

A new look at uncertainty in end member mixing models for streamflow partitioning

Adams, G.A.¹, P.S. Cornish³, B.F.W. Croke^{1,4}, M.R. Hart³, C.E. Hughes², A.J. Jakeman⁴

¹ The Department of Mathematics, The Australian National University, Australian Capital Territory

² Institute for Environmental Research, The Australian Nuclear Science and Technology Organisation, Lucas Heights

³ School of Natural Sciences, University of Western Sydney, Hawkesbury Campus

⁴ Integrated Catchment Assessment and Management Centre (iCAM), The Fenner School of Environment and Society, The Australian National University, Australian Capital Territory

Email: geoffrey.adams@anu.edu.au

Abstract: Chemical and isotope based hydrograph separation methods, such as end member mixing analysis (EMMA¹), have been utilized in recent decades because they offer a better explanation of stream chemistry than can be obtained from flow partitions derived using graphical or mathematical filter hydrograph partitioning techniques (Chapman and Maxwell, 1996; Pinder and Jones, 1969). However, use of such constituent based hydrograph separation methods results in uncertain streamflow partitions, attributable, in part, to uncertainty in the characteristic chemical or isotope content of source waters (Figure 1) (Soulsby *et al.*, 2003). This paper explores additional causes of uncertainty when using EMMA for flow partitioning.

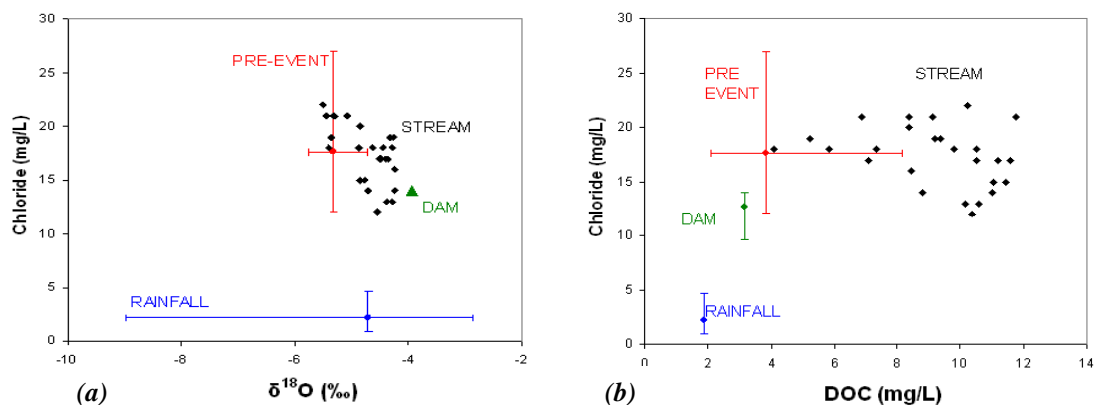


Figure 1: Mixing diagram showing the mean and standard deviations of measurements of constituent content for three end members in a NSW Southern Highlands catchment

This paper shows that:

- hydrograph partitioning based on measured end member constituent concentrations should not be assumed to be correct just because the stream chemistry is adequately explained;
- propagating uncertainty estimated from sample variations will underestimate uncertainty in flow partitions; and
- numerical problems during computation introduce considerable uncertainty into the solved end member contributions. Sampling and measurement errors exacerbate the problem.

Keywords: EMMA, end member mixing analysis, hydrograph, partition, water quality, uncertainty

¹ The term EMMA is sometimes used in the context of having more constituents than required to fully determine the mass balance equations. It is used here in the context of any streamflow mass balance.

1. INTRODUCTION

The conceptual framework of a new parsimonious nutrient mobilisation and export model currently under development requires that streamflow for small hillslope catchments be partitioned into surface, sub-surface and base flow components. This is because the sub-surface and surface flow paths have very different phosphorus (P) mobilisation and transport characteristics, and soil at different depths has different P content and buffering characteristics. This paper briefly reviews the uncertainties inherent in the use of chemical and isotope mass balance techniques, such as end member mixing analysis (EMMA), to partition stream flow according to flow path or source. The paper also demonstrates that uncertainties in mass balance partitioning of streamflow into partitions such as surface runoff, sub surface runoff and base flow (sourced from groundwater) may be greater than previously recognised.

Early work on partitioning stream flow according to source partitions was, in effect, based on response time – for example, graphical analysis of hydrographs or mathematical filtering (Chapman and Maxwell, 1996; Grayson *et al.*, 1996; Nathan and McMahon, 1990; Pilgrim and Cordery, 1993). It was recognised however that flow separation based on response time does not necessarily accord with the actual flow paths, nor distinguish “old water” from “new water” (Chapman and Maxwell, 1996; Pinder and Jones, 1969), with Vanni *et al.* (2001) observing that most (hydrograph) techniques for estimating baseflow are “somewhat” arbitrary. Earlier, Appleby (1970) had described (response time based) baseflow separation as “that fascinating arena of fancy and speculation ...”.

As far back as 1969, Pinder and Jones (1969) estimated the baseflow contribution to runoff by mass balance of dissolved ions, having inferred the chemical signatures of overland runoff and subsurface water from sampling discharge at high and low flows. Hooper *et al.* (1990) used a mass balance approach, EMMA, to investigate hydrograph partitioning. Other constituent based techniques for hydrograph partitioning utilise the additional information available from variations in environmental isotope concentrations in catchment, rain and discharged water (Clark and Fritz, 1997; Machavaram *et al.*, 2006).

The most basic hydrograph partitioning using chemical tracing uses one constituent to enable the partitioning of the discharge hydrograph into two partitions. If concentration data for a second constituent, which exhibits suitably different characteristics to the first constituent, is available, then the discharge hydrograph can be partitioned into three source partitions (Clark and Fritz, 1997). It is often found that three partitions are necessary to adequately explain the chemical signature of the receiving waters (Jung *et al.*, 2009). The characteristic concentration of each constituent in each source partition, or *end member*, is generally assumed to have been identified from analysis of water samples from the end members. The mass balance equations governing flow rate mass balance and constituent load mass balance used in the EMMA, based for instance on Clark and Fritz (1997), are¹:

$$Q_t = Q_s + Q_{ss} + Q_{bf} \quad (1)$$

$$Q_t \cdot C_{1t} = Q_s \cdot C_{1s} + Q_{ss} \cdot C_{1ss} + Q_{bf} \cdot C_{1bf} \quad (2)$$

and

$$Q_t \cdot C_{2t} = Q_s \cdot C_{2s} + Q_{ss} \cdot C_{2ss} + Q_{bf} \cdot C_{2bf} \quad (3)$$

where;

- the numerical subscripts refer to the constituents,
- subscript *s* refers to flow path 1 (surface flow), subscript *ss* refers to flow path 2 (sub-surface flow), subscript *bf* refers to an additional sub-surface flow path (e.g. base flow from groundwater), and subscript *t* refers to the total discharge rate (i.e. stream flow rate),
- *Q* is discharge rate,
- *C* is constituent concentration (or a proxy, for example electrical conductivity).

With total stream discharge rate measured, and the total stream constituent concentrations and end member constituent concentrations known, the set of three equations has three unknowns: Q_s , Q_{ss} , and Q_{bf} . With

¹ Other researchers use equivalent equations using dimensionless flow proportions, rather than discharge rate.

suitable input data, apparently satisfactory explicit solutions can be found for these unknown flow rates, and the chemical behaviour of the stream is thus apparently explained.

In such an approach, the conceptualized end members are implicitly considered to be the sole sources of stream water. Further, for the purpose of solving the equations, the identified concentrations of each constituent in each end member are assumed to perfectly characterise the concentration of the constituents in water sourced from each end member, across space and time. (Uncertainty in end member characteristics is discussed in the next section.) The use of mixing diagrams¹ enables the viewer to readily see if the stream chemistry response can be explained by some combination of flow rates from the selected end members. If so, the stream constituent concentrations will lie within the area bounded by the end member concentrations (Christophersen *et al.*, 1990), thus in Figure 1a the observed stream Cl concentration and $\delta^{18}\text{O}$ can be explained by flow from the selected end members. If however the mixing diagram shows that the stream chemistry is outside the area bounded by the chemistry of the end members, then we cannot satisfactorily explain the stream chemistry from any mixture of the assumed stream flow sources (as in Figure 1b). The discrepancy may be due to use of unrepresentative end members, estimation of uncharacteristic end member chemistry or use of constituents that do not behave conservatively. The differences between the stream chemistry, and that of the identified end members, may assist the targeting of additional field work aimed at identifying flow paths, as in Christophersen *et al.* (1990). For the case in which the stream chemistry lies within the area bounded by the chemistry of the end members, we may estimate the proportional contributions of each end member to the total stream flow by solving the mass balance equations.

2. QUANTIFYING THE UNCERTAINTY IN HYDROGRAPH PARTITIONING

There is, however, a fundamental cause of uncertainty with this widely used approach to streamflow partitioning, related to how well the adopted characteristic concentrations of the hypothesised water sources represent the catchment-wide end members. The identified end member chemistry may typically be based on a number of point measurements, in space and time, within the catchment. It is however widely recognised in other areas of hydrology that point measurements of various catchment characteristics do not necessarily represent the overall catchment characteristic due to problems of scale and heterogeneity (e.g. Beven, 1989; Grayson *et al.*, 1995; Sobieraja *et al.*, 2004).

To exemplify the difficulty in usefully characterising end member chemical signatures, consider the base flow from a groundwater end member from an intensive study in a water supply catchment in the Southern Highlands, NSW². It might seem reasonable to assume that under dry, low flow, conditions, all flow is

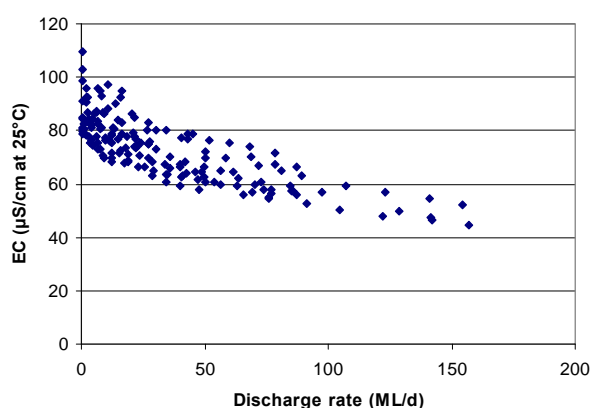


Figure 2. Electrical conductivity versus flow rate in a 1km² NSW Southern Highlands catchment

sourced from the groundwater end member, and that the stream chemistry represents an integration of the groundwater chemistry for the whole catchment. Yet Figure 2, a graph of electrical conductivity (EC)³ versus flow rate for one of six sites in the study, shows that there is no one characteristic EC at low flows. This is likely due to the heterogeneous lithology, and changing contributions to base flow from different sub-areas with time. The other five sites of the study showed similar variability in base flow characteristics for EC, and the concentrations of Cl⁻ and Ca²⁺ ions.

Some researchers have estimated the effect of uncertainty of end member characteristics on the estimates of stream

¹ A mixing diagram is an X-Y plot comparing the concentrations of the two constituents for each of the conceptual *end members*, and the stream.

² The study is an ARC funded linkage project involving the Sydney Catchment Authority, the University of Western Sydney, and the Australian National University.

³ EC is used here to represent the soluble constituents of water.

flow contribution from each end member. However, as discussed in the next section, these approaches have only considered the uncertainty that has been identified due to variation in sample concentrations, not uncertainty that has not been identified. For example, Soulsby *et al.* (2003) used a “Bayesian Markov Chain – Monte Carlo” approach to quantify the uncertainty in estimated source contributions due to variation in the concentrations of chemicals in end member samples. They estimated that the contribution of the groundwater end member to stream flow was between 4 and 25 % of total flow. Rice and Hornberger (1998) applied Monte Carlo sampling with assumed, arbitrary, sampling errors to indicate the uncertainty in hydrograph partitioning. While acknowledging the difficulty in adequately defining catchment-wide characteristics of the solute constituents of selected end members (e.g. use of chemistry of single wells to characterise soil water or groundwater end members for the whole catchment being studied), Hooper *et al.* (1990) presented a formulation for a first order approximation of the effects of uncertainty of end member concentrations on estimated source contributions to the total hydrograph. Other researchers such as Genereux (1998) and Uhlenbrook and Hoeg (2003) have summarised previous work in estimating partitioning uncertainty, and applied Gaussian error propagation techniques for estimating partitioning uncertainty. All of these uncertainty estimation methods have relied on estimates or determinations of sampling or measurement error.

Various techniques have been used to ensure the validity of EMMA. For example, Bernal *et al.* (2006) restricted the EMMA to data sets in which the stream chemistry was totally bounded by the presumed end member chemistry, then determined the goodness of the EMMA model by comparing predicted and recorded stream concentrations. It is demonstrated below however that such tests are not conclusive, as good agreement between stream and modelled concentrations can be obtained from a poor (incorrect) model.

3. ADDITIONAL CAUSES OF UNCERTAINTY

Uncertainty in the estimation of end member contributions to total stream flow is not only affected by how precisely the end member chemistry is characterised, but also by the precision and accuracy of the stream chemistry measurements. For example, Soulsby *et al.* (2003) estimated by mass balance that groundwater contributed 6 % of total flow in a stream. The calculated contribution could however be as much as 17 % if the measured stream chemical concentrations were adjusted by typical measurement error (Adams, in prep.).

A more fundamental source of uncertainty in mass balance approaches to streamflow partitioning may arise from the difficulty in reliably characterising end member properties, even where the end members are appropriately identified. The use of a relatively small set of sample results to characterise the range of end member properties is likely to result in underestimation of the true uncertainty in the properties. This can then lead to unidentifiable uncertainty in streamflow partitioning, as illustrated later in this section.

An indication of the problems that can occur with EMMA is the different partitionings that can result when different constituents are used in the EMMA. This is illustrated in the two mixing diagrams in Figure 1, which show the same three end members, but two different pairs of constituents for the NSW Southern Highlands catchment referred to above. Figure 1a shows the selected end members can provide a feasible explanation of the observed stream chemistry. Yet in Figure 1b, with data from the same period, it is apparent that the same selected end members cannot explain the stream chemistry, and there must therefore be an unidentified source of flow to the stream. To a certain extent Christophersen *et al.* (1990) avoided exposure to this problem by solving an over-determined system for all constituents simultaneously (through minimising a least squares objective function). However their problem manifested itself in a different way, with the EMMA model failing to predict stream concentrations for constituents for which the stream chemistry was bounded by the end members (Hooper *et al.*, 1990). Other researchers have compared the results of three end member, two constituent partitioning for different pairs of constituents, with Rice and Hornberger (1998) concluding that the use of different pairs of tracers can result in inconsistent partitioning. Scarff (2008), when using isotope tracers to partition streamflow in the 1 km² Southern Highlands catchment referred to above, found the estimated proportion of pre-event water varied from 16 % of total discharge to 56 % of total discharge in a two component partitioning, depending on how an environmental isotope ($\delta^2\text{H}$) was characterised in the event-water end member. In three component partitioning, different pairs of tracers resulted in inconsistent partitioning, with tracers $\delta^2\text{H}$ and $\delta^{18}\text{O}$ giving an event water contribution of ~50 %, compared with tracers Cl and $\delta^{18}\text{O}$, which gave an event water contribution of only 13 %. This is due in part to uncertainty in end member characterisation of pre-event waters and also demonstrates that when characterising an evaporated end member (farm dams), combining $\delta^2\text{H}$ with $\delta^{18}\text{O}$ provides valuable information concerning the age of discharged water.

The following discussion demonstrates that while it should be possible to infer end member characteristics from stream flow and concentration data, numerical problems thwart this. It is then shown that the same

numerical problems add uncertainty to EMMA. As discussed, inability to accurately characterise the constituent concentrations in the nominated end members is a significant source of uncertainty in flow partitioning. If the appropriate end member characteristics could be inferred from the stream constituent concentration data, rather than deriving them from analysis of a small number of end member samples, then this uncertainty could be avoided. A consequence of such an approach would be that the inferred “*end members*”¹ may not be explicitly identified with a particular flow path or source. However, as is discussed in Adams (in prep.), the definitions of the end members are, in practice, rather arbitrary.

Equations 1 to 3 provide three independent equations with three unknowns and, hence, are a fully determined set of equations, with no scope to uniquely solve additional unknowns such as end member constituent concentrations. However, if it is considered that the end member constituent concentrations are constant with time, then additional equations can be derived for additional samples at different times and discharges, as summarised in Table 1. Table 1 shows that flow contributions and characteristic constituent concentrations for two partitions could potentially be identified using total stream discharge and total stream constituent concentration data for two constituents from four different samples.

Table 1: Use of data from four samples to solve for flow partitions and partition constituent concentrations – solving for two partitions using two constituents

Sample	Unknowns	No. of unknowns	No. of independent equations*
1	$Q_s, Q_{ss}, C_{1s}, C_{1ss}, C_{2s}, C_{2ss}$	6	3
2	As for sample 1, plus for sample 2 Q_s, Q_{ss}	8 (including for sample 1)	6 (including for sample 1)
3	As for sample 2, plus for sample 3 Q_s, Q_{ss}	10 (including for previous samples)	9 (including for previous samples)
4	As for sample 3, plus for sample 4 Q_s, Q_{ss}	12 (including for previous samples)	12 (including for previous samples)

* Based on Equation 1 to Equation 3, omitting the base flow terms

A third, base flow partition, could be identified on the same basis if base flow constituent concentrations were determined *a priori* from low flow (i.e. base flow) data, and a corresponding flow rate identified, below which all flow is assumed to be from the third (groundwater) end member, and above which the base flow rate is subtracted before other end member flow contributions are determined. Hence the unknown concentrations and flow contributions from each partition can be identified by solving the 12 equations², if the end member concentrations are assumed to remain constant over four differing samples. This assumption will not generally be strictly true in practice – however it seems more plausible than the standard EMMA assumption that spot concentration measurements (both spatial and temporal) are sufficient to generally characterise the end member constituent concentrations. (Recognising that the effective end member concentrations will in fact vary with time, the concentrations C1s, C1ss, C2s and C2ss could be solved for the second or third time/sample, rather than the first time/sample, to minimize the maximum time difference between the solution time and sample times.) If concentration data for a third, suitably different, constituent are available, then, provided the base flow threshold flow rate and concentrations are determined *a priori*, the total flow can be partitioned into three partitions using just three samples/times. It turns out that the problem will typically be ill-posed, with many different combinations of end member concentration and flow partitioning yielding apparently perfect solutions.

This is best demonstrated using error-free synthetic data, as in the example

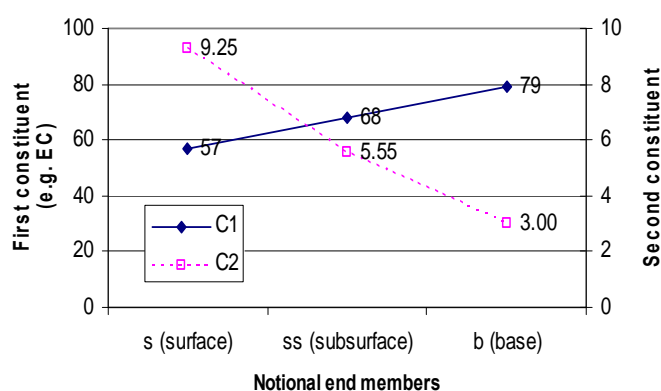


Figure 3: Assumed end member constituent concentrations for the synthetic example

¹ The emphasis is to highlight that the end members are now notional, rather than conceptual.

² The majority of the 12 equations are non linear, and so an optimisation method is required to solve them.

below. The end member constituent concentrations (Figure 3) were selected to provide distinct chemical signatures in the different notional end members. The synthetic data (Table 2) derived using those concentrations and randomly generated flow data, are error free, and necessarily determined to a much higher precision than is realistic for real data, as explained below.

Table 3 shows some of the end member concentrations and flow partitioning that could account for the synthetic set of data shown in Table 2 (Adams, in prep.). If either, errors are present in the data (in practice always the case), a realistic precision applied to the data, or the end member concentrations change with time, the results of the flow partitioning will be significantly worse than for the ideal case. Indeed, it is unlikely in any of these situations that the EMMA will identify the “correct” solution. The example below assumes that the stated end member concentrations apply for all four samples, enabling constituent concentrations in three end members, and separate end member contributions to total stream flow for the four samples, to be solved.

In this synthetic example, the “observed” stream chemistry could be explained by flow rates in the first partition ranging from 1.8 to 3.2, and in the second partition ranging from 0.7 to 2.2 (arbitrary units).

The data are presented at such high precision for two purposes: the high precision allows readers to confirm that the solutions presented in Table 3 could each explain the “observations” at timestep 2 in Table 2; and, it

Table 2: Randomly selected synthetic flow, and corresponding concentration data derived from the concentration data defined in Figure 3 (arbitrary units)

Timestep/sample	Total stream flow* [†] (Q _t)	First constituent stream concentration* (C _{1i})	Second constituent stream concentration* (C _{2i})
1	9.06695423	65.32720463	6.5017942
2	4.36138076	63.60362760	7.1384698
3	14.92431670	60.13170859	8.2286622
4	1.51814833	69.47685626	5.3683600

*The data need to be passed to the solver at high precision to enable the correct solution to be included in the set of feasible solutions. [†]Q_b fixed at 0.416.

allows the reader to confirm that if the presented data are rounded down to realistic measurement precision (say one significant figure in this example), the “correct” concentrations and flows (middle row Table 3) no longer best explain the “observations” in Table 2.

Table 3: Examples of possible solutions which adequately explain the synthetic “observations” in Table 2.

C1s	C1ss	C2s	C2ss	Qs*	Qss*
57	66	9.25	6.222727	1.762166	2.183215
57	68	9.25	5.550000	2.159114	1.786267
57	84	9.25	0.168182	3.217642	0.727738

* Qs and Qss are given for the second sample of Table 2.

[†]The concentrations in the middle row (bold) were shown in Figure 3, and were used to derive all synthetic results in Table 2.

4. CONCLUSION

It is shown that “correct” end member constituent concentrations cannot be accurately inferred from stream concentration data or, indeed, from error free artificial data, although in practice, field measurements can be used to constrain the range of inferred end member concentrations to some extent. A corollary of this is that hydrograph partitioning based on end member concentrations estimated from analysis of spot samples cannot be assumed to be valid simply because it appears to explain the stream chemistry – a multitude of different partitioning solutions could explain the observed stream chemistry time series, if uncertainty in the end member characterisations is recognised. Numerical error during computation increases the uncertainty in the end result.

A further outcome of these results is that estimated uncertainty in hydrograph partitioning based on propagating measurement or sampling errors is likely to be underestimated. This is because the true uncertainty in the end member characteristics, in space and time, cannot generally be identified with a small set of samples.

ACKNOWLEDGMENTS

This research, into a new nutrient mobilisation and export model, is funded by the Australian Research Council, in conjunction with the Sydney Catchment Authority, the University of Western Sydney and the Australian National University. We thank the two anonymous reviewers for their helpful comments.

REFERENCES

- Adams, G.A., in prep. Development of a simple hydrological modelling methodology for estimating nutrient generation from rural land. PhD Thesis, Australian National University, Canberra.
- Appleby, F.V., 1970. Recession and baseflow problem. *Water Resour. Res.*, 6(5): 1398-1403.
- Bernal, S., Butturini, A. and Sabater, F., 2006. Inferring nitrate sources through end member mixing analysis in an intermittent Mediterranean stream. *Biogeochemistry*, 81(3): 269-289.
- Beven, K., 1989. Changing ideas in hydrology -- The case of physically-based models. *J. Hydrol.*, 105(1-2): 157-172.
- Chapman, T.G. and Maxwell, A.I., 1996. Baseflow separation - comparison of numerical methods with tracer experiments, *Hydrology and Water Resources Symposium*. Institution of Engineers, Australia, Hobart, pp. 539-545.
- Christophersen, N., Neal, C., Hooper, R.P., Vogt, R.D. and Andersen, S., 1990. Modelling streamwater chemistry as a mixture of soilwater end-members -- A step towards second-generation acidification models. *J. Hydrol.*, 116(1-4): 307-320.
- Clark, I.D. and Fritz, P., 1997. *Environmental Isotopes in Hydrogeology*. Lewis Publishers, New York, 328 pp.
- Genereux, D., 1998. Quantifying uncertainty in tracer-based hydrograph separations. *Water Resour. Res.*, 34(4): 915-919.
- Grayson, R.B., Argent, R.M., Nathan, R.J., McMahon, T.A. and Mein, R.G., 1996. *Hydrological recipes: Estimation techniques in Australian hydrology*. Cooperative Research Centre for Catchment Hydrology, 125 pp.
- Grayson, R.B., Blöschl, G. and Moore, I.D., 1995. Distributed parameter modelling using vector elevation data: Thales and TAPES-C. In: V.P. Singh (Editor), *Computer models of watershed hydrology*. Water Resources Publications, Colorado, pp. 669-696.
- Hooper, R.P., Christophersen, N. and Peters, N.E., 1990. Modelling streamwater chemistry as a mixture of soilwater end-members -- An application to the Panola Mountain catchment, Georgia, U.S.A. *J. Hydrol.*, 116(1-4): 321-343.
- Jung, H.Y., Hogue, T.S., Rademacher, L.K. and Meixner, T., 2009. Impact of wildfire on source water contributions in Devil Creek, CA: evidence from end-member mixing analysis. *Hydrol. Process.*, 23(2): 183-200.
- Machavaram, M.V., Whittlemore, D.O., Conrad, M.E. and Miller, N.L., 2006. Precipitation induced stream flow: An event based chemical and isotopic study of a small stream in the Great Plains region of the USA. *J. Hydrol.*, 330(3-4): 470-480.
- Nathan, R.J. and McMahon, T.A., 1990. Evaluation of automated techniques for base flow and recession analyses. *Water Resour. Res.*, 26(7): 1465-1473.
- Pilgrim, D.H. and Cordery, I., 1993. Chapter 9: Flood runoff. In: D.R. Maidment (Editor), *Handbook of hydrology*. McGraw-Hill Inc., New York, pp. 9.1-9.42.
- Pinder, F.G. and Jones, J.F., 1969. Determination of the groundwater component of peak discharge from the chemistry of total runoff. *Water Resour. Res.*, 5(2): 438-445.
- Rice K. C. and Hornberger G. M., 1998. Comparison of hydrochemical tracers to estimate source contributions to peak flow in a small, forested, headwater catchment. *Water Resour. Res.*, 34: 1755-1766.
- Scarff, S., 2008. *The Application of Isotope and Chemical Tracers in the Determination of Flow Pathways of Water in a Small Agricultural Catchment in the Southern Highlands, NSW, Australia*. Honours Bachelor of Environmental Science Thesis, University of Wollongong, Wollongong.
- Sobieraja, J.A., Elsenbeerb, H. and Cameron, G., 2004. Scale dependency in spatial patterns of saturated hydraulic conductivity *Catena*, 55(1): 49-77.
- Soulsby, C. *et al.*, 2003. Identifying and assessing uncertainty in hydrological pathways: a novel approach to end member mixing in a Scottish agricultural catchment. *J. Hydrol.*, 274(1-4): 109-128.
- Uhlenbrook, S. and Hoeg, S., 2003. Quantifying uncertainties in tracer-based hydrograph separations: a case study for two-, three- and five-component hydrograph separations in a mountainous catchment. *Hydrol. Process.*, 17(2): 431-453.
- Vanni, M.J., Renwick, W.H., Headworth, J.L., Auch, J.D. and Schaus, M.H., 2001. Dissolved and particulate nutrient flux from three adjacent agricultural watersheds: A five-year study. *Biogeochemistry*, 54: 85-114.