

Data-Based Mechanistic Modeling of Flow-Concentration Dynamics of Non-Point Source Pollution: A Case of Upper Vaal Water Management Area

George, M. Ochieng¹. Frederick, A.O. Otieno². Julius, M. Ndambuki³

¹ Contact Author: Department of Civil Engineering, Tshwane University of Technology, Private Bag X680, Pretoria 0001, RSA. Tel: +27-12-3824773. Fax: +27-12-3825226. Email: ochiengmm@tut.ac.za

² (Prof): Faculty of Engineering & the Built Environment, Tshwane University of Technology, Private Bag X680, Pretoria 0001, RSA. Tel: +27-12-3825120. Fax: +27-12-3825568. Email: otienofao@tut.ac.za

³ (Prof): Department of Civil Engineering, Vaal University of Technology, RSA

ABSTRACT

This study aimed at applying the Data Based Mechanistic (DBM) modelling approach to develop a simple, parsimonious and discernable Flow-Concentration (F-C) model that can be partitioned into the various identifiable pathways associated with the pollutant at the catchment scale. An attempt was made to model the occurrence of acid mine drainage (AMD) in the Vaal River by using an indicator water quality parameter (sulphate concentration). The optimal Instrumental Variable (IV) methods of identifying and estimating discrete and continuous-time transfer function models as implemented in the CAPTAIN MATLAB[®] Toolbox were applied to the time series data in order to identify the appropriate model. A discernable F-C model was developed of three parallel pathways: the “quick-route” depletion pathway with a residence time $T = 5.06$ months; a “slow-route” build-up pathway with a residence time $T = 8.73$ months and a direct term component pathway regressed in 9 months. The resulting model showed that it is possible to use the DBM modelling approach to address the problem of representing the potential lag between polluting activity and its effect as well as provide more salient information about the system dynamics. This kind of information (i.e. the residence times and the advective time delays in the system) could prove useful for the catchment managers in making informed decisions including laying out remediation measures.

Keywords: Flow-Concentration Modelling, Non-Point Source Pollution, Data Based Mechanistic Modelling, Transfer Function models, Water Quality, Flow Pathways

INTRODUCTION

The challenge of water quality management associated with the principle of sustainable development has been of concern to many researchers and managers in the last decade [1]. In efforts to meet this challenge, the application of technologies as important tools for water quality management has been on the rise.

Among the technologies applied such as optimisation approaches, risk assessment and computer based techniques such as the Geographical Information System (GIS), water quality modelling is considered a principal technology for water quality management given its ability to relate causes (inputs) to effects (outputs) and predict the effects of control actions and changes to pollution sources [1]. Water quality models have thus emerged as useful tools for water quality simulation and prediction [1, 2, 3], an important recipe for judicious management of water as a precious resource. These models can project consequences of alternative management, planning, or policy-level activities such that effective management schemes can be identified.

Successful modelling of non-point source pollution (NPSP) exports from catchments eludes most conventional deterministic catchment water quality models. Apart from their complexity, over-parameterisation, intensive data and calibration requirements and overall cost limitations, studies show that complex models are not particularly appropriate in estimating catchment contaminant exports, even in fully instrumented catchments [4]. An identified part of the complexity in NPSP management lies in the inability of most models to represent the potential lag between polluting activity and its effect. This is particularly so for groundwater contamination and discharge into surface water resources [5]. Problems in NPSP modelling are further exacerbated by the dynamics involved in the NPSP process, the mechanisms of its occurrence are generally considered complex. The complexity of this problem may be due to the fact that the process is hydrologically driven (hence stochastic) and contaminants might enter a water body through the surface runoff or groundwater flow (base flow) rather than from discrete points [6]. By extension, decisions made from most NPSP models are inconclusive [7], in the sense that the modeler/decision maker might not be in a position to discern the various proportionate contributions to the total load from the various flow pathways (i.e. instantaneous flow – from surface runoff; quick flow – from intra flow; slow flow – from base flow and the constant or slowly varying contribution from the geological formation of the river course, also commonly referred to as the background concentration). This is a common scenario experienced while modelling NPSP using conventional deterministic catchment runoff models.

Although deterministic models still dominate environmental modelling, there is widespread consensus that simple modelling approaches using as few parameters as possible to represent key identifiable catchment hydrologic responses using the readily available catchment data is a promising strategy in catchment response modelling e.g. [8, 9].

This study aimed at modelling NPSP in particular the occurrence of AMD in the Vaal River by using an indicator water quality parameter. In line with the South African practice, total dissolved solids (TDS) could have been modeled as an indicator of AMD presence in the Vaal River. However, given its possible general interpretation (combining both point and non-point sources), it was opted for a parameter that could be directly linked to AMD effects/presence in surface waters. In this context, sulphate concentration was chosen as an indicator since it is a major anion associated with mining activities in areas affected by mining pollution.

In an endeavour to tackle the shortcomings of deterministic models, attempts were made to apply simple modelling techniques that have the ability to model a discernable F-C relationship that can be partitioned into the various identifiable pathways. This would go a long way in further understanding the contribution of the various flow pathways to the total pollutant load in the stream and if possible also determine the respective residence times. In this context, a Data Based Mechanistic (DBM) modelling approach was selected and applied in the Vaal River in the Upper Vaal water management area as a case study.

DBM model

The methodologies underpinning the DBM modelling approach are well documented e.g. [10, 11]. The general DBM model in transfer function (TF) terms for a single input single output takes the form:

$$y_t = b + \frac{B(z^{-1})}{A(z^{-1})}u_{t-\delta} + \xi_t \dots\dots\dots(1)$$

where: y_t is the measured output (e.g. pollutant concentration); b is a constant or very slowly varying output component (may or may not be present depending on the catchment response as inferred by the data); u_t is the measured input (e.g. measured river flow). The TF polynomials are defined as:

$$\begin{aligned} A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots\dots\dots + a_nz^{-n} \\ B(z^{-1}) &= b_0 + b_1z^{-1} + b_2z^{-2} + \dots\dots\dots + b_mz^{-m} \dots\dots\dots(2) \end{aligned}$$

where: z^{-1} is the backshift operator (i.e. $z^{-i}y_t = y_{t-i}$); δ is a pure advective time delay, measured in sampling intervals, which is introduced to allow for any temporal delay that may occur between the incidence of a change in u_t and its first effect on y_t and ξ_t is the stochastic noise (residuals) representing measurement noise and unmeasured effects and is defined as follows:

$$\xi_t = y_t - \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})}u_t \dots\dots\dots(3)$$

$\hat{A}(z^{-1})$ & $\hat{B}(z^{-1})$ are the estimates in Eq. 2. The model order is defined by $\{n, m\}$ while the triad $\{n, m, \delta\}$ defines the structure of the model.

METHODOLOGY

Data availability and treatment (pre-processing)

The results presented in this paper are for station number C1H007 located at Vaal River at Vaal at Goedgeluk. This station has a catchment area of 4686 km² and forms part of the larger Upper Vaal water management area in drainage region C [12]. The hydrological data (river flow) used in this study was obtained from the daily recorded data courtesy of the Department of Water Affairs and Forestry (DWAF). The water quality data used was also obtained from DWAF. Use was also made of the water quality software (Water Quality on Disc) from the South African Council for Scientific and Industrial Research (CSIR). This is a consolidated national surface water quality database on CD-ROM. The water quality data is obtained from the Hydrological Information System (HIS) supplied by the Directorate of Hydrology and the Institute for Water Quality Studies.

Streamflow data treatment

Even though the selection criteria used was thought to be sufficient to provide for “natural” streamflow data, possible instances could still occur where the observed discharge series are

superposed and altered by different impacts such as small unregulated farm dams. According to [13], this impact can be interpreted as noise or outliers. In this study, the said impacts were interpreted as possible noise in the data. The data obtained was thus assumed to be reliable and consistent and was not subjected to analysis for possible outliers.

Water quality data treatment

Given that the data obtained was so sparsely populated with very uneven monitoring frequency, a monthly time step in reference to the F-C modelling was applied. To accomplish this, the data had to be grouped in months (monthly in commensurate dates). This was done using the Statistics Toolbox in MATLAB[®] [14]. Fig. 1 presents the pre-processed monthly data used in the modelling.

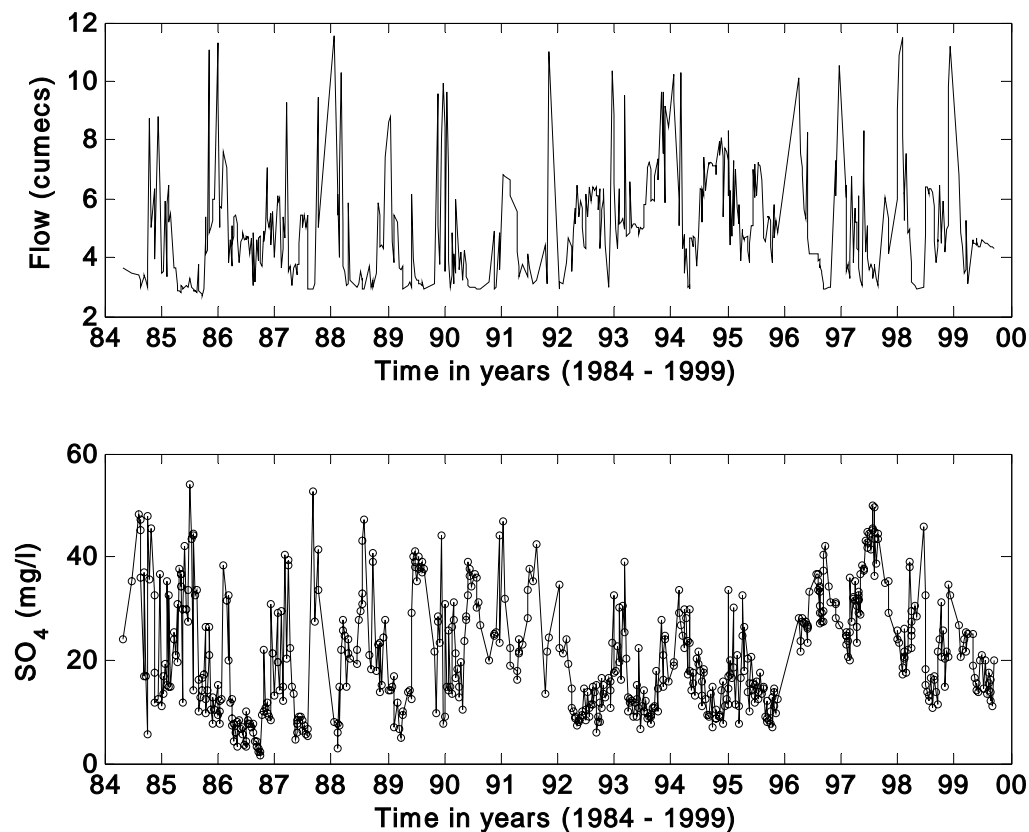


Fig. 1: Pre-processed monthly averaged flow and water quality data for station C1H007

DBM model identification and estimation

The DBM modelling philosophy proposes the model building i.e. identification, estimation and validation using time series data based on a suitable generic model class that is both capable of explaining the data in a parametrically efficient manner and producing a model that can be interpreted in physical terms. The DBM modelling methods described in [15, 16] and the references therein were applied in the model structure identification and parameter estimation of a single input single output (SISO) discrete time TF model to represent the flow-concentration dynamics in the catchment. The sulphate concentration was the expected output and the measured flow entering the model was the single input variable. The identification and optimisation

methods were implemented in the CAPTAIN MATLAB[®] Toolbox that applies the optimal Instrumental Variable (IV) methods to identify and estimate discrete and continuous time transfer function models.

Based on the data in Fig. 1, the Simplified Revised Instrumental Variable (SRIV) algorithm in the CAPTAIN MATLAB[®] Toolbox was used to identify and estimate several candidate models. The models were then analysed for quality in order of the resulting graphical fits and the returned values of the identification criteria that are: the Young Information Criterion (YIC) and the coefficient of determination (R_T^2) which is based on the variance of the model errors. Each model was also tested for meaningful physical interpretation of its resulting TF.

RESULTS

Of all the tried models, the following SISO TF model was objectively selected as the most appropriate one to represent the F-C dynamics:

$$C_t = \frac{-0.3213 + 0.4204z^{-1}}{1 - 1.6879z^{-1} + 0.7105z^{-2}} q_{t-10} + \xi_t \dots\dots\dots(4a)$$

OR

$$C_t = \hat{C}_t + \xi_t \dots\dots\dots(4b)$$

$$\Rightarrow \hat{C}_t = \frac{-0.3213 + 0.4204z^{-1}}{1 - 1.6879z^{-1} + 0.7105z^{-2}} q_{t-10} \dots\dots\dots(4c)$$

where the term C_t denotes the observed monthly sulphate concentrations as the output in mg/ℓ and q_t denotes the average monthly discharge as the input in m³s⁻¹. \hat{C}_t is the estimated/modelled monthly sulphate concentration in mg/ℓ. The deterministic part of the model \hat{C}_t has a YIC value of -3.24 and $R_T^2 = 22\%$ implying that based on the model error variance, the model explains 22% of the data. The error term ξ_t , quantifying all stochastic influences including measurement noise, was modelled as a AR (1) process of the following form:

$$\xi_t = \frac{1}{1 - 0.6187z^{-1}} \hat{e}_t \dots\dots\dots(5)$$

where $\hat{e}_t = C_t - \hat{C}_t$ is the estimated model error. The standard errors of the parameters characterising the model \hat{C}_t are given in Table 1. Fig. 2 presents a comparison of the deterministic part of the F-C model \hat{C}_t with the observed water quality values C_t .

Table 1: Standard errors of the F-C model transfer function parameters

Parameter	a_1	a_2	b_0	b_1
Estimate	-1.6879	0.71049	-0.3213	0.4204
Standard error	0.1915	0.1721	0.2489	0.3322

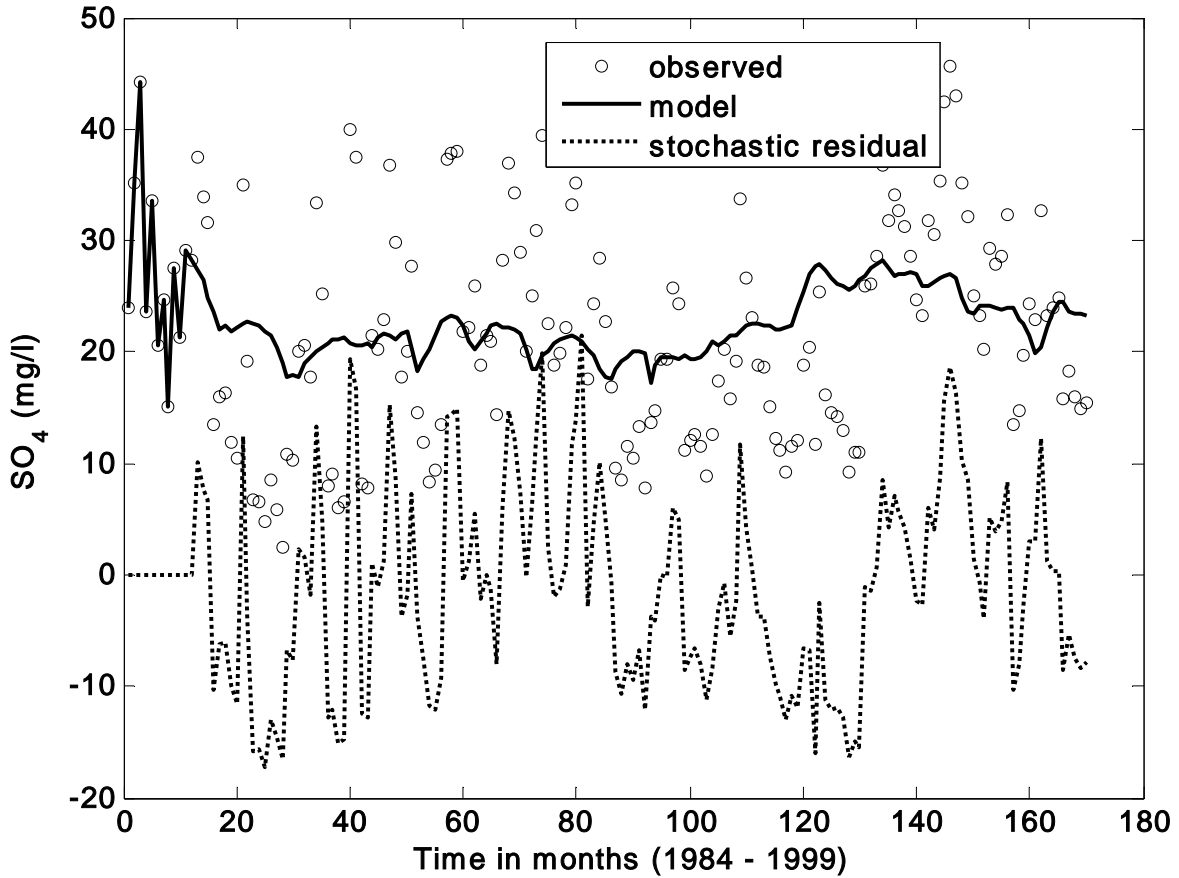


Fig. 2: Comparison of the F-C model transfer function output and the observed water quality data

DISCUSSION

The identified model explains 22% of the water quality data over the entire period. Despite the poor fit of the model to the data, it can be appreciated that the model estimates the data very well in the first 14 months; thereafter, the model follows the trend of the data quite well but does not give an accurate approximation. This behaviour could be attributed to the possible problem of lumped data where important information content on the system behaviour is lost. This observation also points to reiterate the problems encountered in lumped models; in this case, the data is lumped on a monthly basis and therefore critical dominant modes that could have been picked on a daily time scale could then be missed. Further, the uneven monitoring schedule of the water quality also contributes to the poor fit of the model since some months could have more data than others and hence the monthly averages would be definitely biased.

Another possible reason for a poor fit would be the fact that the SO_4 concentration is here related directly to the total streamflow. This is done under the assumption that all the components of flow are contributing to the build up in concentration. Possibilities are that only one of the flow components, for instance the groundwater, could be significantly contributing to the concentration build-up. It is possible that identifying and quantifying the major contributing flow path and modelling the flow in the respective pathway with the concentration would probably give a better model. However on average, the model represents the trend to some extent over the entire period.

On the basis of the ensuing discussion, it is argued that if the model was to be used for short time predictions (for instance periods less than a year), then a fair representation of the catchment dynamics would be obtained by the model. The decision to present the fit for the longer period was deliberate. This was in an attempt to investigate the longevity of the model in capturing long term catchment dynamics with respect to water quality in relation to observed flow. The rationale for this decision being that: as one of the objectives of the study was to model AMD related pollutants, it would be of significance if the model being used has a long term representation vis a vis the catchment's hydrological response. The sense in this argument is in the fact that AMD effects take longer to be manifested in the affected stream given that "drainage" as opposed to "wash off" would naturally have a longer time of concentration. In a longer time span, the model could be viewed as inaccurate. But this is expected as it would be ambitious for one to expect greater fit over a 15 year horizon given the data type and spread. However, it captures the catchment's flow-concentration dynamics adequately for the early period and could still be useful for predictive work over a limited time horizon. The model could thus be regarded as a fair representation of the average flow-water quality dynamics in the catchment especially in a shorter time period. To examine the model adequacy, an adequacy test was conducted on the model.

Model adequacy test

In determining the model adequacy, a correlation analysis was applied to the model stochastic residuals and the input variable so as to ascertain one of the Gaussian assumptions in respect of stochastic disturbances. Under this assumption, the error exists only in response (output) and not in the predictor data (input). Hence it is expected that a cross correlation exist with the response and not the predictor. In this regard, cross correlation analysis was done between the stochastic error and the input variable. Fig. 3 and Fig. 4 present the correlation analysis results. Fig. 3 presents the ACF of the stochastic residuals of the model while Fig. 4 presents the CCF between the stochastic residuals and the recorded streamflow input.

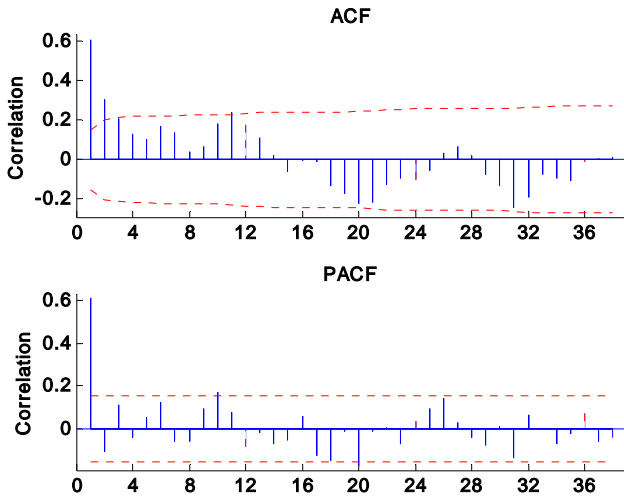


Fig. 3: Auto-correlation function of the F-C model stochastic residuals

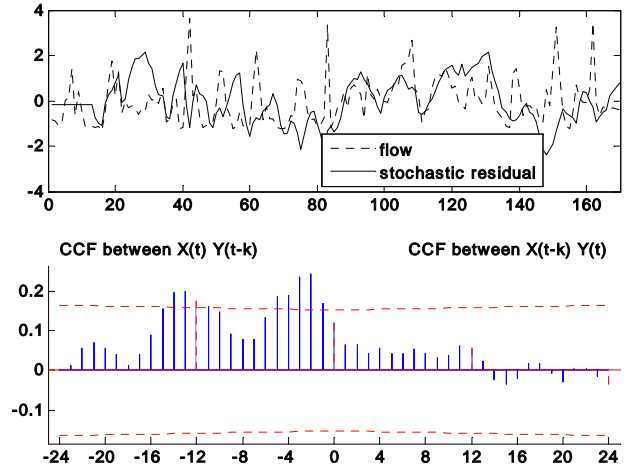


Fig. 4: Cross-correlation function between input (flow) and stochastic residuals

From Fig. 3, it can be observed that there is a significant auto-correlation in the stochastic residuals at the 1st and 2nd lags. Thereafter the auto-correlation is insignificant as required. Given the model performance with respect to the input-output data, the observed auto-correlation in the model residuals would be expected. The evident auto-correlation could also be as a result of possible input feedback in the system. From Fig. 4, some significant correlation is evident on the negative side that could also imply an input feedback existing in the system. This feedback could be as a result of some unobserved components characterising the catchment hydrologic dynamics. These could be nonlinear relationships between flow and water quality or possible chemical interactions within the system that at this stage could not be discerned. Examples of such interactions could be the processes of sediment adsorption and release of the sulphates in the system. However there is no positive cross correlation between the stochastic residuals and the input. Thus it could be assumed from the results that the stochastic residuals ξ_t and the flow input q_t are reasonably uncorrelated as required.

Mechanistic interpretation of the F-C model

By partial fraction expansion techniques, Eq. 4c was decomposed into different possible parallel pathways. The resulting single order transfer functions were of the following form:

$$\hat{C}_t = \frac{5.5251}{1-0.8854z^{-1}} q_{t-10} - \frac{17.6952}{1-0.8024z^{-1}} q_{t-10} + \left[\begin{array}{l} 12.1701 + 9.3066z^{-1} + 7.0615z^{-2} + \\ 5.3065z^{-3} + 3.9395z^{-4} + 2.8790z^{-5} + \\ 2.0605z^{-6} + 1.4322z^{-7} + 0.9535z^{-8} + \\ + 0.5917z^{-9} \end{array} \right] q_{t-10} \dots(6)$$

The model identified in this case is one with three parallel pathways. The three parallel pathways are identified here as the “quick-route” depletion pathway, “slow-route” build-up pathway and a direct term component pathway that is regressed in 9 months. These pathways combine additively to yield the modelled concentration \hat{C}_t . For brevity, Eq. 6 can be rewritten as follows:

$$\hat{C}_t = y_t^{(q)} + y_t^{(s)} + y_t^{(d)} \dots\dots\dots(7)$$

$$\Rightarrow y_t^{(q)} = \frac{5.5251}{1 - 0.8854z^{-1}} q_{t-10}; y_t^{(s)} = -\frac{17.6952}{1 - 0.8024z^{-1}} q_{t-10};$$

$$y_t^{(d)} = \left[\begin{array}{l} 12.1701 + 9.3066z^{-1} + 7.0615z^{-2} + 5.3065z^{-3} + 3.9395z^{-4} + 2.8790z^{-5} + \\ 2.0605z^{-6} + 1.4322z^{-7} + 0.9535z^{-8} + 0.5917z^{-9} \end{array} \right] q_{t-10}$$

Where: $y_t^{(q)}$ is the “quick-route” depletion component, $y_t^{(s)}$ is the “slow-route” build-up component and $y_t^{(d)}$ is the direct term regressed in 9 months. The details of the individual first order TF in the decomposition are given in Table 2. Fig. 5 presents the sulphate concentration pathway components generated by the parallel decomposition.

Table 2: Decomposed transfer function components of the F-C model

Transfer Function (TF)	Decomposed TF components			
	RT	SSG	TC	% \hat{C}_t
TF(A) – Build-up component	0.8854	48.21	8.73	26.28
TF(B) – Depletion component	0.8024	89.55	5.06	48.81
TF(C) – Direct terms component	0	45.70	0	24.91

RT = Root or eigen value of the decomposed TF; SSG = Steady State Gain; TC = Residence time; % \hat{C}_t = Percent of \hat{C}_t , the modelled sulphate concentration = $\frac{SSG}{\sum SSG} * 100$

Note: For brevity, the direct term component is considered here as a lumped unit being the summation of all the individual direct terms given in $y_t^{(d)}$. This component is herein referred to as transfer function C.

In examining Eq. 6, it can be observed that there is a large advective time delay of 10 months. This is the time taken for the input effects to manifest on the output. Using some basic ideas on the generation and dynamics of AMD, we could attempt to directly link the time delay to rainfall causing flow and the consequent changes in concentration of the AMD related constituent. Further, the decomposition produces real eigen values (the roots of the $A(z^{-1})$ polynomial) from which three distinct parallel pathways can be distinguished: the “quick-route” depletion pathway associated with the 0.8024 eigen value and resulting in a residence time of $TC = 5.06$ months; the “slow-route” build-up pathway associated with the 0.8854 eigen value resulting in a residence time of $TC = 8.73$ months and a direct term component pathway whose effect is regressed in 9 months. As a prelude to a further analysis of the possible scenario, a brief discussion on AMD dynamics and abandoned mine behaviour follows.

Acid producing materials composed of pyrite react at much slower rates and produce only small amounts of AMD when left in an anaerobic condition underground [17]. Generally, pyrite oxidizes at a rate of between 4% and 8% per year and at an average rate of 7% per year

implying that the acid production is a much more gradual process. AMD is a serious problem in areas of extensive surface and underground coal mining such as the Upper Vaal WMA that is rich in coal deposits especially in the Mpumalanga region. In underground coal mining, the water table is lowered to a great depth sometimes up to 100 m [18]. Once the mines are closed, the groundwater flow regime begins to change due to the refilling of the depression cones. The recovery of the water table could take a long time depending on the tortuosity and heterogeneous permeability of the old mine workings as well as the hydrology of the catchment. The recovering water table brings with it a cocktail of contaminants such as AMD [19].

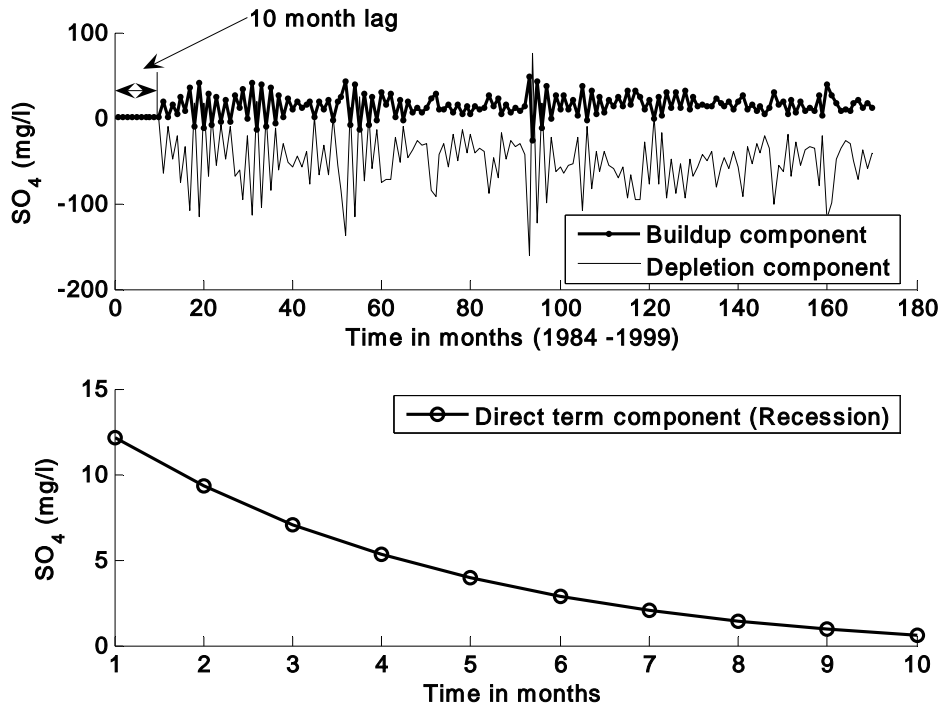


Fig. 5: Sulphate concentration pathways generated by parallel decomposition of the F-C model transfer function

The model components in their decomposed parallel pathways are presented as a block diagram in Fig. 6.

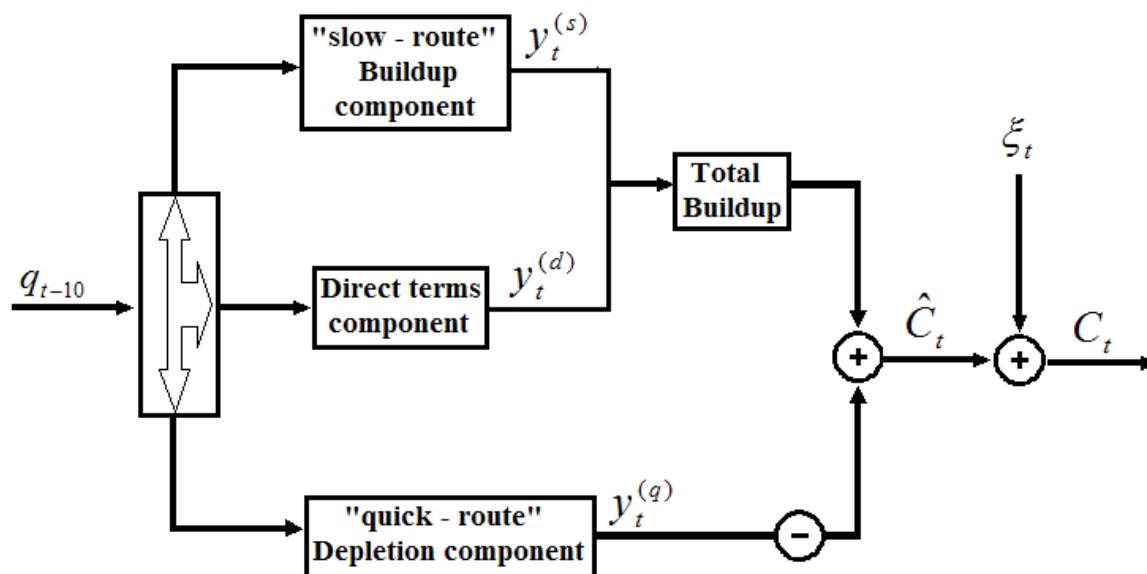


Fig. 6: Block diagram of the F-C model showing the decomposed parallel sulphate concentration flow pathways

From the foregoing, a possible physical interpretation of the model would be as follows. For the sulphate concentration build-up route, the average rainfall in a given month generates an average monthly flow as well as a deep percolation component that triggers an initial AMD production. The process would continue with subsequent rains and resulting flow. The process would be gradual given that the acid production is a much more gradual process. The deeply percolated water contributes to the recovery of the water table. The full recovery of the water table would be a gradual and continuous process depending on the hydrology of the region as well as the heterogeneous permeability of the old mine workings. In this case it would be assumed that the recovery takes an average period of 10 months. After the 10 months, the AMD is then continuously discharged into the stream via the groundwater component of measured flow at a more uniform rate for a period of 8.73 months. Hence the 10 month advective time delay in the model. The uniform rate of build-up is evident in Fig. 5 where the build-up component is seen to portray a relatively uniform pattern with minimum variations in high and low values. It is also implied from Eqs. 6, 7 and Fig. 5 that after the full recovery of the water table there exists in the system constant positive concentration values added into the stream each month and regressed in 9 months. This quantity is added into the system as a constant amount in each month but the value decreases in each successive month for a period of 9 months. A possible physical interpretation to this route would be that it is a mean trend build-up route discerned from the overall monthly build-up route. Observation of the plot of build-up component in Fig. 5 shows that this component exhibits a “sinusoidal” pattern that is uniform through out the observation period. Further, it could be observed that it is no coincidence that the regression period of 9 month is approximately equal to 8.73 months, the calculated time constant of the “slow-route build-up”. Hence it could be deduced that the model discerns from the total build-up route a mean monthly value for each month which is here presented as a direct constant term that decreases in each successive month. The observed recession makes reasonable hydrological meaning. This is because the impact of the average monthly rainfall that caused the AMD effect manifested in the stream 10 months later is expected to decay with time at some given rate over a period of time. In this case the duration is the time of concentration which is approximately 9 months. Further, it is also implied that a concurrent process with a negative sign which is inferred here as the depletion

route takes place in the stream. A possible physical interpretation to this route would be that as the AMD from groundwater is continuously injected into the system and assuming an initial low concentration in the stream, the added sulphate concentration is continually dispersed in the stream resulting into some form of dilution of the incoming flux.

The dilution effect can also be inferred from Eq. 4a where the first term in the numerator (b_0) which is a “gain” parameter is a negative value. The inferred dilution would occur in conjunction with the dilution resulting from direct precipitation. This dilution process takes place for a period of 5.06 months concurrently with the build up process. It is argued here that, the 5.06 months depletion period is continual and not continuous. A continual process occurs repeatedly and regularly with possible breaks between processes, while a continuous process occurs without interruption/gaps/breaks. Hence it is continual because hydrologically, the dilution would depend on the season as well as the flow regime in the stream.

In its present state, the model could only be used as a preliminary guide to making management decisions. It should however be noted that this model could be improved further if we had a daily time step data for the concentrations. This would have enhanced the model by allowing the incorporation of other states that affect the stream concentrations including evaporation. In its improved state, important management decisions could be made by inferring to the ensuing interpretations. For instance, with regard to the build-up route, knowing the time lag between the closure of a mine and the first AMD impact on the receiving stream and also knowing the expected duration of the impact could aid the management in setting up the right mitigation measures.

CONCLUSION

The results and the ensuing discussions have exhibited a further extension of the DBM modelling important attributes as utilised in simple flow-concentration modelling. In this study, the resulting DBM model explains 22% of the data deterministically, the rest being stochastic. It is also noted that the 22% fit is for the whole period considered from 1984 to 1999. However in the first 14 months, the model estimated approximate 100% of the data deterministically.

Despite the poor overall deterministic model accuracy, attempts were made to analyse not only the model fit but also its mechanistic interpretation of the system. It has also been observed that the DBM modelling approach has potential of discerning not only the contaminant flow pathways but also give an indication of the potential lag between the occurrence of the polluting element and the effects especially with respect to groundwater pollution resulting from non-point sources. In this case, the potential lag between the occurrence of rainfall contributing to flow and the consequent AMD production and manifestation in the stream was identified as 10 months. Three distinct pathways identified were; the concentration build-up, concentration dilution/depletion and the direct term contribution (mean trend build-up) pathways. In each case, the percent composition of each pathway was also quantified. Where the total build-up pathway (build-up + direct term) constituted 51.19% of the modelled sulphate concentration and the depletion pathway constituted 48.81% of the modelled sulphate concentration.

The resulting model has shown that it is possible to use the DBM modelling approach to help solve the problem mostly encountered in non-point source pollution modelling at the catchment scale that is, the inability of most models to represent the potential lag between polluting activity and its effect especially with respect to groundwater contamination and discharge into the surface water resources.

In terms of the non-point source pollution modelling and management, the resulting model has further shown ability to provide more salient information about the system dynamics. This kind

of information that is, the residence times and the advective time delays in the system, could prove useful for the catchment managers in making informed decisions including laying out remediation measures. It suffices to mention that these decisions can only be made once the current model has been improved. As mentioned earlier, the mechanistic interpretations presented here are for the deterministic part of the model that explained only 22% of the data. It is argued that this model could still be improved further. This as also earlier stated could be achieved by identifying possible nonlinearities in the system and allowing for the incorporation of other states that affect streamflow concentration such as evaporation.

ACKNOWLEDGEMENTS

This paper is based on partial work undertaken towards a doctorate study. We would like to thank Tshwane University of Technology for awarding a bursary to the principal author to undertake his studies. The data provided by the Department of Water Affairs and Forestry (DWAF), and Council for Scientific and Industrial Research (CSIR), South Africa are also greatly appreciated.

REFERENCES

- 1) Huang, G.H., and Xia, J. (2001), "Barriers to sustainable water-quality management", *J. Env. Mgt.*, 61, 1-23.
- 2) Singh, V.P., and Woolhiser, D.A. (2002), "Mathematical modelling of watershed hydrology", *J. Hyd. Eng.*, 7(4), 270-292.
- 3) Merritt, W.S., Letcher, R.A., and Jakeman, A.J. (2003), "A review of erosion and sediment transport models", *Env. Mod. & Software.*, 18, 761-799.
- 4) Letcher, R.A., Jakeman, A.J., Calfas, M., Linforth, S., Baginska, B and Lawrence, I. (2002), "A comparison of catchment water quality models and direct estimation techniques", *Env. Mod. & Software.*, 17, 77-85.
- 5) Pegram, G.C and Görgens, A.H.M. (2001), "A guide to non – point source assessment", WRC Report no. TT 142/01, Pretoria, South Africa.
- 6) Malan, H.L and Day, J.A. (2002), "Development of numerical methods for predicting relationships between stream flow, water quality and biotic responses in rivers", WRC Report no. 956/1/02, Pretoria, South Africa.
- 7) Bradley, S.B and Lewin, J. (1985), "Transport of heavy metals on suspended sediments under high flow conditions in a mineralised region of Wales", *J. Env. Poln.*, 4(B), 257-269.
- 8) Cox, B.A. (2003), "A review of currently available in – stream water – quality models and their applicability for simulating dissolved oxygen in lowland rivers", *J. The Sci. Tot. Env.*, (314 – 316), 335-377.
- 9) Dye, P.J and Croke, B.F.W. (2003), "Evaluation of streamflow predictions by the IHACRES rainfall-runoff model in two South African catchments", *Env. Mod. & Software.*, 18, 705-712.
- 10) Beven, K.J. (2001), "Rainfall-Runoff Modelling. The Primer", John Wiley and Sons Ltd, England.
- 11) Young, P.C and Beven, K.J. (1994), "Data-based mechanistic modelling and the rainfall-flow nonlinearity", *Environmetrics*, 5, 335-363.
- 12) Department of Water Affairs and Forestry. (2004), "National Water Resource Strategy First Edition", Republic of South Africa.
- 13) Basson, M.S., Allen, R.B., Pegram, G.G.S., and van Rooyen, J.A. (1994), "Probabilistic management of water resource and hydropower systems", Water Resources Publication, USA.

- 14) MATHWORKS. (2004) *MATLAB, the language of technical computing*. Version 7.0.0. Release 14. [CD-ROM]. *The MathWorks Inc.* (www.mathworks.com).
- 15) Pedregal, D.J., Taylor, C.J and Young, P.C. (2004), "System Identification, Time Series Analysis and Forecasting", *The Captain Toolbox Handbook v 1.1*. Centre for Research on Environmental Systems and Statistics (CRES) Lancaster University, Lancaster, LA1 4YQ, United Kingdom.
- 16) Young, P.C., Taylor, C.J., Tych, W., Pedregal, D.J and McKenna, P.G. (2004), "The Captain Toolbox", Centre for Research on Environmental Systems and Statistics (CRES), Lancaster University, UK, 2004. (www.es.lancs.ac.uk/cres/captain).
- 17) Demchak, J., Skousen, J and McDonald, L. M. (2004), "Longevity of acid discharges from underground mines located above the regional water table", *J. Env. Qlty.*, 33, 656-668.
- 18) Bilek, F. (2004), "Prediction of groundwater quality affected by acid mine drainage to accompany *in situ* remediation", *J. Appld. Earth. Sci.*, 113(1), April:31- 42.
- 19) Horner, B and Geldenhuys, H. (2005), "Toxic water shock for Gauteng: old mines could poison entire supply with heavy metals, says government report", *Sunday Times (Metro)*, Feb. 27, 1-2.