

Comparison of groundwater sampling methods for State of the Environment monitoring in New Zealand

C. J. Daughney,¹ T. Baker,² A. Jones,³ C. Hanson,⁴ P. Davidson,⁵
M. Thompson,⁶ R. R. Reeves⁷ and G. M. Zemansky⁷

¹ GNS Science, PO Box 30368, Lower Hutt, New Zealand. Corresponding author: c.daughney@gns.cri.nz

² Greater Wellington Regional Council, PO Box 41, Masterton, New Zealand

³ Greater Wellington Regional Council, PO Box 11646, Wellington, New Zealand

⁴ Environment Canterbury, PO Box 345, Christchurch, New Zealand

⁵ Marlborough District Council, PO Box 443, Blenheim, New Zealand

⁶ Ministry for the Environment, PO Box 10362, Wellington, New Zealand

⁷ GNS Science, Private Bag 2000, Taupo, New Zealand

Abstract

This investigation involved a field trial to assess the possible effects of different sampling procedures on the chemistry of groundwater samples from 49 typical State of the Environment (SOE) monitoring sites in the Wellington (including Wairarapa), Marlborough and Canterbury regions. The results indicate that it is essential to adequately purge standing water from the well prior to sample collection to ensure that the water samples will be representative of groundwater within the aquifer. Adequate purging requires removal of three times the volume of water in the well under ambient non-pumping conditions *and* simultaneous stabilisation of pH, conductivity and temperature readings to within ± 0.1 pH units, $\pm 3\%$, and $\pm 0.2^\circ\text{C}$, respectively. The results also indicate that samples should be collected sequentially and directly from a sample point rather than from an intermediate container (e.g., a bucket). There were no significant differences in anion results from filtered versus unfiltered samples. Filtration can be performed either with in-line filters or with syringes and syringe-tip

filters. The results from this investigation will assist with the development of a single best-practice sampling procedure that will allow for meaningful comparison of SOE groundwater quality data at the regional and national scales, even if the samples were collected by different people or at different times.

Keywords

Groundwater quality, State of the Environment monitoring, sampling

Introduction

Groundwater quality is monitored by many organisations in New Zealand, including regional authorities, research organisations, consultancies, industries and individual members of the public. Groundwater quality data are collected for many reasons, including state-of-the-environment (SOE) monitoring, resource characterisation, and consent and compliance monitoring. It is well established that groundwater quality data can be influenced by the way a sample is collected,

preserved and/or transported prior to analysis. For example, the sample's chemistry can be affected by the duration of pumping prior to sample collection (known as purging), pump rate during purging, sample volume, filtration method, and reactivity of materials used in tubing and bottles (e.g., Claasen, 1982; Barcelona *et al.*, 1985; Reynolds *et al.*, 1990; Wilson, 1995; Rosen *et al.*, 1999). Analytical data from samples that are collected using inappropriate sampling methods might be unrepresentative of the groundwater quality within the aquifer.

This investigation pertains to groundwater sampling methods used for SOE monitoring in New Zealand. The 15 regional authorities operate SOE monitoring programmes that vary in terms of number of sites (from about 8 to 300 sites per region, or about 1000 sites across the entire country) and sampling frequency (weekly, monthly, quarterly, annually, or once every two or three years). All regional authorities also participate in the National Groundwater Monitoring Programme (NGMP), which is coordinated by GNS Science and involves quarterly monitoring of 112 sites across the country (Rosen, 1999; Daughney and Reeves, 2005, 2006). Most SOE monitoring sites are privately owned wells that are used for water supply for domestic, farm, municipal or irrigation purposes. Most SOE (including NGMP) sampling involves collection of 1) a field-filtered acid preserved sample, which is usually analysed for major cations and some metals (e.g., iron, manganese, usually above the part-per-billion range); 2) an unfiltered unpreserved (raw) sample, which is usually analysed for alkalinity, laboratory measurement of pH and conductivity, and in some programmes for anions and nutrients; and in some cases; 3) a field-filtered unpreserved sample, which is usually analysed for anions and nutrients if these are not measured in the raw sample. Certain parameters such as pH, conductivity

and water temperature are also usually measured in the field.

Many different methods for SOE groundwater sampling are currently in use in New Zealand, and so groundwater quality data may not be readily comparable between different regions, and even within a single region, older groundwater quality data may not be directly comparable to newer data. This paper describes results from a field trial conducted to assess the possible effects of different sampling procedures on the chemistry of groundwater samples collected for SOE monitoring. The field trial was conducted to provide data that would underpin the development of a standard national groundwater sampling protocol for SOE monitoring that would be specifically relevant to New Zealand, by addressing the following questions (see also Daughney *et al.*, 2006):

1. Are typical purging criteria appropriate? Groundwater sampling guidelines usually specify criteria that must be met to demonstrate that an adequate volume of water has been purged from the well prior to the collection of samples. The need for purging is based on the assumption that stagnant water in the well has different chemical characteristics than groundwater within the aquifer some distance away from the well. This investigation aimed to determine what criteria should be used to demonstrate adequacy of purging (i.e., once the purging criteria are met, samples collected at different times afterwards will yield identical analytical results). Some SOE monitoring sites are pumped continuously and thus do not need to be purged prior to sampling; however, such sites were avoided in this field trial so that the importance of purging could be addressed.
2. Should samples be collected from an intermediate container? In the NGMP, it has been traditional to collect the filtered

preserved, filtered unpreserved and raw samples from a single intermediate container, like a bucket or pail (Rosen *et al.*, 1999). This is to ensure that the calculated charge balance error can be used as a check for analytical accuracy and sample integrity (Freeze and Cherry, 1979). However, use of the intermediate container exposes the sample to air and light, and so degassing and oxidation might affect the concentrations of some analytes, such as iron, manganese, or alkalinity. For these reasons, the SOE samples in some regions are collected sequentially from the sample point or pump outlet, while the pump is operating, instead of from an intermediate container. This investigation aimed to determine if samples collected sequentially from the pump outlet would yield acceptable charge balance errors, and whether it would differ from charge balance errors calculated for samples collected from a single intermediate container.

3. Do samples for analysis of anions need to be field-filtered? Many previous investigations have demonstrated the importance of field-filtration for samples that are to be analysed for cations and metals (e.g., Claasen, 1982; Wilde *et al.*, chapters variably dated), but the importance of field-filtration for anions, nutrients or silica remains unclear. Historically, most SOE monitoring in New Zealand (including the NGMP) has relied on an unfiltered unpreserved sample for alkalinity. This is because the standard laboratory analytical method for alkalinity (APHA 2320) specifically states that samples should not be filtered, in order to avoid possible loss of dissolved carbon dioxide during filtration (American Public Health Association, 2001). However, analysis of alkalinity in an unfiltered sample is potentially biased by inclusion of titratable suspended organic carbon or mineral particulates (Langmuir, 1997). Many SOE

programmes in New Zealand (including the NGMP) involve collection of filtered unpreserved samples for analysis of other anions such as Cl, SO₄ and NO₃, whereas other SOE programmes analyse these anions in unfiltered unpreserved samples. Some SOE programmes involve collection of a filtered acid-preserved sample for SiO₂, whereas in other cases it is analysed in a filtered unpreserved sample.

4. Are in-line filters more appropriate than syringe-tip filters? For most SOE programmes in New Zealand, the filtration step is conducted using syringe-tip filters (pore size 0.1 to 0.45 µm). However, filling a syringe with sample water, affixing a filter, then pushing the water through the filter into a bottle might be an unnecessarily complex operation that has the potential to bias analytical results due to aeration and excessive handling of the sample. Thus, some SOE programmes (particularly overseas) conduct field-filtration using an in-line filter that is attached directly to the sample point using a short length of clean tubing.

Methods

Sites

The field trial involved 49 groundwater monitoring sites in gravel (the majority) or sand aquifers in Canterbury (12 sites), Marlborough (12 sites), Wellington (14 sites) and Wairarapa (11 sites). Of the sites sampled, 29 were part of the NGMP and 20 were part of regional authority SOE monitoring programmes. Aquifer confinement varied from unconfined (26 sites) to semi-confined (9 sites) to confined (14 sites). Well depths ranged from 3 to 200 m (the 25th, 50th and 75th percentiles in depth were 10, 19 and 33 m, respectively). All samples were collected by regional authority staff, with typically five to ten sites sampled on any single day.

Sample collection

Purging (removal of standing water the well) was conducted at a pump rate appropriate to each site (minimum, median and maximum rates were 0.02, 0.6 and 240 L/s, respectively). The adequacy of purging at each site was assessed on the basis of number of well volumes of water removed, where one well volume is defined as the volume of water that is in the well casing and screened or open interval during ambient non-pumping conditions. Following standard practice in New Zealand, once more than three well volumes had been extracted from the well, the purged volume was judged to be 'sufficient'; prior to this the purged volume was considered 'insufficient'. The adequacy of purging at each site was also assessed based on changes in field measurements of pH, conductivity and temperature. Following criteria used overseas (Wilde *et al.*, chapters variously dated), field readings are described as either 'stable' or 'unstable', depending on whether or not pH, conductivity and temperature had simultaneously stabilised to within ± 0.1 pH units, $\pm 3\%$ ($\pm 5\%$ if less than 100 $\mu\text{S}/\text{cm}$) and $\pm 0.2^\circ\text{C}$, respectively, based on at least two measurements made at least one well volume or five minutes apart.

Sample collection was initiated at various stages in the purging operation (Table 1), as a means of assessing 1) the importance of adequate purging and 2) the appropriateness

of the purging criteria based on extracted volume compared to stabilisation of field parameters. At 29 sites, sample collection was initiated only after 'adequate' purging, as evidenced by stable field readings *and* sufficient pumped volume. At 18 sites, sample collection was initiated when the purging operation was judged to be 'inadequate', either based on sufficient pumped volume but unstable field readings (7 sites), stable field readings but insufficient pumped volume (8 sites), or unstable field readings in combination with insufficient pumped volume (3 sites). At two sites, certain field data were accidentally not recorded and thus it was not possible to determine the adequacy of the purging operation at the time sample collection was initiated.

The sampling operation involved collection of several sets of samples from each site, using the different methods summarised in Figure 1 (see also Daughney *et al.*, 2006). Samples collected from all NGMP sites and from the SOE sites in Marlborough were analysed at the GNS Science laboratory in Wairakei. Samples from SOE sites in Canterbury were analysed by the Environment Canterbury lab in Christchurch. Samples from SOE sites in Wellington and Wairarapa were analysed by Hill Laboratories in Hamilton.

Data interpretation

'Difference distributions' were used to compare the chemical analyses of the samples collected using the different sampling procedures. As an example, Procedure A might be collection of a sample using an in-line filter, whereas Procedure B might be collection of a second sample from the same site on the same day using a syringe-tip filter. For a particular site, Result A might be a Na concentration of 10.1 g m^{-3} , and Result B might be a Na concentration of 9.8 g m^{-3} . For this site, the relative difference

Table 1 – Assessment of adequacy of purging prior to sample collection at sites included in this investigation.

Overall Assessment	Well volumes extracted	Field readings	n
Inadequate	Insufficient	Unstable	3
	Insufficient	Stable	8
	Sufficient	Unstable	7
Adequate	Sufficient	Stable	29
Unknown			2

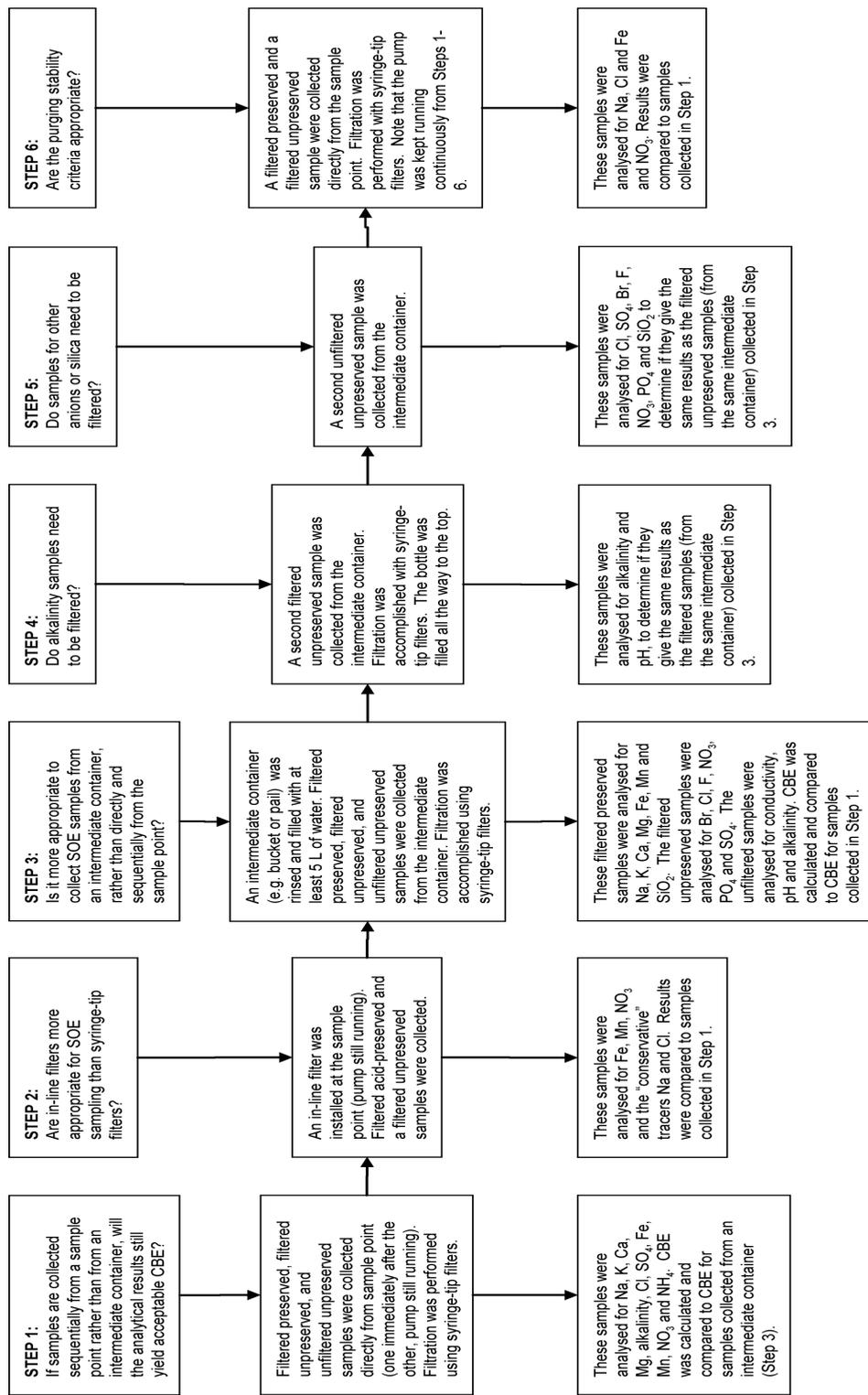


Figure 1 – Flow diagram of sampling procedures followed during this investigation. Top row of boxes describe questions and issues each step in the sampling procedure was meant to address. Middle and bottom rows of boxes describe sampling procedures and laboratory analytical methods, respectively. CBE – charge balance error.

in Na concentration would be 3%, according to the formula:

$$\text{Relative \% Difference} = \frac{\text{Result A} - \text{Result B}}{\text{Average}(\text{Result A}, \text{Result B})} \times 100\% \quad (1)$$

This comparison of Result A to Result B was performed for many different sites, and then the mean, median and standard deviation of the difference distribution were determined. *A posteriori* power analysis (Helsel and Hirsch, 1992) was used to calculate tolerance limits for the mean of the various difference distributions (at the 95% confidence level), to determine whether or not the analytical data differed as a result of the various sampling procedures shown in Figure 1.

Three example difference distributions are displayed in Figure 2. First, there is the case where the mean is close to zero and the standard deviation is small, e.g., 2% and 5% respectively. In this case, the mean would be statistically indistinguishable from zero, and we would conclude that the two different sampling procedures do not lead to a systematic difference in the analytical results. As a second example, there is the case where the mean is not near zero but the standard deviation is large, e.g., 10% and 15% respectively. In this case, the mean is still statistically indistinguishable from zero, but we could conclude that there are large differences in analytical results obtained

at some sites (which give rise to the large standard deviation), and this might indicate a bias relating to the sampling procedures. As a third example, there is the case where the mean is not near zero and the standard deviation is relatively small, e.g., 10% and 5% respectively. In this case we could conclude that the two different sampling procedures do in fact cause a significant and systematic difference in the analytical results. However, even if the mean of the difference distribution is shown to be statistically different from zero, if it is less than about $\pm 5\%$, we would conclude that it is close enough to zero that for SOE groundwater monitoring purposes, the two sampling procedures can be assumed to produce identical analytical results.

Results

Stabilisation of field parameters during purging

At 32 of the 37 sites where stable field readings were obtained, the stabilisation occurred prior to extraction of three well volumes of water. However, at some sites up to 25 well volumes of water had to be extracted from the well in order to achieve simultaneous stabilisation of all three field parameters. Graphs of field parameters versus well volumes purged are provided for four wells (Fig. 3). All of these except well S26/0705 illustrate stabilisation of field parameters within three well volumes of water extracted from the well. Well R25/5190 illustrates perhaps the most expected case, with initial values of field readings rapidly and consistently approaching stable values during the purging operation. In the case of well S26/0705 there is considerable variation in all three field parameters, followed by stabilisation after extraction of more than seven well volumes of water from the well. The nature of this variation is unusual for a groundwater system, indicating possible instrumentation problems rather than changes in actual

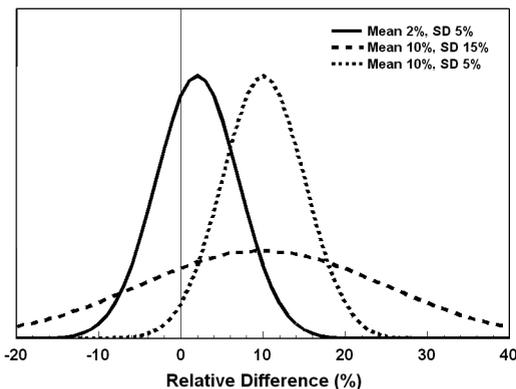


Figure 2 – Example difference distributions.

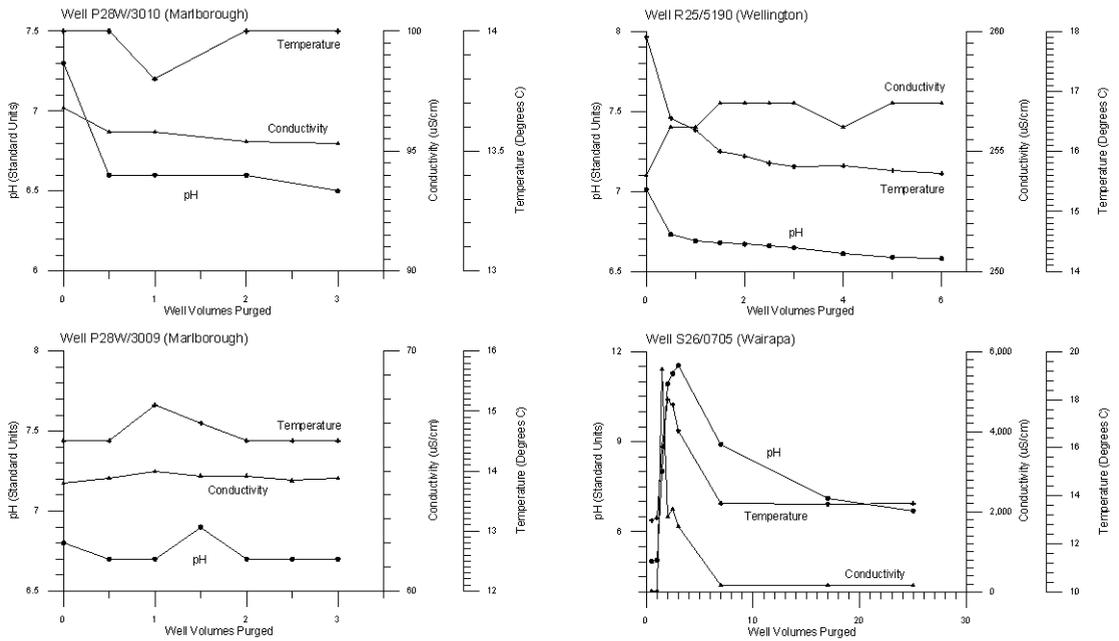


Figure 3 – Graphs of field parameters versus well volumes purged for four sites sampled in this investigation. Wells P28W/3010, P28W/3009 and R25/5190 are typical of most sites in this study. Well S26/0705 is an example of a site that took a long time to stabilise.

groundwater chemistry during purging. The importance of adequate purging and the appropriateness of the different purging criteria are discussed below.

Effect of sample collection from an intermediate container

The calculated charge balance error was within acceptable limits ($\pm 5\%$) for most adequately purged sites, regardless of whether the samples were collected from the intermediate container or sequentially from the sample point. The ability to obtain charge balance errors within acceptable limits when samples are collected sequentially from the sample point is supported by data from the Environment Canterbury SOE programme, which has historically used this approach (Abraham and Hanson, 2003). Similarly, the ability to obtain charge balance errors within acceptable limits for samples collected from a single intermediate container is verified

by data from the NGMP (Daughney and Reeves, 2005).

The data from adequately purged sites suggest that it may be more appropriate to collect samples sequentially from the sample point rather than from an intermediate container. Data from inadequately purged sites were not considered, in order to avoid possible cases where a charge balance error outside acceptable limits was caused by a real difference in the chemistry of the samples that were collected sequentially from the pump outlet. The mean of the distribution of charge balance errors for samples from the intermediate container was -2.1 ± 1.2 (Table 2), which is slightly but significantly less than the ideal value of zero. Arguably, the departure from the ideal value of zero is so slight that it is not environmentally relevant for SOE monitoring purposes, but it may indicate biasing of the analytical results due to use of the intermediate container (see

Table 2 – Distributional parameters for Charge Balance Error for adequately purged sites.

Method sampling	Distribution of Charge Balance Error (%)			
	n	Mean	SD	CI
Sequentially from sample point	29	-0.2	2.2	0.8
From intermediate container	29	-2.1	3.1	1.2

n, SD and CI refer to the number of sites, the standard deviation of the distribution, and 95% confidence interval around the mean, respectively.

discussion about effect of filtration on anion results). By comparison, the charge balance error values associated with samples collected directly from the sample point are more narrowly distributed around the ideal value of zero.

Effect of filtration method

The % difference was calculated according to Equation 1, with Result A referring to the sample collected directly from the pump outlet using an in-line filter, and Result B referring to the sample collected from the pump outlet using a syringe-tip filter. Data from inadequately purged sites were excluded in order to avoid possible cases where apparent differences in the analytical results were related to purging procedure instead of filtration method. In addition, data from sites at which concentrations were below or near the detection limit were excluded, in order to avoid cases where observed differences in results were caused by analytical uncertainty instead of the filtration method.

When data from all of the adequately purged sites are assessed together, there are no significant systematic differences in the analytical data that can be positively related to filtration method (Table 3). The means of the difference distributions are close to zero for all of the analytes considered, and the standard deviations are relatively small. We

Table 3 – Distributional parameters for difference distributions comparing filtration methods at adequately purged sites.

Analyte	n	Mean	SD	CI
Na	28	0.6	1.8	0.7
Cl	27	-0.1	1.8	0.7
NO ₃	26	1.1	4.1	1.6
Fe	20	4.2	11.4	5.3

n, SD and CI refer to the number of sites, the standard deviation of the distribution, and 95% confidence interval around the mean, respectively.

conclude that the two filtration methods do not affect Na, Cl or NO₃ concentrations by more than $\pm 5\%$ or Fe concentrations by more than $\pm 10\%$. Most of the sites considered in the field trial had Mn concentration below the analytical detection limit, and so comparisons based on Mn are not meaningful and are not tabulated or discussed further in this paper.

Effect of filtration of anion results

The % difference was calculated according to Equation 1, with Result A and Result B referring to the filtered and unfiltered samples collected from the intermediate container, respectively. Again, we exclude data from sites at which concentrations are below or near the analytical detection limit. This analysis does not exclude data from inadequately purged sites, because the sampling procedures being compared both involved collection of water from the same intermediate container, and this would not be influenced by the extent of purging.

When data from all sites are considered together as a single group, there are no statistically significant differences in alkalinity or anion results obtained from filtered samples compared to unfiltered samples (Table 4). For alkalinity, pH and SiO₂, the difference distributions have mean values near zero and relatively small standard deviations, and thus we conclude that the filtration, compared to lack of filtration, does not lead to a systematic difference in the analytical results. For all

Table 4 – Distributional parameters for difference distributions comparing filtration to lack of filtration for samples collected from an intermediate container.

Analyte	n	Mean	SD	CI
Alkalinity	49	0.1	1.7	0.5
pH	33	-0.9	3.5	1.2
Cl	49	0.6	25.7	7.4
SO ₄	47	1.4	21.8	6.4
NO ₃	42	1.2	39.4	12.3
PO ₄	35	-10.3	50.0	17.7
F	33	-13.5	31.0	11.0
Br	14	-21.7	38.8	22.4
SiO ₂	43	1.3	6.5	2.0

n, SD and CI refer to the number of sites, the standard deviation of the distribution, and 95% confidence interval around the mean, respectively.

other analytes, the standard deviations of the difference distributions are relatively large. This means that it is not possible to identify cases where the means of the difference distributions are significantly different from zero, but the large standard deviations indicate that there are large differences in analytical results obtained at some sites, which might indicate a bias relating to the sampling procedures.

When data from each site are considered independently (including cases of both adequate and inadequate purging), at most sites there is little difference in the results obtained for filtered versus unfiltered samples. However, there are 18 cases across ten sites where analyte variations exceed $\pm 10\%$ for the filtered compared to the unfiltered sample (Table 5). Several lines of evidence suggest that the observed differences are actually related to the excessive sample handling involved with use of the intermediate container, rather than to filtration. First, there are several instances where significant variations are observed for Na and/or Cl, even though these analytes were specifically selected because they were not expected to be influenced by filtration, and might be used somewhat like conservative tracers. Second,

for each analyte, neither sampling procedure leads to results that are consistently higher or lower than the other. Third, comparisons to results from samples collected in other steps of the field trial (Table 5) reveal what appear to be spurious analytical results, and wherever such spurious results could be identified, they were for samples that had been collected from the intermediate container rather than directly from the sample point. Finally, as noted above, the mean of the distribution of charge balance errors for samples from the intermediate container was found to be slightly but significantly less than the ideal value of zero (Table 2), which could also be caused by analyte variations arising from the use of the intermediate container.

Importance of purging and appropriateness of purging criteria

The % difference was calculated according to Equation 1, with Result A and Result B referring to the first samples and last samples collected directly from the pump outlet using syringe-tip filters, respectively. To assess the importance of purging, this analysis was performed first for all adequately purged sites considered as a single group, and then for all inadequately purged sites as a single group. To assess the appropriateness of the different purging criteria, the analysis was repeated first for inadequately purged sites where sample collection was initiated after sufficient pumped volume but while the field readings were still unstable, and second for sites where the field readings were stable but the pumped volume was insufficient. In all cases, data were excluded if concentrations were below or near the detection limit.

When the data from all adequately purged sites are considered together as a single group, there are no systematic significant differences in the Na, Cl, NO₃ or Fe results for the first versus last samples collected from the pump outlet (Table 6). For each analyte, the

Table 5 – Assessment of cause and significance of observed variations in analytical results. This table presents results for samples collected by various methods, but only at the sites where significant variations had been identified.

Purging	Site Name	Analyte	Analytical Results ¹					Cause of Variation ²
			A	B	C	D	E	
Adequate	J39/0109	NO ₃	2.9	2.8	3.9*	2.8	3.1	1
	J40/0080	Fe	0.20	0.18	0.41*		0.14	1
	M35/1382	Na	3.8	3.8	3.3		3.7	1
	O31/0156	PO ₄	1.523		1.095*	1.511		1
	P28W/3010	NO ₃	1.10	1.10	1.10	0.42*	0.96	1
		Cl	2.5	2.5	2.6	8.0*	2.5	1
		SiO ₂			12.1	16.5		?
	R26/6587	Cl	20.8	22.4	19.1	17.9	19.2	1
	S26/0705	Na	16.3	16.4	14.9*		16.4	1
	S27/0299	Cl	8.1	8.1	2.5*	8.1	8.2	1
		SO ₄	3.7		1.1*	3.7		1
		Fe	0.03	0.03	0.04		0.04	?
	S27/0607	F			0.38	0.33		?
		PO ₄	1.670		3.426*	1.937		1
	T26/0538	Cl	28.7	28.8	29.2	25.7*	28.6	1
Inadequate	P28W/3069	Na	11.8	11.8	8.1		8.2	2
		Cl	8.0	8.0	2.9	2.5	2.8	2
		NO ₃	0.04	0.04	1.9	1.1	1.9	2
		SO ₄	5.1		10.6*	5.0		1
		SiO ₂			16.0	12.8		?
	P28W/4402	Na	14.6	14.6	11.7		11.8	2
		Cl	5.3	5.3	8.0	2.8*	8.0	1,2
		NO ₃	0.7	0.71	0.04	2*	0.04	1,2
		SO ₄	4.2		5.2	10.7		?
		Fe	<0.02	<0.02	0.51		0.63	2
	P28W/4403	Na	3.9	3.9	14.5		14.7	2
		Cl	3.6	3.6	5.2	5.2	5.2	2
	P28W/4404	Na	8.1	8.1	4.0		3.9	2
		Cl	2.9	2.8	3.6	3.6	3.6	2
		NO ₃	1.9	1.9	0.31	0.31	0.31	2
R25/5100	Fe		0.36	0.28		0.72	?	
	NO ₃	<0.002	<0.002	0.007	<0.002	0.003	1	

¹ Analytical results for A) first samples collected from the sample point using syringe-tip filters, B) samples collected from the sample point using in-line filters; C) filtered samples from the intermediate container; D) unfiltered samples from the intermediate container; E) last samples collected from the sample point using syringe-tip filters. All results reported in g m⁻³ (as N for NO₃ and as P for PO₄).

² Assessment of whether variation is caused by 1) spurious results (denoted by *) obtained for samples from the intermediate container or 2) inadequate purging.

difference distribution has a mean value near zero and a relatively small standard deviation. We conclude that the purging criteria are appropriate and effective, in that once met, samples collected at different times afterwards will yield identical analytical results.

When the data from all inadequately purged sites are considered together, the difference distributions again have mean values that are statistically indistinguishable from zero, but the standard deviations are much larger than observed for the adequately purged sites (Table 6). This indicates that there are several inadequately purged sites at which analyte concentrations varied substantially between the first and last samples collected from the sample point. An assessment of individual results reveals 12 cases across four sites where Na, Cl, NO₃ and/or Fe concentrations vary by more than ±10% (Table 5). This reinforces the conclusion reached above, namely that it is essential to satisfy the purging criteria prior to sample collection, in order to obtain representative analytical results.

For inadequately purged sites where sample collection was initiated after the field readings were stable but the pumped volume was insufficient, most of the difference distributions have mean values near zero and relatively small standard deviations (Table 6). In comparison, for sites where sample collection was initiated after sufficient pumped volume but while the field readings were still unstable, the difference distributions have much larger standard deviations. This suggests that it is more appropriate to assess the adequacy of purging based on the stabilisation of field parameters than on the total volume of water that has been purged from the bore. Note that neither criterion on its own is as effective as both together, so we conclude that it is best to satisfy both types of purging criteria prior to sample collection.

Table 6 – Distributional parameters for difference distributions comparing first versus last samples collected directly from the pump outlet, as a function of adequacy of purging.

Analyte	Adequately Purged				Inadequately Purged				Stable, Insufficient*				Unstable, Sufficient*			
	n	Mean	SD	CI	n	Mean	SD	CI	n	Mean	SD	CI	n	Mean	SD	CI
Na	29	0.4	2.1	0.8	18	1.0	34.5	17.2	8	0.1	4.0	3.4	7	1.6	57.9	53.5
Cl	29	0.7	2.1	0.8	18	-0.3	27.5	13.6	8	-0.3	3.7	3.1	7	-0.8	46.0	42.6
NO ₃	25	0.2	5.8	2.4	14	-11.7	69.2	39.9	5	-0.5	5.0	6.2	7	3.6	125.9	116.4
Fe	20	-0.5	10.5	4.9	17	0.8	71.8	36.9	7	10.7	28.3	26.2	7	0.7	109.8	101.8

n, SD and CI refer to the number of sites, the standard deviation of the distribution, and 95% confidence interval around the mean, respectively.

* Inadequately purged sites that were sampled when the field parameters were stable but the pumped volume was insufficient, compared to sites that were sampled when the field parameters were unstable but the pumped volume was sufficient.

Summary and conclusions

The data indicate that the criteria used to assess adequacy of purging are achievable and appropriate for typical SOE sites in New Zealand. Specifically, a well must be purged until at least three well volumes of water have been pumped from the well *and* field measurements of pH, conductivity and temperature have simultaneously stabilised to within ± 0.1 pH units, $\pm 3\%$ ($\pm 5\%$ if less than $100 \mu\text{S}/\text{cm}$) and ± 0.2 °C, respectively, based on at least two measurements made at least one well volume apart. The results from the field trial clearly show significant variation in the chemistry of samples collected at some sites at different times prior to adequate purging of the well. It can be concluded that the purging criteria specified above must be met in order to obtain representative groundwater samples at typical SOE sites in New Zealand.

The data indicate that samples should be collected sequentially from a tap or similar sample point while the pump is running, rather than from an intermediate container. The results from the field trial revealed some small but statistically significant shifts in analytical data from samples collected sequentially from the sample point compared to an intermediate container. In addition, the mean charge balance error was slightly but significantly different from the ideal value of zero for the samples collected from the intermediate container. Most importantly, in all cases where a potentially spurious analytical result could be identified, the sample in question had been collected from an intermediate container.

Otherwise, the different sampling methods tested in this investigation did not cause systematic, significant biases in the analytical results. The data indicate that anion samples do not need to be filtered, and that SiO_2 samples do not need to be acid preserved. However, the standard laboratory analytical method for alkalinity specifies that

the sample should not be filtered (American Public Health Association, 2001). Thus, until more data are collected to indicate otherwise, we recommend collection of an unfiltered unpreserved sample for alkalinity. We also recommend a filtered preserved sample for cations and metals, and either an unfiltered unpreserved sample or a filtered unpreserved sample for other anions. Note that samples to be analysed for anions may still require filtration at some stage, simply to protect the laboratory analytical equipment, and in some instances it may be more convenient to perform the filtration in the field instead of the laboratory. Where field filtration is performed for samples to be analysed for cations or anions, either in-line filters or syringe-tip filters can be used.

We acknowledge that the conclusions reached in this investigation have so far been demonstrated only for the range of conditions encountered in the field trial, namely Fe from <0.02 to 1.9 g m^{-3} , alkalinity from 12 to 237 g m^{-3} as HCO_3 , Cl from 2.5 to 41 g m^{-3} , SO_4 from <0.03 to 38.6 g m^{-3} , NO_3 from <0.002 to 12.2 g m^{-3} as N, PO_4 from <0.004 to 0.33 g m^{-3} as P, F from <0.05 to 0.17 g m^{-3} , Br from <0.04 to 0.19 g m^{-3} , and SiO_2 from 10.9 to 59.1 g m^{-3} . The ranges of SO_4 , NO_3 and PO_4 measured in the field trial should encompass most groundwaters in New Zealand, as demonstrated by the similar ranges for these parameters measured through the NGMP (Daughney and Reeves, 2005). Some groundwaters in New Zealand can have alkalinity or concentration of Fe, Cl, Br, F or SiO_2 that are approximately double the maximum values measured in this field trial (Daughney and Reeves, 2005). However, we have no data that suggest that any effects of sampling methods would be more pronounced for any of these analytes at higher concentrations. In the future it may be important to re-assess the conclusions reached in this investigation, to take account of data that become available from other SOE sites and regions of New Zealand.

Acknowledgements

We thank Phil Abraham and Zella Smith (Environment Canterbury), Carol Mills (Marlborough District Council) and Brent King (Greater Wellington Regional Council) for assistance with collection of the samples. We also thank Moya Appleby, Marshall Muller, Ann Noddings and Colette Dedual at the GNS Science Wairakei Water-Gas Laboratory, Sue Steel, Paul Woods and the rest of the staff at the Environment Canterbury analytical lab, and the staff at Hill Laboratories (Hamilton) for analysis of the samples. The members of the New Zealand Groundwater Forum are thanked for their helpful comments and suggestions. This manuscript was greatly improved by the recommendations of two anonymous reviewers. This research was supported by funding from Ministry for the Environment, Greater Wellington Regional Council, Environment Canterbury, Marlborough District Council, and a grant from the New Zealand Foundation for Research, Science and Technology (C05X0304).

References

- Abraham, P.; Hanson, C. 2003: Groundwater quality monitoring – annual summary 2002-2003. Environment Canterbury Technical Report U03/86.
- American Public Health Association 2001: Standard Methods for the Examination of Water and Wastewater. 20th Edition. American Public Health Association, Washington.
- Barcelona, M.J.; Helfrich, J.A.; Garske, E.: 1985. Sampling tubing effects on groundwater samples. *Analytical Chemistry* 57: 460-464.
- Claasen, H.C. 1982: Guidelines and techniques for obtaining water samples that accurately represent the water chemistry of an aquifer. U.S. Geological Survey Open File Report 82-1024. Reston, Virginia.
- Daughney, C.J.; Reeves, R.R. 2005: Definition of hydrochemical facies in the New Zealand National Groundwater Monitoring Programme. *Journal of Hydrology (NZ)* 44: 105-130.
- Daughney, C.J.; Reeves, R.R. 2006: Analysis of long-term trends in New Zealand's groundwater quality based on data from the National Groundwater Monitoring Programme. *Journal of Hydrology (NZ)* 45: 41-62.
- Daughney, C.J.; Reeves, R.; Zemansky, G.; Baker, T.; Jones, A.; Hanson, C.; Davidson, P. 2006: Field trial to support development of a nationally standardised protocol for State of the Environment groundwater sampling in New Zealand. GNS Science Consultancy Report 2006/136.
- Freeze, R.A.; Cherry, J. A. 1979: Groundwater. Prentice Hall, New Jersey.
- Helsel, D.R.; Hirsch, R.M. 1992: Statistical Methods in Water Resources. Studies in Environmental Science: 49. Elsevier, Amsterdam.
- Langmuir, D. 1997: Aqueous Environmental Geochemistry. Prentice Hall, New Jersey.
- Reynolds, G.W.; Hoff, J.T.; Gilham, R.W. 1990: Sampling bias caused by materials used to monitor halocarbons in groundwater. *Environmental Science and Technology* 24: 135-142.
- Rosen, M.R. 1999: The importance of long-term, seasonal monitoring of groundwater wells in the New Zealand National Groundwater Monitoring Programme (NGMP). *Journal of Hydrology (NZ)* 38: 145-169.
- Rosen, M.R.; Cameron, S.G.; Taylor, C.B.; Reeves, R.R. 1999: New Zealand guidelines for the collection of groundwater samples for chemical and isotopic analysis. Institute of Geological and Nuclear Sciences Science Report 99/9: Wellington, New Zealand.
- Wilde, F.D.; Radtke, D.B.; Gibs, J.; Iwatsubo, R.T. (Chapters variously dated): U.S. Geological Survey Techniques of Water-Resources Investigations. Accessed May 2005 at <http://water.usgs.gov/owq/FieldManual/>
- Wilson, N. 1995: Soil water and groundwater sampling. CRC Press, Lewis Publishers, Boca Raton, Florida.