

Using factor analysis and end-member mixing techniques to infer sources of runoff generation

CORINNA ABESSER^{1,*}, RUTH ROBINSON¹ & CHRIS SOULSBY²

¹ *Department of Geography and Geosciences, University of St Andrews, St Andrews KY16 9AL, UK*
c.abesser@bgs.ac.uk

² *Department of Geography and Environment, University of Aberdeen, Aberdeen AB24 3UF, UK*

** now at: British Geological Survey, Maclean Building, Crowmarsh Gifford, Wallingford, OX10 8BB, UK*

Abstract Factor analysis and (two-component) end member mixing analysis (EMMA) were applied to high resolution stream chemistry data from three catchments to infer sources of iron- and manganese-rich runoff in the catchment area and to evaluate their temporal and spatial influence on the streamwater quality. Two Fe and Mn sources were identified: (a) an organic soilwater source, associated with Fe and Mn accumulation in the organic-rich upper soil horizons; and (b) a deep soilwater/groundwater source arising from reduced metal mobilization in the deeper soils. While: (a) provides the dominant Fe and Mn inputs during storm events; (b) becomes important just before peak flow when riparian groundwater is displaced into the stream, giving rise to high Fe and Mn concentrations. The timing of such high Fe and Mn loadings in the stream runoff has important implications for the water supply management of the catchment.

Key words EMMA; factor analysis; hydrology; iron; manganese; Scotland, UK

INTRODUCTION

Upland catchments are the major sources of potable, industrial and agricultural water supplies, particularly in Scotland where 96% (by volume) of the public water supplies are derived from surface waters (MacDonald, 1994). The increasing environmental pressures on upland water quality arising from: (a) acidification (Kirchner & Lydersen, 1995); (b) high sediment loads (Carling *et al.*, 2001); and (c) mobilization of colour and metals (Heal *et al.*, 2002) have provided a major impetus for water quality research in the UK upland regions during the last few decades (Neal *et al.*, 1997b; Langan & Hurst, 2004). While iron and manganese have generally been included in these studies, the main focus of the research has been on the impact of acid deposition and aluminium mobilization and comparatively little work has been dedicated to water quality problems related to the occurrence of (naturally) high loadings of Fe and Mn. In recent years, the increased presence of Fe and Mn in stream runoff and water bodies has become a growing concern to many water authorities in the UK (Little & McFadzean, 1991; Schofield *et al.*, 1991) and the EC Drinking Water Directive (CEC Commission of European Communities, 1980) has set the mandatory maximum admissible concentration of Mn and Fe in drinking water at 50 $\mu\text{g L}^{-1}$ and 200 $\mu\text{g L}^{-1}$, respectively. Temporary deterioration in water quality due to high Mn levels has

occurred in the raw waters of Loch Bradan, southwest Scotland (Little & McFadzean, 1991), and Fe and Mn problems were also experienced in the Megget Reservoir during an excessive draw down of the water level in 1997/98 (Abesser, 2003). While it is well-established that catchment hydrology is a dominant control on streamwater quality and metal loading in such upland regions, relatively little is known about the mechanisms and pathways by which these elements are delivered to streams.

In this paper, the statistical techniques of *R*- and *Q*-factor analysis in combination with the relatively simple technique of End Member Mixing Analysis (EMMA) are applied to high-frequency storm event data from three streams in the Megget catchment of Scotland with the aim to: (a) gain a better understanding of the underlying mechanisms of runoff generation in the Megget catchment during storm events; and thus (b) identify the sources of Fe and Mn in the catchment area; and (c) evaluate their temporal and spatial influence on the streamwater quality.

STUDY AREA

The Megget catchment is located in the Scottish Uplands, about 50 km south of Edinburgh (NGR 319 622). It covers an area of about 45 km² and holds one of Scottish Water's largest supply reservoirs. Smoothed hillsides, moraine hummocks and ridges form the sides of the U-shaped Megget valley with altitudes ranging between 330 m and 847 m above sea level. Temperatures in the catchment area range between -10°C and 29°C and the average annual rainfall is 1524 mm with higher rainfalls occurring in autumn and winter. The geology of the catchment is comprised of intensively folded and fractured Silurian greywackes, siltstones and shales, which are overlain by glacial and post-glacial (alluvial) deposits. The most abundant soil types are peaty podzols on the steep slopes and blanket peats of 0.5–1 m thickness on the high grounds. Peaty gleys are locally developed on impermeable drift and/or alluvial deposits. The catchment displays a relatively flashy hydrological runoff regime although lag times of around 9 h have been observed for the largest stream, the Megget Burn (Grzybowski, 2001). Groundwater storage and transport in the area are mostly confined to the shallow, weathered zone and to the superficial glacial and post-glacial deposits (Robins, 2002). Semi-natural moorland and grassland communities form the dominant vegetation of the steep hill slopes and the dominant agricultural activity in the area is sheep farming.

METHODOLOGY

Event-based sampling was carried out on three streams in the catchment area over a two-week period in October 2000. Samples were collected at 2 to 4 hourly intervals and analysed for pH, colour, alkalinity, conductivity (SEC), turbidity, major ions (Mg, Ca, SO₄, NO₃), acid available and total metals (Al, Fe, Mn, Zn) and SiO₂. Hydrological conditions were monitored at the Winterhope Burn site where stream level changes were recorded in 30-min intervals.

Simultaneous *R*- and *Q*-mode factor analysis was performed for each stream using the complete, log-normalised and standardized data set. For Fe, Mn and Al, only the

total concentrations were included in the analysis in order to minimize effects of data redundancy. The analysis was carried out in Minitab following a procedure proposed by Walden & Smith (1995), as detailed in Abesser (2003). The main objective of factor analysis is to reduce the dimensionality of the original data by looking for underlying trends inherent in the data. This is done by transforming the original set of variables (e.g. chemical parameters) or samples (e.g. all streamwater samples collected during the sampling programme) into a smaller set of mutually uncorrelated “factors”, which contain the essential information of the original set of variables (or samples). The contribution of the original variables (or samples) to a particular factor can then be inferred from the resulting matrix of “factor loadings”. In the *R*-mode technique, which is used to investigate the intercorrelations among the input variables (e.g. alkalinity, Ca, pH), the size of the resulting factor loadings is related to the amount of variance contributed by a *variable* to a particular factor. In the *Q*-mode technique, which explores the relationships among samples (e.g. sample 1, sample 2, ...), the resulting factor loading is related to the amount of variance contributed by a *sample* to a particular factor.

Simple two-component End-Member Mixing Analysis (EMMA) was carried out on the data, using alkalinity as the conservative tracer (Soulsby *et al.*, 2003), and the hydrograph split was established according to the classical two-component mixing model (Christophersen & Neal, 1990). The alkalinity of the groundwater end-member was inferred from the three lowest flow events at each particular stream, **Error! Reference source not found.** while the soilwater end-member was defined based on a range of literature values (Table 1) from sites with similar soil and catchment characteristics. Following the hydrograph separation, streamwater time series for the three sites were constructed based on recorded and modelled flow data.

Table 1 End-member alkalinities ($\mu\text{Eq L}^{-1}$).

	Groundwater end-member*	Soilwater end-member**	
Winterhope Burn	269	-75	-50
Peaty Muckle Burn	269	-75	-50
Shielhope Burn	389	-75	-50

(*calculated from alkalinities at three lowest flows, **estimated from literature values in Ferrier *et al.* (1990), Soulsby *et al.* (2003); Soulsby & Dunn (2003); Soulsby (unpublished data)).

RESULTS

R- and *Q*-mode factor analysis

Following the rule of accepting all components with Eigenvalues >1 meant that two factors were selected from the *R*-mode analysis, each of which explains greater variance than the original standardized variables and together explain 85–90% of the variance in the original data (Table 2).

Table 2 The first two factors selected from *R*-mode analysis for the three streams.

Elements	Winterhope Burn		Shielhope Burn		Peaty Muckle Burn	
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2
Colour	0.765	-0.406	0.969	0.104	0.912	-0.192
Conductivity	-0.909	-0.255	-0.971	-0.174	-0.839	-0.157
pH	-0.909	-0.191	-0.872	-0.246	-0.772	0.009
Alkalinity	-0.945	-0.238	-0.841	0.019	-0.854	-0.140
Mn	0.687	-0.571	0.731	-0.660	0.487	-0.848
Fe	0.448	-0.860	0.871	-0.422	0.550	-0.771
Al	0.947	-0.149	0.917	-0.247	0.941	-0.214
Ca	-0.807	-0.239	-0.938	-0.192	-0.721	-0.590
Mg	-0.733	-0.419	-0.962	-0.152	-0.899	-0.330
SiO ₂	-0.966	-0.188	-0.914	-0.151	-0.926	-0.320
SO ₄	-0.953	0.086	-0.952	-0.152	-0.933	0.0515
Eigenvalues	7.728	1.687	9.032	0.882	7.343	2.001
Variance explained (%)	70	15	82	8	67	18
Total variance		85%		90%		85%

***R*-mode analysis—variable factor loading**

The first factor is by far the most dominant factor in all streams accounting for 67% to 82% of the variability and divides the original variables into two distinct groups with high negative loadings on pH, alkalinity and base cations and high positive loadings on Al and colour, Fe and Mn (Table 2). The high positive loadings on Al, colour (and TOC), Fe and Mn reflect the strong influence of the soilwater component and decomposition in the upper horizons. During rainfall events, stream runoff is derived mainly from overland flow and subsurface storm flows generated in the upper soil horizons. Silica, alkalinity and base cations, indicative of bedrock weathering, have high negative loadings on factor 1 which is consistent with the dilution of base-rich groundwater and deeper mineral soils during storm periods. Hence, factor 1 reflects flow related changes in stream chemistry and is interpreted as the dilution of stream baseflow when soilwater is displaced by rainfall infiltration during storm events.

The loadings on the second factor are highest for Fe and Mn in all streams and capture between 8% and 18% of the variability within the original data sets. It is interpreted as the mobilization of Fe and Mn at low pH within the soil and/or related to reductive solution of Fe and Mn oxyhydroxides in the soils under low redox conditions; the importance of such processes on stream chemistry has been widely recognized (Reid *et al.*, 1981; Heal *et al.*, 2002).

***Q*-mode analysis—Sample factor loadings**

Water samples from all streams show a similar spread of loading values on factor 1 and factor 2 (Fig. 1). The temporal changes in these loadings and their relation to flow are considered in Fig. 2 and Fig. 3. From Fig. 2 it is clear that factor 1 is closely related to flow, reinforcing the importance of flow path changes during storm events and the

increasing influence of inputs from the upper, organic-rich soil horizons on streamflow composition. The importance of changes in flow path during events is also reflected in the factor plots of Fig. 4 for a single storm event. In this graph, *R*- and *Q*-mode factor loadings are plotted in the same two-dimensional factor space, which allows the sample loadings to be explained by reference to the position of the original variables. Hence, samples with high negative loading on factor 1 indicate a runoff composition dominated by high alkalinity, SiO₂ and base cation ion concentrations, while samples with high positive loadings reflect a high proportion of Al and colour in the stream runoff. From Fig. 4 it is readily apparent that samples have high negative

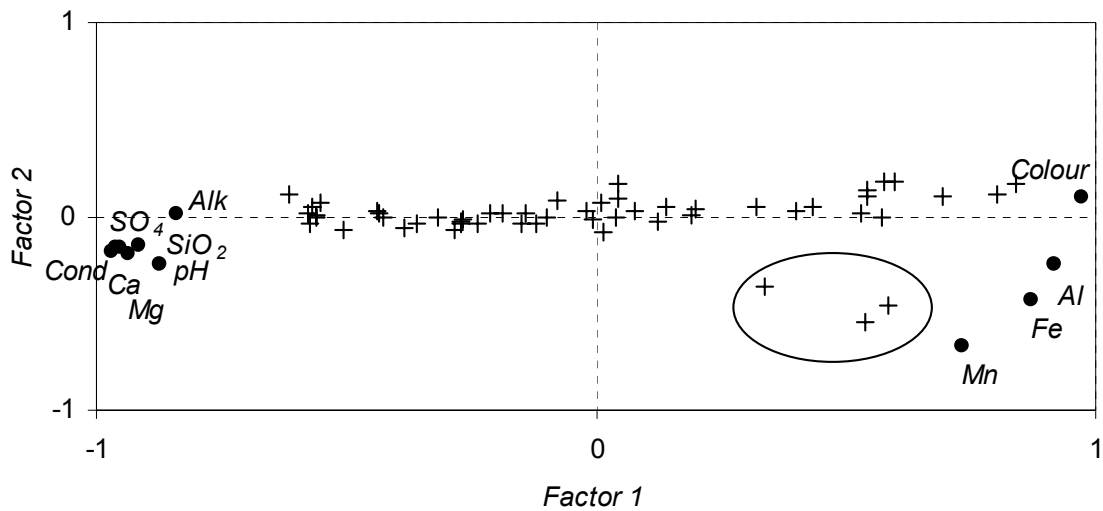


Fig. 1 Sample factor loadings for Shielhope Burn with high-flow events encircled. Plus signs represent each sample.

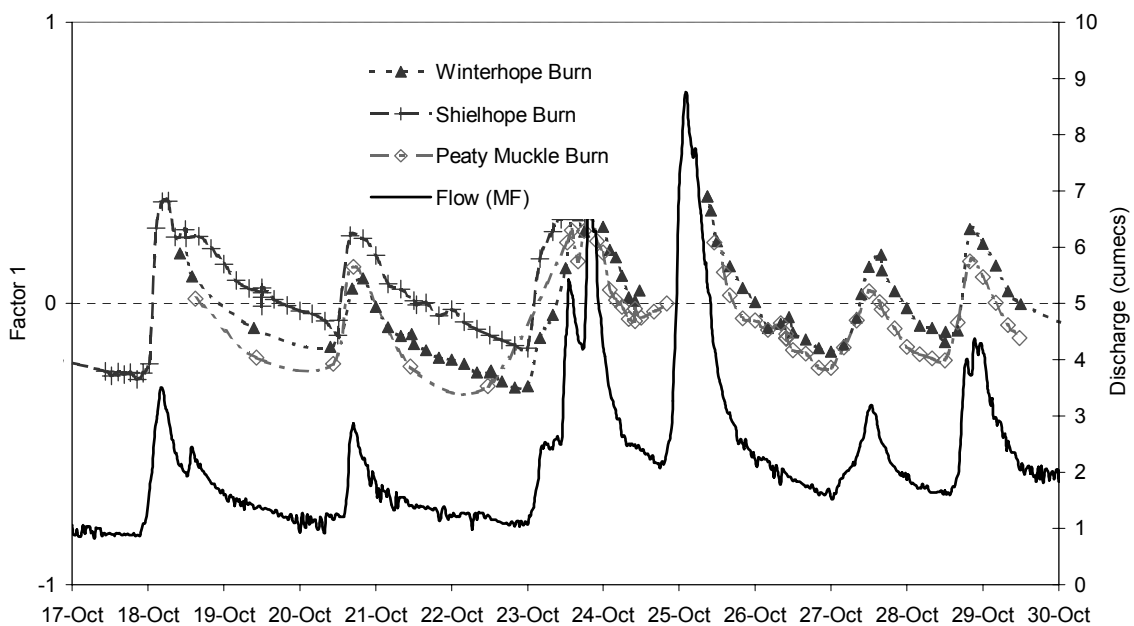


Fig. 2 Sample loadings of all streams on factor 1 and their relation to flow.

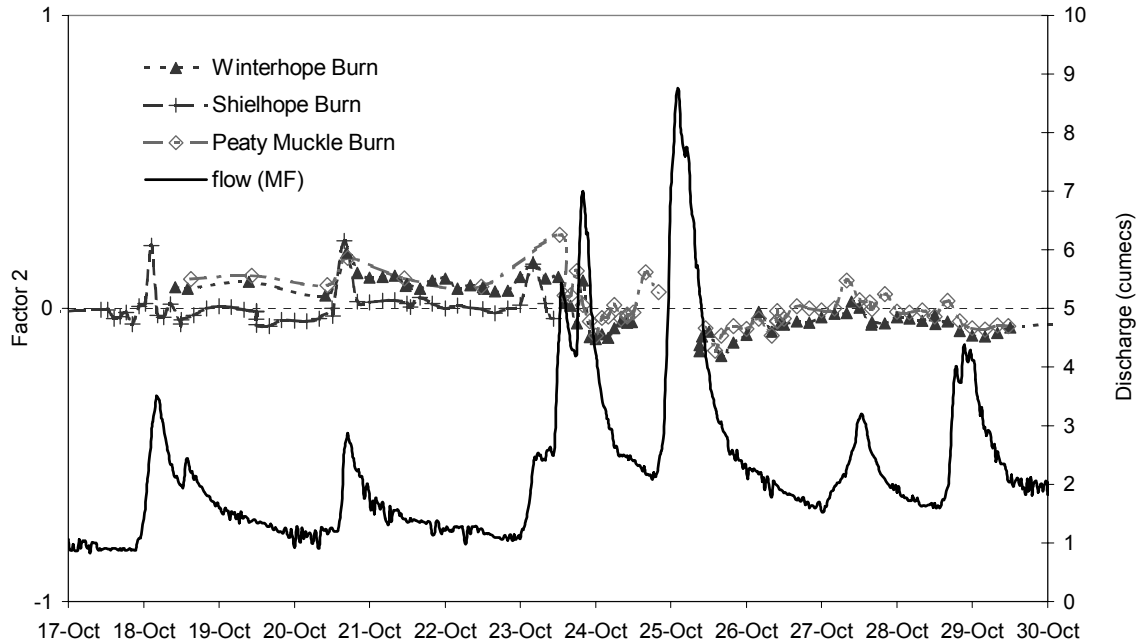


Fig. 3 Sample loadings of all streams on factor 2 and their relation to flow.

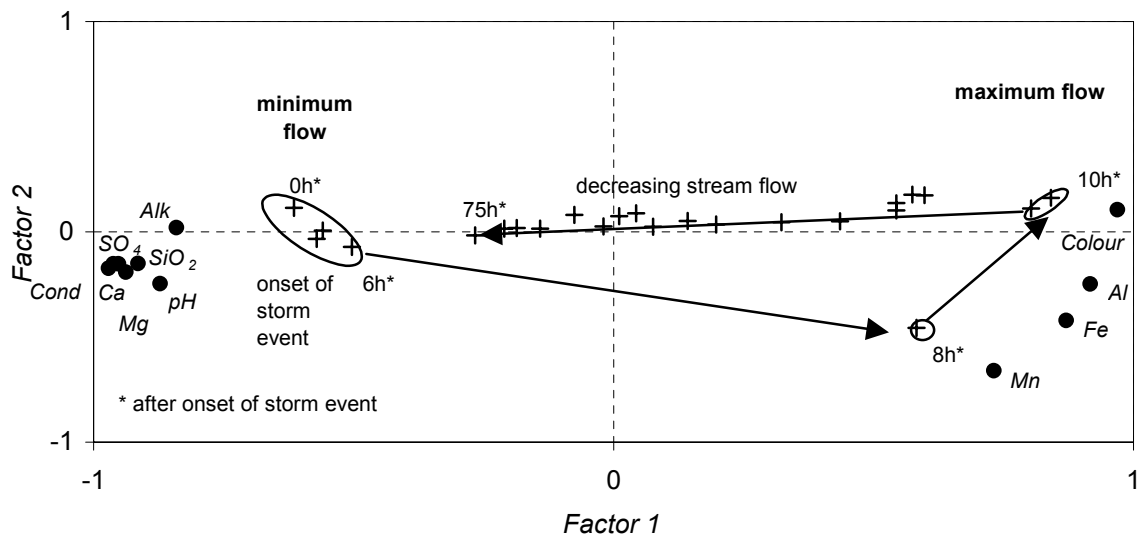


Fig. 4 Shielhope Burn sample factor loadings for the rainfall event on 18 October (the arrows indicate the chronological order in which the samples were collected and the time of sample collection is given relative to the beginning of hydrograph rise).

loadings (close to baseflow composition) on factor 1 when discharge is low, but as the hydrograph rises, sample loadings on factor 1 steadily increase, and reach a maximum at peak flow. The associated maximum in colour and Al concentrations in the original data indicates that contributions from the organic-rich soil horizons are highest at peak flow. On the recession limb of the hydrograph, factor 1 loadings, and thus soil water inputs, gradually decline towards the initial negative values and baseflow composition.

Factor 2 also shows some relation to flow (Fig. 3), but appears to be less closely associated than factor 1. Factor 2 loadings in all streams are relatively small, except for

a few samples collected during storm events. In the factor diagrams, these samples are easily identified as a small group of samples with high loadings on factor 2 (see encircled samples in Fig. 1). During storm events, abrupt changes, noted as high factor 2 loadings, occur just before the hydrograph peaks and are followed by a rapid decline to pre-storm values (Fig. 4). The sudden change in factor 2 loadings coincides with maximum Fe and Mn values in the original data and is interpreted as the displacement of deeper Fe and Mn-rich soilwater from the waterlogged gleyed soils in the riparian zone (Reid *et al.*, 1981). This process is most pronounced in Shielhope Burn, which has the highest proportion of riparian gleyed soils.

End-member mixing analysis

The results of the hydrograph separations, exemplified for Winterhope Burn, are shown in Fig. 5 for the intensive stream sampling period. The graph shows that inputs from the upper, organic-rich soil horizons have a major influence on the stream's hydrochemistry (Fig. 5). It is apparent that soilwater contributions during the monitoring period are often large enough to dominate stream chemistry and to override the groundwater signal, although significant proportions of deep water are also reaching the stream. It was found that groundwater inputs at individual streams make up between 26% and 43% of streamflow at maximum flow. Groundwater inputs dominate the stream chemistry during periods of low flow, as well as during hydrograph rise (Fig. 5), and generally peak about 1–2 h before the peak of the soilwater end-member, which typically occurs at maximum flow. This early peak is probably associated with the displacement of deep soilwater and groundwater as subsurface flow from the steeper hillslopes enters the riparian zone (Giusti & Neal, 1993). The process is most pronounced in Shielhope Burn, due to the high proportion of waterlogged gley soils in the riparian area, which promote the reduction of Fe and Mn and subsequent flushing of Fe and Mn into the stream along with displaced riparian groundwater (Reid *et al.*, 1981).

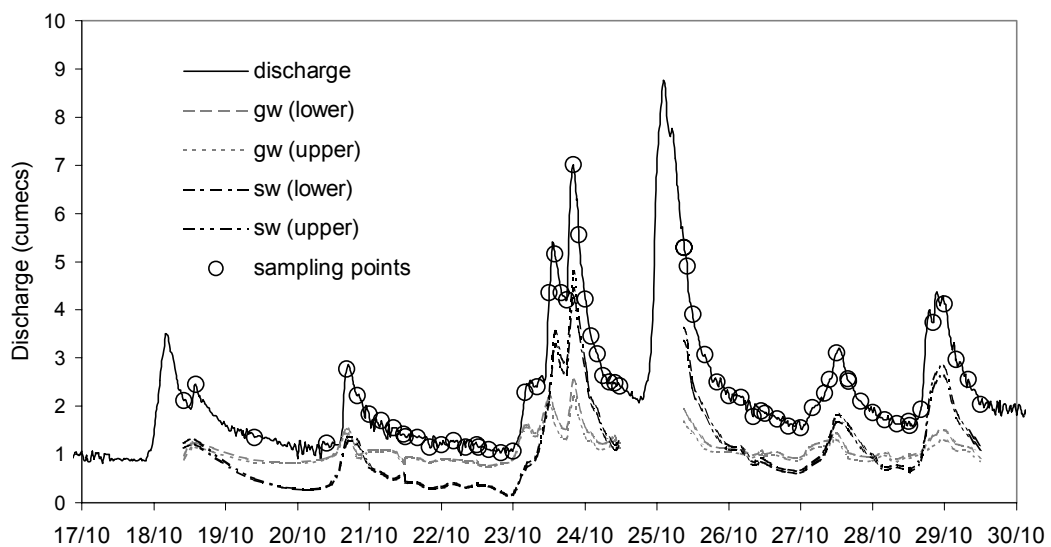


Fig. 5 Hydrograph split for Winterhope Burn based on event data.

Uncertainty in the hydrograph split depends almost entirely upon the end-member selection, which in the case of soilwater is chosen from literature sources (Table 1). Therefore, the proportion of groundwater contribution to flow is calculated and plotted for both an upper (alkalinity $-50 \mu\text{Eq L}^{-1}$) and a lower (alkalinity $-75 \mu\text{Eq L}^{-1}$) soilwater end-member (Fig. 5). Errors in groundwater proportion are small (1–5%) and approximately constant with time, with higher divergence occurring at high flows.

DISCUSSION AND CONCLUSIONS

Factor analysis and EMMA were successfully applied to the data and demonstrate that flow related changes in soilwater inputs (associated with changes in flow pathways) exert a major control on stream chemistry during storm events. Uncertainty in EMMA related to the end-member definition were shown to be relatively small ($< 5\%$) and the trends shown are not sensitive to the end member values. Hence, the fundamental processes are well presented by the technique, even if the end-members may not be defined accurately.

Two Fe and Mn sources were identified within the studied catchments; (1) organic soilwater sources associated with Fe and Mn accumulation in organic-rich soil horizons and (2) deep soilwater/groundwater sources related to metal mobilization at low pH and/or reductive solution of Fe and Mn oxides and hydroxides under low redox conditions in deeper soils of the riparian zone. The organic soilwater source is typically the dominant Fe and Mn source during storm events, where runoff from organic soil horizons increases. Maximum contributions from this source are associated with maximum soilwater inputs, which typically occur at peak discharge. Hence it appears that autumn and winter storm events present critical periods for high Fe and Mn concentrations in the stream runoff of the Megget catchment.

Groundwater contributions also increase during storm events and such buffering of streamwater chemistry by groundwater inputs at high flows has also been reported in other Scottish upland catchments (Giusti & Neal, 1993; Soulsby *et al.*, 1998). It is mainly attributed to inputs from riparian areas and from valley bottom glacial deposits fed by recharge from the upper parts of the catchment. As this subsurface flow from the steeper hillslopes enters the riparian zone, Fe and Mn associated with the deep soilwater/groundwater source are flushed into the stream along with the displaced riparian groundwater. These Fe and Mn loadings may be markedly higher during the first autumn storms, particularly after a dry summer period, due to prolonged time available for the accumulation of Fe and Mn in the soils (Neal *et al.*, 1997a).

From a management perspective, the above findings imply that autumn and winter storm events present the most critical periods, during which the bulk of Fe and Mn is delivered to the streams and subsequently to the reservoir and show that metal fluxes into a reservoir could be modelled and predicted for this (and similar) upland catchments. However, any variability in storm frequency and magnitude, a predicted outcome of current climate change scenarios (Werritty *et al.*, 2001), will affect the accumulation and retention of metals in soils, and the rate and total amount of metals entering the stream environment and reservoir.

Acknowledgements Funding and facilities for this study have been provided by East of Scotland Water (now part of Scottish Water) and by the University of St Andrews.

REFERENCES

- Abesser, C. (2003) Water quality functioning of an upland reservoir and its catchment area-with particular emphasis on iron and manganese. PhD thesis, University of St Andrews, St Andrews, UK.
- Carling, P. A., Irvine, B. J., Hill, A. & Wood, M. (2001) Reducing sediment inputs to Scottish streams, a review of the efficacy of soil conservation practices in upland forestry. *Sci. Total Environ.* **265**(1), 209–229.
- CEC Commission of European Communities (1980) Council Directive 80/778/EEC relating to the quality of water intended for human consumption. *Official J.* **L/229**, 1–23.
- Christophersen, N. & Neal, C. (1990) Linking hydrochemical, geochemical and soil chemical processes on the catchment scale, an interplay between modelling and field work. *Water Resour. Res.* **26**, 3077–3086.
- Ferrier, R. C., Jenkins, A., Miller, J. D., Walker, T. A. B. & Anderson, H. A. (1990) Assessment of wet deposition mechanism in an upland Scottish catchment. *J. Hydrol.* **113**, 285–296.
- Giusti, L. & Neal, C. (1993) Hydrological pathways and solute chemistry of storm runoff at Dargall Lane, southwest Scotland. *J. Hydrol.* **142**, 1–27.
- Grzybowski, D. S. (2001) A study of catchment discharge from tributaries that supply the Megget Reservoir in the Southern Uplands of Scotland. Dissertation Thesis, University of St Andrews, St Andrews, UK.
- Heal, K. V., Kneale, P. E. & McDonald, A. T. (2002) Manganese in runoff from upland catchments, temporal patterns and control on mobilization. *Hydrol. Sci. J.* **47**(5), 769–780.
- Kirchner, J. W. & Lydersen, E. (1995) Base cation depletion and potential long-term acidification of Norwegian catchments. *Environ. Sci. Technol.* **29**, 1953–1960.
- Langan, S. J. & Hurst, D. (2004) An analysis of the long-term variations in streamwater quality for three upland catchments at Loch Dee (Galloway, SW Scotland) under contrasting land use. *Hydrol. Earth System Sci.* **8**(3), 422–435.
- Little, D. I. & McFadzean, C. J. (1991) Bradan water treatment works extensions and upgrading. In: *Water and Environmental Management* (ed. by M. D. F. Haigh & C. P. James), 178–188. Ellis Horwood, Chichester, UK.
- MacDonald, T. D. (1994) Water supply. In: *The Fresh Waters of Scotland, a National Resource of International Significance* (ed. by P. S. Maitland, P. J. Boon & D. S. McLusky), 279–296. Wiley, Chichester, UK.
- Neal, C., Hill, T., Hill, S. & Reynolds, B. (1997a) Acid neutralization capacity measurements in surface and groundwaters in the Upper River Severn, Plynlimon, from hydrograph splitting to water flow pathways. *Hydrol. Earth Syst. Sci.* **3**, 687–696.
- Neal, C., Wilkinson, J., Neal, M., Harrow, M., Wickham, H., Hill, L. & Morfitt, C. (1997b) The hydrogeochemistry of the headwaters of the River Severn, Plynlimon. *Hydrol. Earth Syst. Sci.* **1**(3), 583–617.
- Robins, N. S. (2002) Groundwater quality in Scotland, major ion chemistry of the key groundwater bodies. *Sci. Total Environ.* **294**(1–3), 41–56.
- Schofield, T., Perkins, R. & Simms, J. S. (1991) Frankley water-treatment works development, pilot studies. *JIWEM* **5**, 370–380.
- Soulsby, C., Chen, M., Ferrier, R. C., Helliwell, R. C., Jenkins, A. & Harriman, R. (1998) Hydrogeochemistry of shallow groundwater in an upland Scottish catchment. *Hydrol. Process.* **12**, 1111–1127.
- Soulsby, C. & Dunn, S. (2003) Towards integrating tracer studies with a conceptual rainfall-runoff model, recent insights from a sub-arctic catchment in the Cairngorm Mountains, Scotland. *Hydrol. Process.* **17**, 403–416.
- Soulsby, C., Rodgers, P., Smart, R., Dawson, J. & Dunn, S. (2003) A tracer based assessment of hydrological pathways at different spatial scales in a mesoscale Scottish catchment. *Hydrol. Process.* **17**, 759–777.
- Walden, J. & Smith, J. P. (1995) Factor analysis, a practical application. In: *Statistical Modelling of Quaternary Science Data* (ed. by D. Maddy & J. S. Brew), 39–63. Quaternary Research Association, Cambridge, UK.
- Werritty, A., Black, A. & Duck, R. (2001) Climate Change: Flooding occurrence review. Scottish Executive. 01-May-03, <http://www.scotland.gov.uk/cru/kd01/lightgreen/ccfo-06.asp>