/EA study data set is the lack of clear and reliable information regarding the as-built construction details of each of the sites. This includes a lack of specific information about the engineering detail of the drainage system. Similarly, verified drainage areas for the six highway sites were not available. Detailed survey work to obtain this information – for example via dye tracing – would likely have involved road and carriageway closures, and could not be supported by the project sponsors at the time. Best estimates of the catchment area (based on available drawings and/or site reconnaissance) for each outfall are provided in Table 1. Based on a mass-balance between measured volumes of rainfall and runoff, 'effective catchment areas' were determined for each of the six sites, which are also shown in Table 1. The effective catchment area is defined as the average area required to produce the measured runoff from the monitored rainfall, assuming that there was 100% runoff (i.e. 100% impermeable with no initial losses). The latter approach is limited in that it ignores initial losses, but the lack of any consistent agreement between the engineering best estimates and the mass-balance calculations suggests that catchment area should not be employed for model development. The lack of detailed information regarding highway catchments and drainage design is an acknowledged limitation and the HA are undertaking a detailed survey of all their assets.

A typical small scale receiving water, HA37, is shown in Figure 1a, while Figure 1b shows the highway and drain at HA12. At each of the sites, a tipping-bucket rain gauge measured the variation in rainfall intensity with time. Sensors installed in the highway drain (Figure 1c), just upstream of the outfall, recorded temporal variations in the turbidity, depth and velocity of the highway runoff.

In addition, 24 x 1-litre samples were taken from the drain by an automatic sampler for each storm at each site. Over a 125 minute period during a storm, 10 samples were taken at 2 minute intervals, followed by 5 samples at 5 minute intervals and 8 at 10 minute intervals, with a final sample 24 h later. This pre-determined sampling pattern did not always cover entire storm events. The sampler was triggered when the discharge, related to the depth above the temporary installed weir (Figure 1c), and turbidity of the water exceeded certain limits, so the sampled events are inherently biased towards 'larger' storms.

The runoff samples were centrifuged at 3000 rpm for 15 minutes, and the retained solids weighed to determine the sediment concentration. Centrifuging was adopted rather than the British Standard (BS) filtration method owing to the difficulty in retrieving the particulates from filter papers to perform chemical analysis. Additionally, the majority of the particles in the runoff were less than 45 µm in diameter and would not be retained by BS filtration.

Storm events were isolated from the continuously recorded raw data if the discharge and turbidity readings exceeded set values. The antecedent dry weather period was defined for each storm as the time from the end of the last rainfall event.

Although a minimum of 10 events was sampled from each of the six sites, only 21 storms were judged suitable for the present purpose. Reasons for the rejection of specific events mainly related to equipment failures (e.g. drifting or erratic depth sensor data, missing rainfall data (due to damaged or stolen rain gauges)) or poorly-timed runoff sampling. The 21 events encompass five out of the six sites; no records from HA12 were included. A summary of the events is provided in Table 2. Event notation is in the form of Site Name (e.g. HA01), followed by the date in yymmdd format. The estimated storm load was determined from the discrete sample load multiplied by the instantaneous flow rate, integrated for the 23 discrete samples taken over 125 minutes after the first sample.

The rainfall depths and durations for the 21 storms have been compared with the long-term data record for the relevant locations (FEH CD-ROM, NERC, 1999). The events range in depth from 1.2 to 15.4 mm, and in duration from 0.4 to 20.2 hrs. Rainfall depths correspond to between 12.5 and 139% of the expected 1 yr return period event, with the mean value being 58% (median 56%). Events sampled at sites HA01, HA08 and HA37 include events larger and smaller than the 1 yr return period event, whereas the data for HA09 and HA11 corresponds only to small (i.e. return period < 1 yr) events. It may be concluded that the data set provides a reasonable representative sample of rainfall events for river impact (water quality) applications.

The 21 storms included several multi-peaked events. As the TSS samples typically corresponded to one specific peak only, the relevant sub-event was isolated from the complete storm and the partial storm data (see Table 2) was employed in the model development. This is consistent with the ultimate aim of identifying a methodology that can be applied to predict TSS concentrations and loads associated with design (i.e. single peaked) rainfall events.

2.2 An assessment of modelling approaches

The model development comprised four phases. **Phases 1 and 2** focused on the Kim *et al.* (2005) model. **Phase 1** comprised a set of basic sensitivity analyses intended to characterise the model's inherent response to modifications to its four parameters. In **Phase 2**, a direct evaluation against the UKHA/EA dataset was undertaken. Preliminary findings from Phases 1 and 2 were reported by Stovin *et al.* (2010), and a summary of key conclusions is reproduced here.

The Stovin *et al.* (2010) study suggested that the Kim *et al.* (2005) model might usefully be reduced to a simpler two-parameter form, in which TSS is dependent upon the normalised cumulative proportion of total runoff volume, TSS(t)=f{v(t)}:

$$TSS(t) = v(t)\beta^* \, Exp[-\alpha v(t)] \qquad \text{(Eq. 6)}$$

Phase 3 of the modelling work therefore focused on the systematic identification of the two parameters, α and β*. The *lsqcurvefit* function in MATLAB (2007) was utilised to identify the best-fit parameter values, based on the monitored Q and TSS time-series data. This was done independently for each monitored storm event. For generic and practical model applications, it is necessary to identify suitable parameter values for application to unmonitored catchments. Regression analyses were therefore undertaken to establish potential mechanisms for estimating the parameter values from catchment and/or storm event characteristics. Scatter plots were generated to explore any potential dependencies of α and β* on the storm event characteristics identified in Table 2. However, these failed to reveal any clear dependencies; therefore further comprehensive statistical analysis was not felt to be justified. Peak TSS concentration, α and β* were examined for evidence of site-specific variations, but similarly this did not provide strong support for the inclusion of site-specific parameters within the model. The limited number of storm events also cautions against too much parameter fitting. Instead, a single set of generic values for α and β* was obtained by applying the *lsqcurvefit* function in MATLAB (2007) to the combined data set (all events). The validity and usefulness of this generic model is discussed in section 3.3.

One potential limitation of the (simplified) Kim *et al.* (2005) approach is that, mathematically, it can only predict a single peak in TSS. However, real runoff data often include complex temporal variations with multiple peaks in both discharge (Q) and turbidity (and TSS where available). Indeed, preliminary visual inspection of the UKHA/EA storm event data set suggested that in many cases TSS varied systematically in proportion to the measured flowrate. Therefore, an alternative TSS modelling approach was evaluated, in which TSS is dependent simply upon Q, TSS(t)=f{Q(t)}:

$$TSS(t) = kQ(t)^{n} \qquad \text{(Eq. 7)}$$

Again, MATLAB's *lsqcurvefit* function was utilized to identify the best-fit values of k and n. In this case the identified parameters varied widely, and no systematic dependencies with the obvious candidate variables emerged. This approach is therefore not discussed further.

All the modelling approaches described above require an accurate temporal runoff (Q(t)) profile as input. Although the TSS model development work made use of measured runoff profiles, it would be beneficial to provide a modelling approach that may be utilised to evaluate TSS load profiles (i.e. Q x TSS) for completely unmonitored catchments. **Phase 4** therefore focused on the potential to develop a suitable rainfall-runoff model. Jones *et al.* (2008) showed that a reasonable estimate of runoff for these systems could be generated from the rainfall record via a storage routing approach. However, the method was reliant on the use of catchment area data, which **–** as described above for this study

– cannot be relied upon. The application of the modelling approach described here is therefore limited by the requirement that runoff data is available, either from field monitoring or as a result of hydraulic modelling undertaken during the drainage design. Although all the analysis presented here was based on measured runoff data, there are many commercial drainage design tools in existence that could be deployed to estimate the temporal runoff profile from a highway outfall in response to design rainfall event.

3 RESULTS AND DISCUSSION

3.1 Sensitivity analysis of the Kim *et al.* (2005) model

Prior to any evaluation of the model's applicability to predict TSS concentration profiles from UK highways, it is important to check that the fundamental characteristics of the two datasets are comparable. Figure 2 compares the ranges of TSS EMC and mass loading values between the two data sets. The UKHA/EA data presented here is taken directly from the original UKHA/EA study report (Gaskell *et al.*, 2007), and therefore includes events that were subsequently removed or trimmed to generate the subset of 21 storms considered in the present context.

In Figure 2 a high degree of comparability in EMC values is observable, and this suggests that it is not unreasonable to attempt to apply the Californian model in a different regional context. The mass loading values are consistently around one order of magnitude lower in the UKHA/EA data set compared with the Kim *et al*. (2005) data set. This may reflect the fact that their data was collected in highly urbanized catchments, whereas the UKHA/EA data was collected in rural sections of trunk roads. It may also reflect uncertainties in the calculation of both total sediment mass and – certainly in the case of the UK data – catchment area. The UKHA/EA values were determined from a maximum of 24 flow samples, integrated with the runoff flow record; the Kim *et al*. (2005) data appears to have been predicted via use of their model fitted to observed sample values to enable interpolation and integration. Differences may also indicate sediment deposition in the UKHA/EA system at some point between the carriageway and the outfall/monitoring location.

In the model, the four parameters (α, β*, δ and γ*) are determined via regression-based relationships. Figure 3 shows, for each of the Kim *et al.* (2005) four model coefficients, how they vary in relation to the storm characteristics ADD, Trun and ARV. The figure also shows the range of the relevant storm characteristics experienced in both the Kim *et al.* (2005) study (open circles) and the present UKHA/EA study (+ symbols). In most cases there is a good range of overlap between the two data sets, although the Kim *et al.* (2005) data show fewer short ADD events and a significant number of long (>20 day) ADD events (which are not included in the graph for clarity). This is significant because for ADDs of less than 1.98 days it may be seen that the value of δ, which describes the initial TSS concentration, is negative. This is concerning, as negative TSS concentrations are physically not possible. At the other extreme, values of δ in excess of 400 mg/l (10 days ADD) seem high for a 'baseline' TSS concentration. Two of the remaining three parameters (β* and γ*) show both positive and negative values for the recorded ranges of storm data; α is always positive.

Figure 4 shows how selected combinations of these parameter values impact on the form of the predicted temporal concentration profile. The profiles are shown as a function of proportional storm runoff (v(t)). A 'baseline' parameter set was selected, with α = 10, β* = 1500, δ = 25 and γ* = -5. These values were chosen on the basis that they generate a profile of the type typically monitored for TSS during storm events, i.e. with a rapid rise to a peak concentration during the early part of the storm (first flush), followed by an exponential-type decay as the easily-eroded surface sediments become exhausted. These values are also typical of those used in Kim *et al.*'s own sensitivity analysis (their Figure 3). In each of the other profiles, just one of the parameters has been varied. It may be seen that an increase in α produces a decrease in the peak (the opposite of what is shown in Figure 3 in Kim *et al,* 2005). Use of a negative value of β* causes the profile to be approximately reflected vertically about y = δ, with the profile exhibiting an initial dip (dilution effect). Any alteration in δ displaces the profile vertically, with the value δ = -100 (which is possible for short ADD) generating a profile for which TSS is negative throughout the storm event. Increasing γ* from its baseline value of -5 to 90 generates a profile in which TSS increases towards the end of the storm. This is not commonly observed in monitored TSS profiles. A negative value of γ* (-90) results in a more rapid decline, with TSS predictions being negative for more than 50% of the total runoff volume.

Kim *et al.* (2005) state that one of the benefits of their modelling approach is the flexibility inherent within the model to represent the wide range of temporal contaminant concentration profiles that are observed in reality. However, the potential of the model to generate profiles that are entirely negative and/or showing increasing levels of TSS towards the end of the event must raise some doubts about its generic credibility.

3.2 Model Testing against the UKHA/EA Sediment Study Data Set

For the preliminary evaluation described in Stovin *et al.* (2010), three storms were selected from the UKHA/EA data set. The events correspond to the three largest drainage areas. Event HA01-050724 is representative of many of the medium to large long duration events, with multiple peaks in the rainfall and runoff response. Events HA11-060420 and HA37-050811 were both short duration, with the latter event having an unusually high peak intensity. The HA11 event has a very 'clean' almost design storm profile, whereas the HA37 event exhibits a double-peak in runoff. None of the selected events has a particularly long antecedent dry period, although they are all typical for this data set (see Figure 3). Initially TSS for each storm event was modelled using the parameter values derived from the relevant storm characteristics (Trun, ADD and ARV) according to the Kim *et al.* (2005) published relationships. However, the predictions were generally quite poor. TSS concentrations in the HA01 event were generally over predicted, with unrealistic final concentration levels in excess of 200 mg/l, approximately an order of magnitude greater than the observed data. Predictions for the HA11 and HA37 events, on the other hand, were both characterized by a fall in TSS at the start of the event, where the monitored data suggests a significant first flush.

It is not clear in Kim *et al.* (2005) that validation of the model using the derived parameter estimation relationships was undertaken. Their own storm characteristics would generate several storms for which the predicted values of TSS are negative and/or exhibit an initial drop to a minimum value at around v(t) = 0.2.

3.3 Proposed simplified model

Table 3 summarises the parameter values (α and β*) that were identified as best fitting the model (Eq. 6) to the observed TSS data for each of the 21 storm events. The R_t^2 parameter (Eq. 8, Young *et al.,* 1980) provides a measure of the goodness of fit of the predicted temporal concentration profile $p(t)$ to the measured data $c(t)$.

$$R_t^2 = 1 - \left[\frac{\sum_{t=1}^{n} \left(c(t) - p(t) \right)^2}{\sum_{t=1}^{n} (c(t))^2} \right] \qquad \text{(Eq. 8)}$$

A value of R_t^2 of 1.0 indicates a model that explains the data perfectly; values less than 1.0 indicate weaker explanatory (or predictive) capability. However, there are no absolute criteria for determining whether a model is sufficiently accurate for a specific application, and R_t^2 is more often used as a comparative measure of accuracy. In practical engineering terms different application-specific criteria may also be applied to determine whether a model's predictive capabilities are fit-for-purpose. Experience with the types of data sets being considered here suggests that R_t^2 values in excess of 0.95 indicate an "excellent" model fit, whereas R_t^2 values below 0.6 tend to indicate a model with weaknesses in terms of its practical predictive capability. This would correspond to models that over- or under-predict the peak concentration or timing by more than 50%, or significantly misrepresent the shape of the concentration profile. It may be seen that the TSS = f{v(t)} modelling framework generates an acceptable model ($R_t^2 \geq 0.6$) in 95% (20 out of 21) of the events.

An understanding of the physical process underpinning the observed TSS profiles would suggest that the parameter values for α and β* might be dependent to some extent on key event variables, including the Antecedent Dry Weather Period (ADWP), storm depth, duration and intensity. However, preliminary explorations using scatter plots (Figure 5) failed to reveal any strong dependencies, and it was felt that the limitations of the data set did not justify further or more statistically rigorous exploration.

The values of α and β* presented in Table 3 do not suggest any strong dependency on site characteristics, with considerable overlap between optimised parameter ranges at all five sites. There is some indication that both values are typically higher at HA37 when compared with the other four sites, but the sample size is too small to justify further statistical exploration of these differences. Similarly, Figure 6 presents the maximum monitored TSS value for each of the sampled storm events (prior to the data set being filtered for problematic rainfall or runoff data). Also indicated (solid square symbol) is the median value for each site. Considerable variation in the peak TSS concentration is observed between individual events. In comparison, the variation between sites is limited, with considerable overlap in observed peak TSS values. Again, there is some evidence of elevated TSS concentration levels at HA37.

Given the limited size of the data set, and the lack of any clear links between the model parameters and either rainfall event or site-specific characteristics, a lumped optimisation exercise was undertaken to identify the single generic values of α and β* that best fitted the complete data set. This identified α = 7.254 and β* = 5940. R_t^2 values associated with the generic model are included in the final column of Table 3.

Figure 7 illustrates the measured and predicted temporal TSS profiles for over one third of the monitored events. The event-specific (i.e. fitted) and generic parameter values have been used to provide two alternative model profiles. In general these plots provide confidence that the single (fixed parameter) functional relationship between TSS and v(t) provides a useful mechanism for capturing both the magnitude and temporal profile of TSS concentrations in highway runoff. The model may be considered to be robust in that the timing of the peak in TSS is generally well-predicted, and the TSS concentrations are always physically plausible (never negative) and generally reasonably accurate ($R_t^2 \geq 0.6$ for 67% (14 out of 21) of the events).

The model performs particularly well for relatively-simple, single-peaked, rainfall events. This is to be expected, as it is inherently limited to predicting a single peak in the TSS profile. Figure 7d demonstrates that in a more complex event, with three rainfall peaks, each of which generates corresponding peaks in the runoff and TSS profiles, the effects of sediment supply exhaustion are such that the decay in predicted TSS following the first peak provides a reasonable match to the observed data.

Although the timing of the peak TSS appears to be consistently good, the generic model parameters tend to generate a peak TSS concentration that is invariant at around 300 mg/l. This is an inevitable consequence of adopting a fixed value for α. The model may be observed to under and over-predict TSS values in some cases (e.g. Figures 7b, g and h). For most sites both over- and under-predictions are observed, but for HA37 the peak TSS concentrations in all events are consistently underpredicted. Of all the sites HA37 is the only trunk road, not having a hard shoulder or breakdown lane, all the others are motorways. Although the traffic loading is relatively low compared with the other sites, higher levels of vehicular acceleration, deceleration, stopping and starting may well lead to higher levels of TSS accumulation on trunk routes compared with motorways. This particular section of road is a major link to ferry terminals, and has a higher than normal Heavy Goods Vehicles (HGV) loading (Table 1), although the higher TSS levels may also reflect differences in the level of treatment occurring in the highway's conveyance system. It is believed that, in contrast to the motorway sites, HA37 does not include filter drains. For several storms, the monitored data for HA37 reveals a double-peak behaviour that does not seem to relate to rainfall. This probably relates some complexity in the drainage system.

Of the seven unsatisfactory ($R_t^2 < 0.6$) predictions using the generic model indicated in Table 3, three were for HA37. For the remaining four events, visual inspection of the predictions suggests that the model provides a reasonable estimate of the temporal profile shape and the timing of the peak. The main problem with the prediction is that, in all four cases, the model overpredicts the TSS concentration values. Measured peaks of approximately 100 mg/l are predicted to be close to 300 mg/l. Although it may be argued that such an estimate would be conservative for the planning of ecological impact mitigation measures, this is acknowledged as a limitation. Three of the four events were characterised by complex multi-period rainfalls.

As indicated in Table 3, there are a number of events for which the event-specific model and/or the generic model did not provide a good fit to the observed TSS data. In addition to the model's limitations with respect to particularly high or low TSS values, other cases of poor fits were predominantly related to multi-peaked, complex rainfall events and/or events in which the TSS samples failed to coincide in time with either the start or the peak of the event.

It may be concluded that the generic two-parameter TSS=f{v(t)} model provides a practical tool for the estimation of TSS temporal profiles in UK Highway drainage outfalls without the requirement for excessive levels of input data or modelling complexity. For single-peak (design-type) rainfall events, the timing of the peak TSS concentration is consistently predicted with a good level of accuracy. The identified generic model parameters result in a peak TSS concentration of approximately 300 mg/l. Although representative of the bulk of the data considered here, it must be appreciated that this value both under- and over-estimates actual monitored peak values. There is clearly scope for further work to improve upon this aspect.

4 MODEL APPLICATION

The present model is not intended to substitute for sophisticated deterministic modelling tools. However, the complex interactions between site and weather characteristics mean that a highly robust and accurate highway runoff prediction tool is not currently available to practitioners in the UK required to assess the potential ecological impacts of highway drainage design options on receiving watercourses. There remains a need for practical estimation methods with limited input data requirements. The proposed model improves upon existing EMC-based tools by providing a plausible estimate of the likely patterns of temporal variation in TSS concentration that will occur. Where practitioners in the field have local site knowledge and experience or other modelling tools at their disposal to predict peak TSS or EMC, it would be perfectly feasible to scale the temporal profile proposed here accordingly. Where no additional information is available, the current model may be considered to provide a plausible approximation to expected temporal variations in TSS.

The following section makes use of a synthetic rainfall profile, both to demonstrate why the temporal variation in TSS might be important for highway runoff impact assessment, and to outline a potential framework that might be adopted to undertake such assessments.

Figure 8 demonstrates the application of the modelling framework to a sample design storm. The storm corresponds to 10 mm rainfall distributed according to a UK symmetrical summer 75% peakedness profile (Flood Studies Report (NERC, 1975)). The catchment area was assumed to be 10,000 m^2. The catchment runoff has been generated assuming no initial losses, using a Muskingum storage routing model, as described in Jones *et al.* (2008) (K = 10 minutes, X = 0). The value of the reach time delay (K = 10 mins) was estimated as being typical of the delay between rainfall and runoff peaks evident in the field data (Figure 7). TSS concentration has been generated using the two-parameter v(t)-based model (Eq. 6), with α = 7.254 and β^* = 5,940. The sediment load profile is the product of runoff and TSS. The total load delivered to the stream during this event is estimated to be 10.9 kg, which – when distributed evenly across the total runoff volume of 1 x 10^6 litres – corresponds to an EMC of 108.6 mg/l. The equivalent EMC-based (constant) concentration profile is also shown, together with the corresponding temporal load profile. Because the peak TSS on the v(t)-based TSS model coincides with the rising limb and the peak of the runoff hydrograph, the total load conveyed to the stream rises sharply, such that at peak loading rate the outfall discharges 5.8 kg (53% of the total storm load) within a 10-minute period. The EMC-based model generates a significantly-reduced peak 10-minute load of 3.5 kg (33% of total storm load). These differences may prove to be critical when the impacts of highway outfalls on small streams are being considered. The peak load may well occur early on the rising limb of the stream hydrograph, when dilution/conveyance potential may be quite low. This may lead to problematic sediment deposits accumulating on the stream bed. Conversely, depending on stream hydrology and hydromorphology, this may mean that the contaminated sediments will be flushed from the immediate vicinity of the outfall as the stream discharge increases, reducing the potential contact time with sediment-based macro-invertebrates. Corresponding streamflow data recorded as part of the UKHA/EA monitoring suggests that, apart from HA01 where the discharge was into a stationary channel which only flowed when there was an overflow event, for all the other sites, the bulk of the sediment was discharged on the rising limb of the hydrograph.

Analysis was undertaken to evaluate the sensitivity of predicted TSS to the model's two parameters, α and β^*. They were each varied by +/- 20%, and the resulting TSS predictions are plotted as a function of v(t) in Figure 9. It may be seen that the basic shape of the distribution is relatively insensitive to either parameter, with the peak TSS occurring at a v(t) of 0.15. Increasing α or decreasing β^* results in a decrease in the peak TSS and also in the total load. An increase in α results in a slightly earlier peak. The peak 10-minute load accounts for a high proportion of the total load (67-73%) in all cases.

The model generates a clear first flush effect. For the model application described above, more than 50% of the sediment load is delivered to the stream within the first 25% of the runoff volume. For all the scenarios considered in the sensitivity analysis a minimum of 44% of the total load is associated with the first 25% of the storm runoff and more than 80% of the storm load is delivered with the first 50% of the runoff volume.

It may be argued that, when combined with a suitable design rainfall and hydraulic modelling tool capable of routing highway runoff to the outfall, the simplified 2-parameter v(t)-based model provides a useful framework for estimating TSS concentrations, and hence temporal load profiles, for use in receiving water ecological impact assessment procedures.

5 CONCLUSIONS

The UKHA/EA highway runoff sediments study TSS data set has been used to evaluate a model that was established from a USA database. The UKHA/EA data comprises rainfall, runoff, turbidity and suspended sediment concentrations for 10 storms recorded at each of 6 sites. Complete datasets were available from 21 storms and were judged suitable for the development of a TSS temporal modelling approach.

The Kim *et al.* (2005) model uses four parameters (derived from storm runoff characteristics) to predict TSS as a function of the cumulative proportion of total runoff. This general approach appears to be valid and useful in the present context. However, the previously published model shows some questionable behaviour, including negative TSS values and initial dips. Preliminary comparisons with the observed UKHA/EA data suggest that the calibrated model does not fit well.

A simplified, two-parameter, variant of the model has been shown to provide a practical means of modelling TSS profiles from UK highway outfalls, and generic parameter values have been identified. The model is reliable for timing of the peak, though further work is required to improve the accuracy with which absolute TSS concentration values are predicted.

6 ACKNOWLEDGEMENTS

This study was performed using data generated during work funded by The Highways Agency and Environment Agency on the project "Accumulation and dispersal of suspended solids in receiving watercourses" (Project 00Y91925).

7 REFERENCES

Aryal, R.K., Jinadasa, H.K.P.K., Furumai, H., Nakajima, F. (2005) "A long-term suspended solids runoff simulation in a highway drainage system", *Water Science and Technology*, 52 (5), 159-167.

Crabtree, B., Dempsey, P., Johnson, I. and Whitehead, M. (2009) "The Development of an Ecological Approach to Manage the Pollution Risk from Highway Runoff", Water Science and Technology, 59, 3, 549-555.

Gaskell, P. N., Guymer, I. and Maltby, L. (2004). Accumulation and dispersal of suspended solids in watercourses: Stage 1 Report. ECUS Ltd. and The University of Sheffield, Sheffield, UK.

Gaskell, P. N., Guymer, I. and Maltby, L. (2007). Accumulation and dispersal of suspended solids in watercourses: Stage 2 Report. ECUS Ltd. and The University of Sheffield, Sheffield, UK.

Guymer, I., Stovin, V., Gaskell, P., Maltby, L. and Pearson, J. (2010) "Predicting the deposition of Highway Derived Sediments in a Receiving River Reach", *17th IAHR-APD Congress*, *Auckland, New Zealand, 21-24 Feb, 2010*.

HMSO (2009) Design Manual for Roads and Bridges, Environmental Assessment, Vol. 11, Environmental Assessment Techniques, Part 10, HD45/09 Road Drainage and the Water Environment. HMSO, London.

Irish Jr., L.B., Barrett, M.E., Malina Jr., J.F., Charbeneau, R.J. (1998) "Use of regression models for analyzing highway storm-water loads", *Journal of Environmental Engineering*, 124 (10), 987-993.

Jones, A., Stovin, V., Guymer, I., Gaskell, P. and Maltby, L. (2008) "Modelling temporal variations in the sediment concentrations in highway runoff", Proc. of 11th Int. Conf on Urban Drainage, Edinburgh, Scotland, Sept.

Kim, L-H., Kayhanian, M., Zoh, K.-D., and Stenstrom, M.K. (2005) "Modeling of highway stormwater runoff", *Science of the Total Environment*, 348 (1-3), 1-18.

Luker, M. and Montague, K. (1994) Control of pollution from highway drainage discharges. CIRIA, London.

Maniquiz, M.C., Lee, S. and Kim, L-H. (2010) "Multiple linear regression models of urban runoff pollutant load and event mean concentration considering rainfall variables", *Journal of Environmental Sciences*, 22, 6, 946-952.

Massoudieh, A., Abrishamchi, A. and Kayhanian, M. (2008) "Mathematical modeling of first flush in highway storm runoff using genetic algorithm", *Science of The Total Environment*, 398 (1-3), 107-121.

MATLAB (2007) MATLAB version R2007b. Natick, Massachusetts: The MathWorks Inc., 2007.

NERC (Natural Environment Research Council), 1975, The Flood Studies Report.

NERC (Natural Environment Research Council), 1999, Flood Estimation Handbook (FEH) CD.

Opher, T. and Friedler, E. (2009) "A preliminary coupled MT–GA model for the prediction of highway runoff quality", *Science of The Total Environment*, 407 (15), 4490-4496.

Sansalone, J.J., Koran, J.M., Smithson, J.A. and Buchberger, S.G. (1998) "Physical characteristics of urban roadway solids transported during rain". *J. of Environmental Engineering*, ASCE, 124 (5), 427-440.

Stovin, V, Guymer, I, Gaskell, P and Maltby, L, (2010), "Evaluation of a highway runoff TSS models against new UK data", *17th IAHR-APD Congress*, *Auckland, New Zealand, 21-24 Feb, 2010*.

US EPA (2000) Science Policy Council Handbook, Risk Characterisation, USEPA Office of Science Policy, Office of Research and Development, WASHINGTON DC 20460.

Young, P., Jakeman, A. and McMurtie, R. (1980) “An instrument variable method for model order identification”, *Automatica,* 16, 281-294.

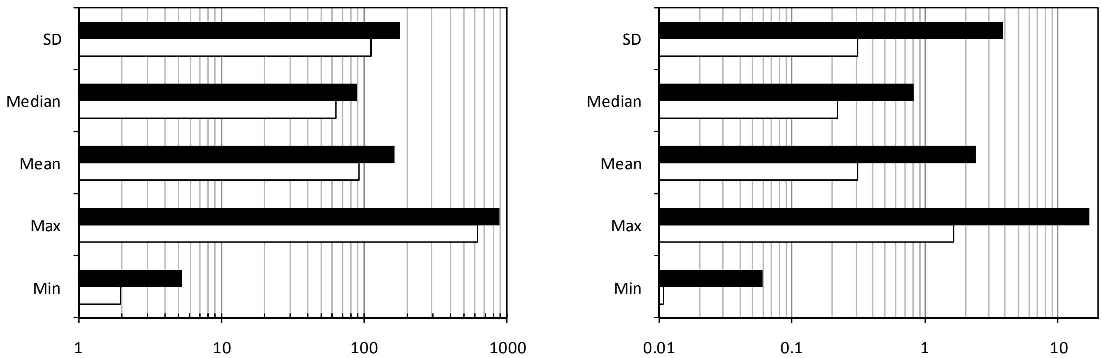
Zanders, J.M. (2004) “Road sediment: characterization and implications for the performance of vegetated strips for treating road run-off”. *Science of the Total Environment*, 339 (1-3), 41-47.

List of Figures

List of Tables

a) Receiving Water at HA37 b) Typical roadside drainage setup c) In-drain runoff sensors

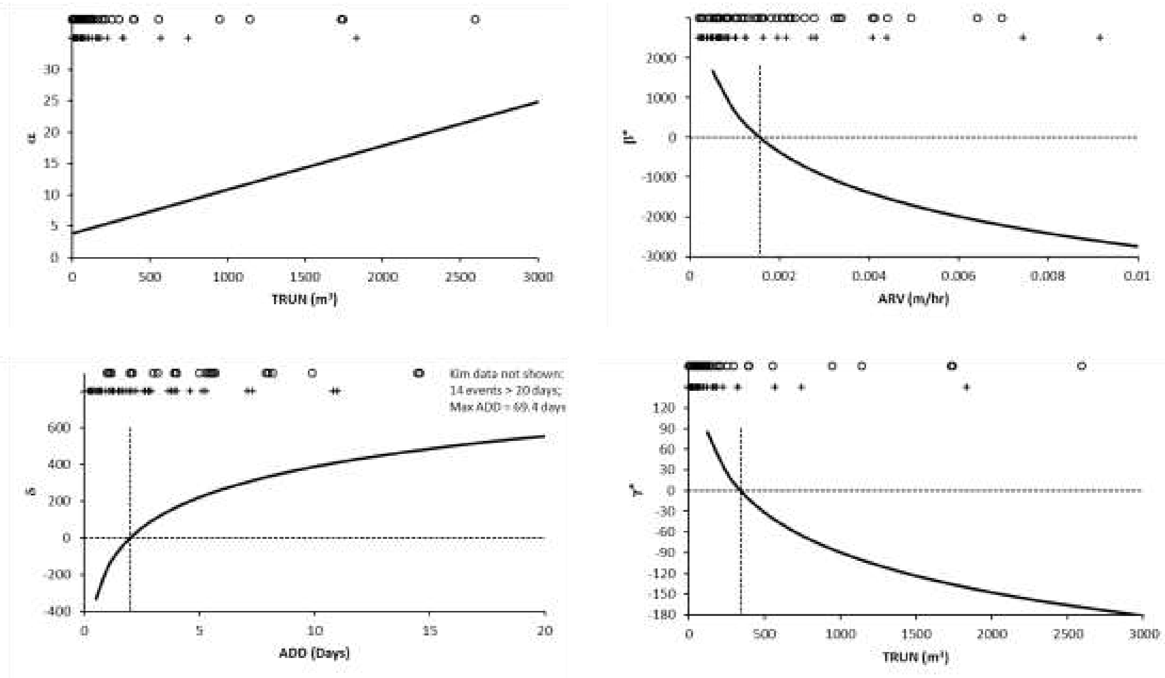
Figure 1 Site Photographs



■ Kim *et al.* (2005) data □ UKHA/EA Data

a) EMC values (mg/l) b) Mass loading values (g/m^2)

Figure 2 Comparison of TSS Concentration and Load Characteristics



○ Kim *et al.* (2005) data points; + UKHA/EA study data points

Figure 3 Sensitivity of the Kim *et al.* (2005) Model Parameters to Storm Characteristics

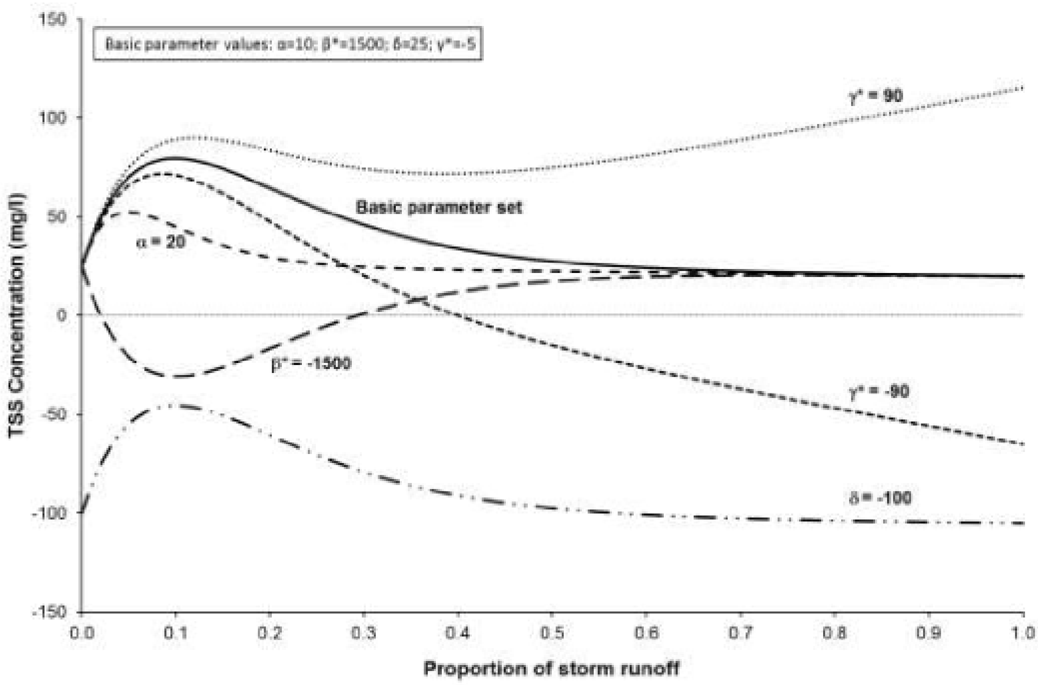


Fig. 4 Sensitivity of the Kim *et al.* (2005) Model to Parameter Values

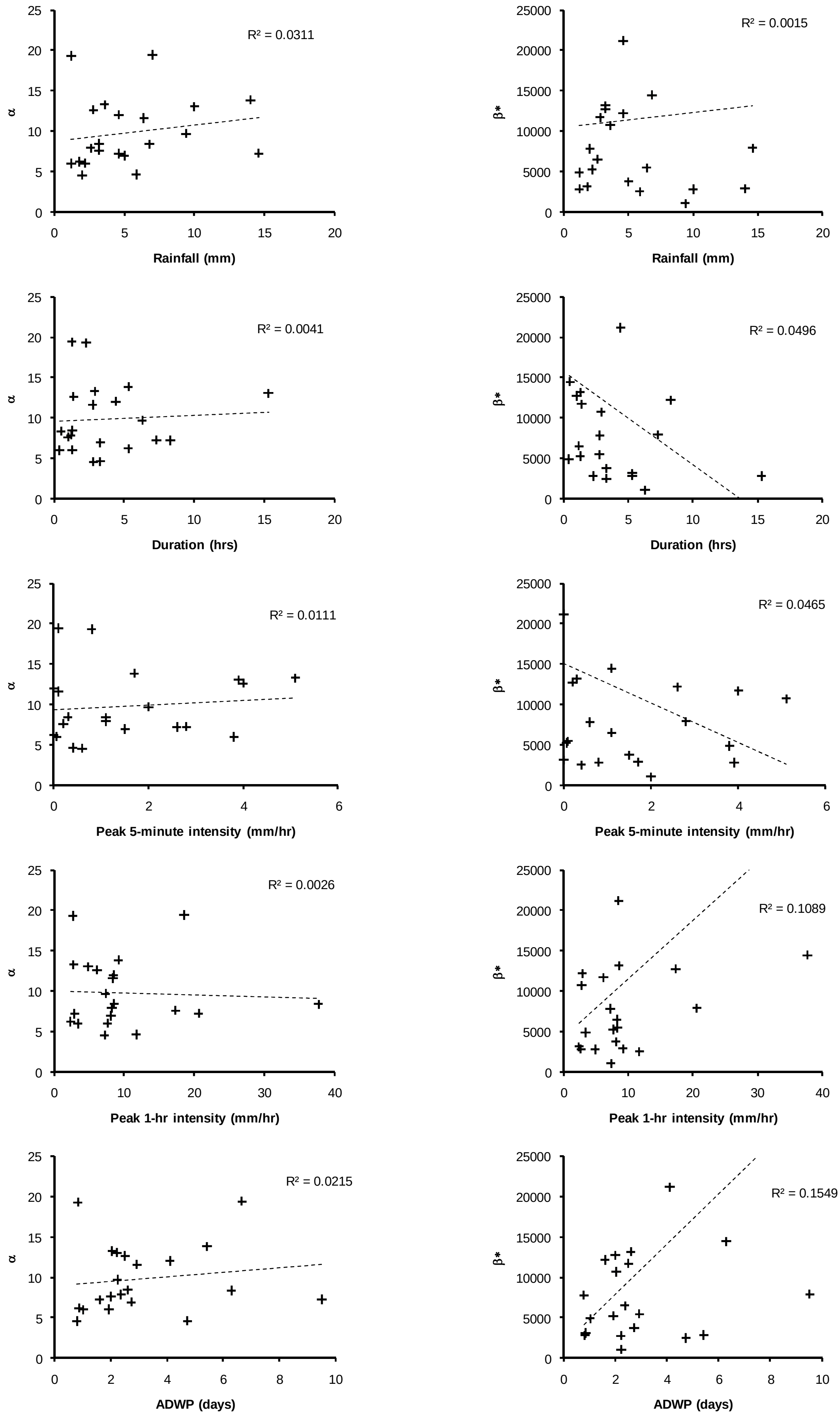


Figure 5 Scatter-plot assessment of potential determinants of model parameters α and β*

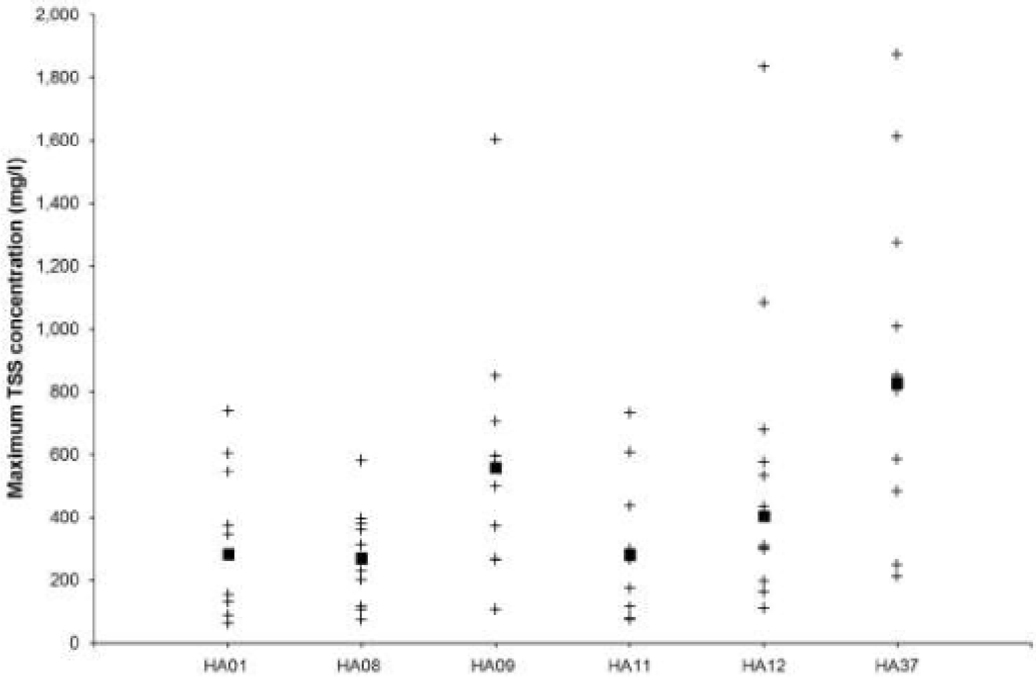
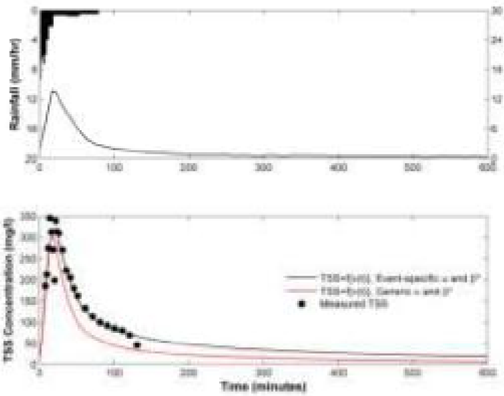
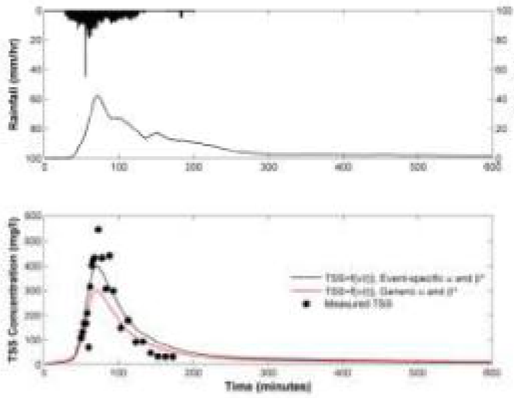


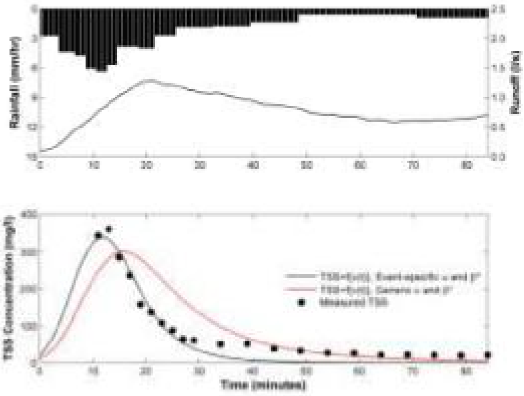
Figure 6   Maximum TSS values for all sampled storm events. Median values are indicated by the solid square



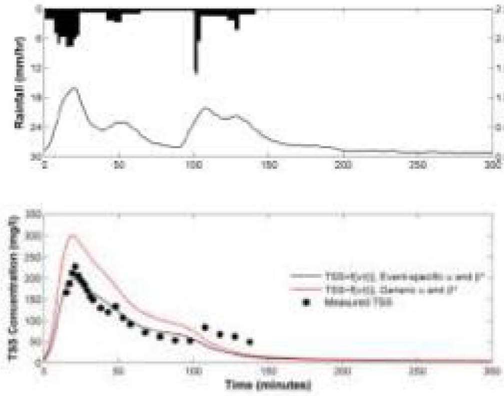
a) HA01_050812



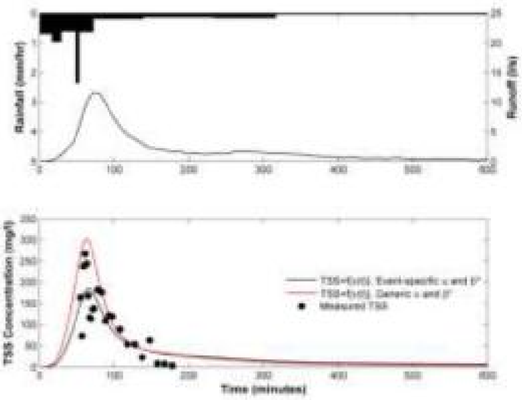
b) HA01_050822



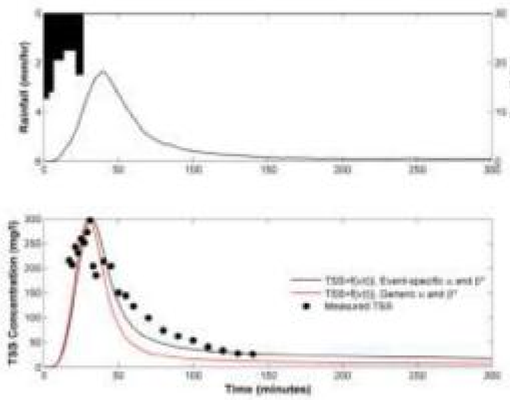
c) HA08_050813



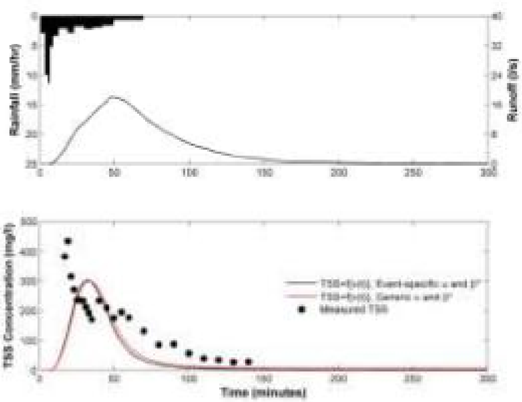
d) HA08_051021



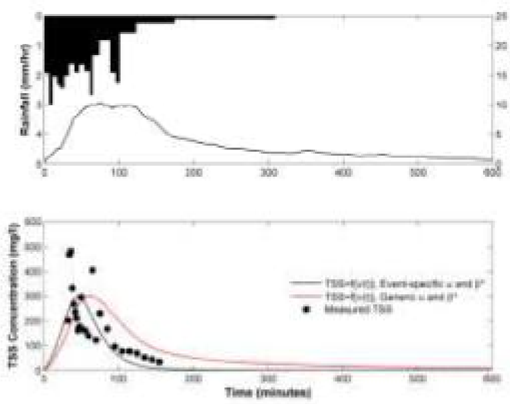
e) HA11_060113



f) HA11_060211



g) HA11_060420



h) HA37_050819

Figure 7 Predicted TSS temporal concentration profiles for selected monitored events

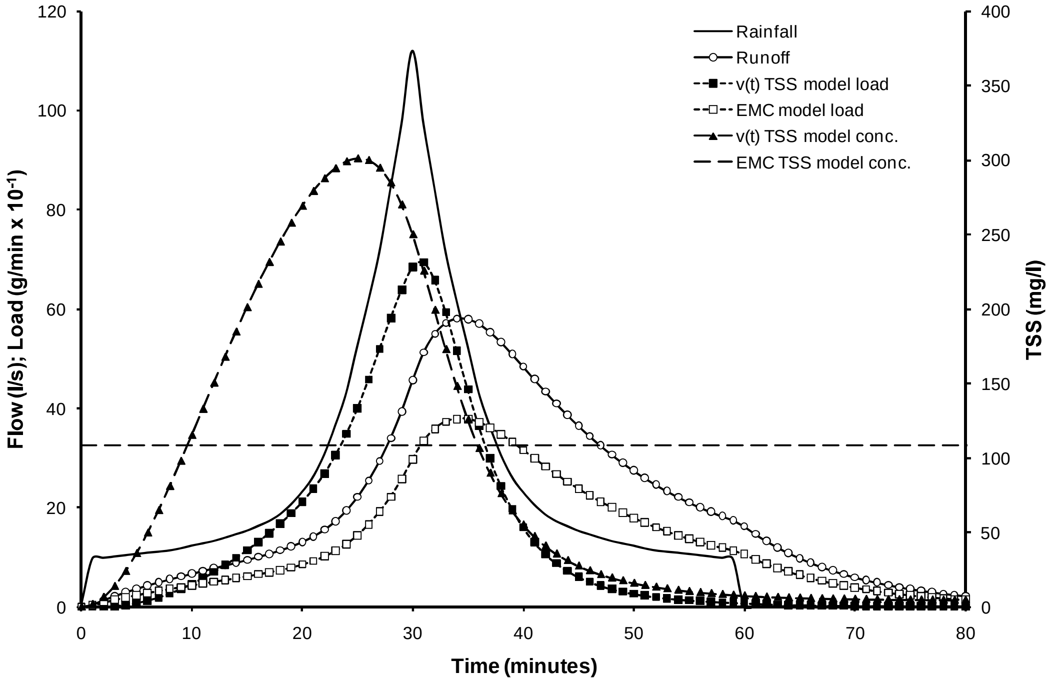


Figure 8 Application of the new TSS model to a design storm

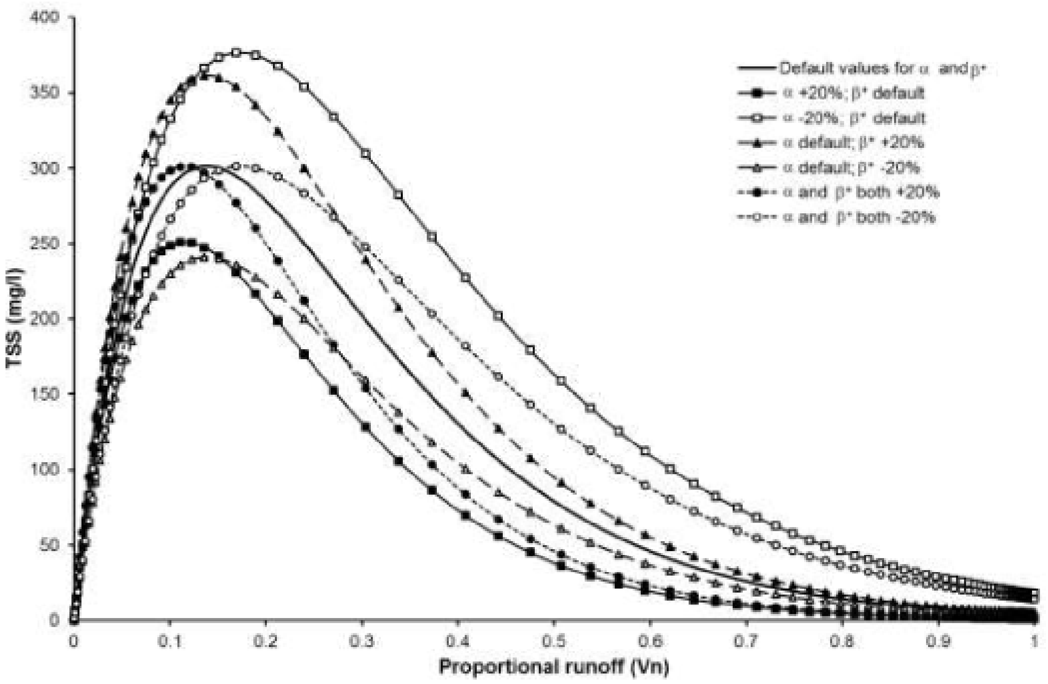


Figure 9 Sensitivity of the proposed TSS model to the parameters α and β*

Site name	HA01	HA08	HA09	HA11	HA12	HA37
Highway	M1 Sheffield	M5 Birmingham	M42 Birmingham	M6 Penrith	A1 Newton Aycliffe	A14 Newmarket
AADT	108500	94000	92000	44800	35981	43954
%HGV	19	20	19	22.5	-	27
Estimated catchment areas (m^2)[a]	19,000	19,500	12,780	>11,000	12,780	Unavailable
Effective Catchment Area (m^2)[b]	23,132	2,556	1,415	48,711	5,054	18,521

[a]Estimated from engineering drawings (where available) and/or site reconnaissance

[b]Estimated from the rainfall and runoff records, assuming no initial losses

	Event characteristics							Sub-event characteristics				
	Rainfall (mm)	Duration (hrs)	ADWP (days)	Runoff volume (m^3)	Storm Load (kg)	EMC (mg/l)	Mass Load (g/m^2)	Rainfall (mm)	Duration (hrs)	ADWP (days)	Peak i 5 mins (mm/hr)	Peak i 1 hour (mm/hr)
HA01_050724	10.0	15.3	3.9	227	1.477	6.5	0.064	10.0	15.3	3.9	4.86	2.22
HA01_050812	4.2	4.3	2.1	64	8.345	130.4	0.361	2.2	1.3	0.1	7.60	1.93
HA01_050822	14.6	7.3	2.8	331	37.658	113.8	1.628	14.6	7.3	2.8	20.63	9.52
HA01_060111	2.0	2.8	0.6	56	11.973	213.8	0.518	2.0	2.8	0.6	7.21	0.79
HA08_050724	14.8	8.4	1.7	35	0.635	18.1	0.248	14.0	5.3	1.7	9.18	5.43
HA08_050813	8.2	4.2	4.0	29	0.391	13.5	0.153	2.8	1.4	4.0	6.12	2.51
HA08_050915	3.2	1.3	0.3	4	0.450	112.5	0.176	3.2	1.3	0.3	8.57	2.60
HA08_050929	1.2	2.3	0.8	4	0.042	10.5	0.016	1.2	2.3	0.8	2.67	0.82
HA08_051021	5.0	3.3	1.5	6	0.486	81.0	0.190	5.0	3.3	1.5	8.08	2.73
HA08_051106	12.8	10.6	0.4	17	0.853	50.2	0.334	5.9	3.3	0.4	11.73	4.74
HA09_051018	9.4	6.3	2.0	7	0.059	8.4	0.042	9.4	6.3	2.0	7.33	2.23
HA09_051230	13.2	17.9	2.6	12	1.004	83.7	0.710	4.6	8.3	2.6	2.89	1.60
HA09_060214	3.2	1.0	0.2	2	0.160	80.0	0.113	3.2	1.0	0.2	17.33	2.00
HA11_060113	1.8	5.3	-	69	4.801	69.6	0.099	1.8	5.3	-	2.35	0.86
HA11_060211	1.2	0.4	3.8	55	7.374	134.1	0.151	1.2	0.4	3.8	3.40	1.03
HA11_060420	2.6	1.2	1.1	73	10.349	141.8	0.212	2.6	1.2	1.1	8.25	2.37
HA37_050811	6.8	0.5	1.1	69	12.984	188.2	0.701	6.8	0.5	1.1	37.71	6.30
HA37_050819	4.0	8.3	5.1	155	9.368	60.4	0.506	3.6	2.9	5.1	2.79	2.05
HA37_050910	15.4	14.3	4.6	184	8.766	47.6	0.473	7.0	1.3	0.1	18.55	6.67
HA37_050915	9.8	14.1	1.2	168	4.671	27.8	0.252	6.4	2.8	0.1	8.36	2.93
HA37_051012	9.4	20.2	3.8	68	5.266	77.4	0.284	4.6	4.4	0.0	8.48	4.11
Mean	7.3	7.1	2.2	78	6.053	79.5	0.344	5.3	3.7	1.6	9.72	3.12
Median	6.8	5.3	1.9	56	4.671	77.4	0.248	4.6	2.8	1.1	8.08	2.37
Max	15.4	20.2	5.1	331	37.658	213.8	1.628	14.6	15.3	5.1	37.71	9.52
Min	1.2	0.4	0.2	2	0.042	6.5	0.016	1.2	0.4	0.0	2.35	0.79
SD	4.89	6.09	1.55	87	8.47	59.1	0.36	3.9	3.5	1.6	8.20	2.26

	Storm Optimised values			Generic model (α=7.25, β^*=5940)
Site	**α**	**β^***	**R_t^2**	**R_t^2**
HA01_050724	13.04	2775	1.00	-12.78
HA01_050812	6.02	5225	0.98	0.95
HA01_050822	7.25	7891	0.92	0.87
HA01_060111	4.54	7807	0.99	0.65
HA08_050724	13.85	2849	0.82	-6.76
HA08_050813	12.63	11693	0.97	0.80
HA08_050915	8.44	13166	0.99	0.87
HA08_050929	19.31	2797	0.76	-53.33
HA08_051021	6.93	3728	0.98	0.74
HA08_051106	4.59	2478	0.85	0.66
HA09_051018	9.70	1031	0.73	-37.18
HA09_051230	7.21	12171	0.94	0.69
HA09_060214	7.61	12731	0.88	0.65
HA11_060113	6.20	3108	0.90	0.60
HA11_060211	5.98	4877	0.92	0.90
HA11_060420	7.88	6499	0.72	0.72
HA37_050811	8.36	14463	0.52	0.37
HA37_050819	13.27	10699	0.82	0.72
HA37_050910	19.42	86501	0.97	0.57
HA37_050915	11.60	5449	0.60	0.02
HA37_051012	12.02	21192	0.91	0.75
Mean	9.80	11387	0.87	-4.69
Median	8.36	6499	0.91	0.66
No. ≥ 0.6			20 [95%]	14 [67%]