Comparison of ammonia-nitrogen concentrations between unpreserved and acid-sulfuric preserved groundwater samples

ME Santamaria Cerrutti M Moreau

GNS Science Report 2021/27
May 2022

GNS SCIENCE
TE PŪ AO

DISCLAIMER

The Institute of Geological and Nuclear Sciences Limited (GNS Science) and its funders give no warranties of any kind concerning the accuracy, completeness, timeliness or fitness for purpose of the contents of this report. GNS Science accepts no responsibility for any actions taken based on, or reliance placed on the contents of this report and GNS Science and its funders exclude to the full extent permitted by law liability for any loss, damage or expense, direct or indirect, and however caused, whether through negligence or otherwise, resulting from any person's or organisation's use of, or reliance on, the contents of this report.

BIBLIOGRAPHIC REFERENCE

Santamaria Cerrutti ME, Moreau M. 2022. Comparison of ammonia-nitrogen concentrations between unpreserved and acid-sulfuric preserved groundwater samples. Lower Hutt (NZ): GNS Science. 15 p. (GNS Science report; 2021/27). doi:10.21420/3RJM-G633.

ME Santamaria Cerrutti, GNS Science, Private Bag 2000, Taupō 3352, New Zealand

M Moreau, GNS Science, Private Bag 2000, Taupō 3352, New Zealand

www.gns.cri.nz

ISSN 2350-3424 (online)
ISBN 978-1-99-001095-8 (online)
http://dx.doi.org/10.21420/3RJM-G633

CONTENTS

TABLES

FIGURES

ATTACHMENTS

ABSTRACT

In 2016, the National Groundwater Monitoring Programme (NGMP) sampling kit was modified by adding a 100 mL sulfuric-acidified sampling bottle for the analysis of ammonia-nitrogen (NH_3-N). This change was prompted by a review applying the most recent version of *Standard Methods* (Rice et al. 2012), which is the basis for the sampling protocol adopted by the programme since 2006. Acidification was recommended to inhibit bacterial activity and prevent salt formation with organic bases. This report aims to inform on (i) the impact of this operational change on measured NH_3-N concentrations at NGMP sites and (ii) the implication of this change on the long-term NGMP dataset.

NGMP samples, including the new bottle, were collected by regional council staff as part of quarterly monitoring. Collected samples were then analysed at the New Zealand Geothermal Analytical laboratory for NH_3-N, one of the monitored parameters using acid-preserved bottles, and for other parameters, including major ions and trace metals. Duplicate analysis of NH_3-N (flow-injection analysis APHA 4500-NH3-H method) was performed on a subset of the unpreserved bottles for cost-effectiveness. The subset was selected to encompass a range of concentrations (<0.003 to 13.7 mg/L) and hydrogeological settings, while ensuring a dataset size to enable statistical testing. The dataset consisted of 251 matched pairs of NH_3-N concentrations (unpreserved and acid-preserved groundwater samples).

Overall, the distribution of concentrations was close to a 1:1 relationship. Identical pairs were observed in 30% of the dataset, and associated percentiles were very consistent between unpreserved and acid-preserved samples. Focusing on the matched pairs with differences, concentrations of acid-preserved samples were smaller than those of the unpreserved samples (Wilcoxon p-value = 0,001; confidence level of 95%). However, the magnitude of this difference fell within the uncertainty of measurement (estimated ±0.003 mg/L), and the calculated effect size was small (Wilcoxon effect size r = 0.29). These suggest that the acid preservation does not significantly impact NH_3-N measured concentrations at NGMP sites and that it is unlikely that the long-term dataset is impacted by the change of sample collection from 2016. Some of the samples recorded a laboratory arrival temperature and sample holding time higher than the threshold listed in *Standard Methods*. However, when sub-setting the matched-pair dataset based on these more stringent criteria, similar results were obtained for statistical testing.

Recommendations of this report are to:

- Accompany any change in operations with an analysis of potential impacts on the monitoring programme resources. For cost-effectiveness, this study supports reverting the NGMP sampling kit from acid-preserved to unpreserved samples for NH_3-N analysis.
- Accompany any operational change that may occur within a programme with data collection in order to assess the impact of this change on the long-term dataset. This study suggests that the NGMP sample collection modification had a negligible impact on the integrity of the long-term dataset and therefore no correction factor is needed.
- Conduct similar comparisons to assess the impact of arrival temperature and sample holding time on laboratory analysis in NGMP samples.

KEYWORDS

National Groundwater Monitoring Programme, groundwater quality, sampling protocol, sample preservation, ammonia-nitrogen dataset

This page left intentionally blank.

1.0 INTRODUCTION

1.1 Groundwater Quality Monitoring

Ammonia-nitrogen (NH_3-N) concentrations are associated with a toxicity threshold of 0.74 mg/L for biota and ecosystem protection (ANZECC and ARMCANZ 2000) and with an aesthetic guideline value of 1.5 mg/L (as NH_3) for drinking water (Ministry of Health 2018). In addition, under low oxygen conditions, NH_3-N is the natural end-product of the reduction of nitrate-nitrogen, a primary indicator of anthropogenic degradation of groundwater quality (Moreau and Daughney 2015).

NH_3-N concentrations are analysed routinely in groundwater samples as part of the National Groundwater Monitoring Programme (NGMP). The NGMP is a long-term research and groundwater monitoring programme operated by GNS Science in collaboration with 15 regional authorities since 1990. The primary aims of the NGMP are to:

1. Provide a long-term national perspective on groundwater quality in New Zealand, including determination of natural 'baseline' groundwater quality.
2. Identify spatial and temporal patterns in groundwater quality and associate them with certain causes, such as human influence.
3. Develop and convey best-practice methods for groundwater sampling, chemical analysis and interpretation.

Currently, the NGMP network consists of 110 monitoring sites across the country, which are sampled quarterly (March, June, September and December rounds). Sampling is guided by a national sampling protocol (Daughney et al. 2006), and groundwater is analysed for a suite of parameters, including major ions, silica, nutrients, selected dissolved metals, pH and electrical conductivity. NGMP groundwater quality data can be accessed publicly through the GNS Science Geothermal and Groundwater database.[1]

Operational changes are not uncommon in long-term monitoring programmes and can affect the network site selection, sampling frequency, analytical suites and methods and data capture and management (Moreau-Fournier and Daughney 2012; Milne et al. 2019). Triggers for operational changes may range from aspects managed within a programme (e.g. site access, budget) to global programme modifications (e.g. review of the programme objectives) or can result from changes outside the programme (e.g. laboratory closure). Operational changes have occurred within the NGMP, having impacts on the long-term dataset quality. For instance, samples were initially analysed at the Department of Scientific and Industrial Research in Christchurch and shifted to the New Zealand Geothermal Analytical Laboratory (NZGAL) upon its establishment in 1993. Another example is linked to the advances in analytical capability and techniques adopted by NZGAL over time (Table 1.1), resulting in an improvement of detection limit and a reduction in uncertainty of measurement.

1 https://ggw.gns.cri.nz/ggwdata/

Table 1.1 Overview of NH_3-N concentration analytical methods used since the National Groundwater Monitoring Programme started (updated from Moreau and Daughney 2021).

Analytical Method	Time Period
Flow-injection analysis (APHA 4500 NH3-H)	December 2013 – current
Flow-injection analysis (LACHAT QUICKCHEM 8500 Series 2 Method 31-107-06-1-B)	December 2012 – September 2013
Manual phenate method (APHA 4500-NH3 F)	September 2007 – September 2012
Phenol/hypochlorite colorimetry (APHA 4500-NH3 G)	March 2002 – June 2007
Automated phenohypochlorite method	September 1998 – December 2001
Ion chromatography	June 1998 – September 1998
Automated phenohypochlorite	June 1996 – March 1998
Absorption spectrophotometry	March 1990 – March 1996

1.2 National Groundwater Monitoring Programme Groundwater Sampling Protocol Review

In 2016, the NGMP groundwater sampling protocol (Daughney et al. 2006) was reviewed against, at the time, the most recent version of the *Standard Methods for the Examination of Water and Wastewater*[2] (Rice et al. 2012; referred to as '*Standard Methods*' onward). *Standard Methods* was first published in 1905 and is regularly peer-reviewed and updated by experts (currently on its 23rd edition). It is a comprehensive reference covering a wide range of water and wastewater analysis techniques, used routinely for environmental purposes (Baird et al. 2017).

Essential information regarding sample collection and dispatch compiled in the current NGMP sampling protocol is derived from *Standard Methods*. While cross-checking sampling requirements of the 2006 national protocol against this standard, differences were noted pertaining to the collection, dispatch and storage of groundwater samples for NH_3-N analysis (Table 1.2). The adjunction of sulphuric acid or hydrochloric acid was recommended as a preservative for the collection of water samples for acting as bacterial inhibitor and preventing salt formation with organic bases (Rice et al. 2012; Baird et al. 2017).

Table 1.2 Difference in sampling bottle type, preservation, transport and storage requirements between the 2006 national protocol (Daughney et al. 2006) and the 2012 edition of *Standard Methods* (Rice et al. 2012).

	2006 National Sampling Protocol	2012 *Standard Methods*
Sampling Bottle Type	Opaque plastic bottle for nutrients and bromide.	Plastic (polyethylene, polypropylene or similar), glass or fluoropolymer (e.g. teflon, PTFE).
Preservation and Transport	Chill to 4°C. No preservative requirements.	Cool below 6°C. Analyse as soon as possible or add H_2SO_4 to pH <2.
Maximum Storage	Analyse as soon as possible after collection.	Analyse 7 (recommendation) to 28 days (regulatory requirement) after collection.

2 A joint publication of the American Public Health Association (APHA), the American Water Works Association (AWWA) and the Water Environment Federation (WEF).

Prompted by this review, the NGMP sampling kit was modified by adding a 100 mL sulfuric-acidified sampling bottle for the analysis of NH_3-N. This change had implications on staff time (sampling, kit preparation) and in consumable cost, effective from May 2016 (Moreau and Cameron 2016).

1.2.1 Study Aim

This study follows up the review and aims to:

1. Assess whether the change in preservation requirement has a significant impact on measured NH3-N concentrations in NGMP samples.
2. Inform future NGMP operations with regard to groundwater sample collection requirements and implications for long-term dataset and data analysis (for the purpose of state and trend reporting), in line with the NGMP objectives.

The design, methodology and statistical tests used to investigate this impact are transferable to other operational changes. Recommendations made based on this study are intended only for the NGMP network, and considerations should be taken if transferring to other monitoring networks (e.g. considering differences in sampling protocols between the NGMP network and others). It is out of the scope of this report to assess if differences between parallel NH_3-N analysis are due to other factors rather than the use of acid-sulfuric preservative (i.e. changes due to different arrival temperatures).

This report includes a description of the graphical and statistical methods for comparing unpreserved and acid-preserved samples, followed by an assessment of the changes of NH_3-N concentrations due to the use of acid-preservative, and provides recommendations for NGMP monitoring operations and data analysis.

2.0 METHOD

2.1 Data Collection and Analysis

2.1.1 Groundwater Sample Collection and Preservation

Groundwater samples were collected nationwide by regional council staff in 2016 as part of routine NGMP sampling at 108 sites in the June round, 109 sites in the September round and 100 sites in the December round. Differences in the number of samples per round reflect cases where samples could not be collected; for example, due to pump failure or lack of access. At the time of collection and writing, the NGMP sampling kit (Figure 2.1) consisted of four Nalgene® sampling bottles for the following analysis:

1. Unfiltered and unpreserved for acidity and alkalinity (yellow label).
2. Field-filtered and unpreserved for anions (green label).
3. Field-filtered and nitric-acidified for cations, dissolved metals and silicon (pink label); sulphuric-acidified and filtered for NH_3-N (white label, bottle cap marked).

Figure 2.1 Photos of the NGMP sampling kit for chemistry prepared at NZGAL (left) and deployed in the field (right). The chemistry set consists of four Nalgene® sampling bottles: yellow label, green label, pink label and white label. Bottle caps for pre-acidified bottles are marked to prevent the swapping of bottle caps in the field. The additional seven bottles shown in the field (right) are collected for age-tracer analysis, undertaken at the GNS Science Water Dating Laboratory. The latter samples are not collected quarterly.

For this study, two bottles were used for NH_3-N analysis: the field-filtered and unpreserved bottle (Bottle 2, referred to as 'NH_3-N concentrations in unpreserved samples') and the field-filtered and sulfuric-acidified bottle (Bottle 4, referred to as 'NH_3-N concentrations in acid-preserved samples'). The latter consists of an acidified 100 ml Nalgene® bottle. Its addition was prompted by the 2016 review, and it has been used since then to enable the analysis of NH_3-N on a field-filtered and sulphuric-acid-preserved groundwater sample. NH_3-N analysis was previously undertaken on a field-filtered unpreserved sample (Bottle 2). Acidification is performed as part of the sampling kit preparation at NZGAL and consists of the adjunction of 0.2 mL of sulfuric acid to bring the sample pH below 2 (Rice et al. 2012).

2.1.2 Laboratory Analysis

Samples collected in the field were then dispatched in chilly bins to the NZGAL, accompanied by the required Chain of Custody record, on which arrival time, date and temperature were recorded by laboratory staff upon sample reception. Samples were then analysed for the regular NGMP suites (Table 2.1), for which NZGAL is accredited for.

Table 2.1 Current list of parameters monitored at National Groundwater Monitoring Programme (NGMP) sites.

Sample	Parameter	Detection Limit (mg/L)	NGMP Analytical Method
Unfiltered, unpreserved (Bottle1)	Alkalinity (as HCO3)	5.0	Titration, APHA 2320B
Field-filtered, unpreserved (Bottle 2)	Dissolved reactive phosphorus	0.002	Flow-Injection Analyser, APHA 4500-P G (modified)
	Chloride	0.05	Ion Chromatography, APHA 4110B
	Bromide	0.02	
	Fluoride	0.02	
	Nitrate (as N)	0.02	
	Sulphate	0.03	
Field-filtered, nitric-acidified (Bottle 3)	Calcium	0.01	Induced Coupled Plasma-Optical Emission Spectrometry, APHA 3120B
	Iron	0.01	
	Magnesium	0.01	
	Manganese	0.001	
	Potassium	0.11	
	Silica	0.05	
	Sodium	0.02	
Field-filtered, sulphuric-acidified (Bottle 4)	Ammonia (as N)	0.003	Flow-Injection Analysis, APHA 4500-NH_3-N

Although not currently released as part of the official NGMP analysis, NZGAL regularly calculates an uncertainty factor accompanying the laboratory measurements (Table 2.2), quantified using a statistical method (Milne 2019).

To create the duplicate analysis dataset, a selection from the 317 field-filtered and unpreserved samples were analysed by flow-injection analytical method (APHA 4500-NH3-H) for cost-effectiveness. The selection for duplicate analyses was made by laboratory staff in consultation with the NGMP manager to ensure that (i) a range of NH_3-N concentrations will be investigated, (ii) the dataset will be representative of a range of hydrogeological conditions across the country and (iii) the number of duplicate analyses is sufficient to allow for robust statistical analysis. In total, 255 sets of duplicate samples were analysed over the May to December 2016 sample collection.

Table 2.2 Uncertainty of measurements (%) for different ranges of ammonia concentrations (mg/L) calculated at the New Zealand Geothermal Analytical Laboratory (Sanderson 2019).

Concentration Range (mg/L)	Uncertainty of Measurement (%)	Example Concentration (mg/L)	Example Uncertainty of Measurement (mg/L)
0.003–0.007	37.3	0.004	0.001
0.007–0.2	6.4	0.10	0.037
0.2–20	3.3	1.0	0.033

2.2 Data Handling

2.2.1 Dataset Quality Check

As part of the NGMP Quality Assurance (QA) procedure, the integrity of the groundwater sample analysis was tested using either a charge balance error (CBE) or an ion sum. These tests are designed to ensure that the groundwater chemistry re-constructed from water analysis satisfies the electro-neutrality principle. The acceptance thresholds are: a CBE equal to or less than 5% or an absolute ion difference equal to or lower than 0.2 meq/L (Moreau-Fournier and Daughney 2010). These thresholds are consistent with the discrete groundwater quality National Environmental Monitoring Standards (NEMS; Milne et al. 2019).

After this QA procedure, four sets of samples were identified as invalid duplicate analyses (sample IDs 2016003799, 2016004149, 2016006427 and 2016006778). These analyses were not uploaded to the Geothermal and Groundwater database nor released as part of the NGMP dataset.

2.2.2 Creating a Matched-Pairs Dataset

Of the checked analysis, NH_3-N concentrations measured at the same site during the same sampling round from the two sets of unpreserved and acid-preserved bottles were assembled to create a matched-pairs dataset (n = 251).

2.2.3 Graphical Examination, Statistical Test Selection and Implementation

To compare the data distribution and illustrate the relationship between matched pairs, scatter plots and boxplots were generated. These graphical examinations were complemented by the following descriptive statistics: percentiles (5^{th}, 25^{th}, median, 75^{th}, 95^{th}), mean, standard deviation and median absolute deviation (MAD).

To test if the difference in medians between unpreserved and acid-preserved samples was equal to zero, the non-parametric Wilcoxon signed-rank test was applied. To focus only on the differences, tied values (i.e. concentration is the same for both unpreserved and acid-preserved samples) were excluded. The signed-rank test assumes that the distribution of the differences is symmetric (not necessarily normal) and was preferred over other non-parametric tests (e.g. the sign test) for being more powerful in detecting differences (Helsel et al. 2020). To fulfil the assumption of symmetry, the matched-pairs dataset was log-transformed.

The signed-rank test is a hypothesis test whereby a probability value (p-value) is calculated to inform whether a difference between paired observations exists and is commonly used to draw a conclusion on its size (Helsel et al. 2020). However, the p-value does not inform on the magnitude of the difference. For this reason, the Wilcoxon effect size was calculated by

dividing the statistic obtained from the signed-rank test (Z value) by the square root of N (number of matched pairs, excluding tied values; Tomczak and Tomczak [2014]). In this study, a confidence level of 95% was used to assess the p-value of the test. This level is commonly used in statistics (Helsel et al. 2020).

The Wilcoxon signed-rank test was performed using the R software (version 3.6.2), using the NADA2 library (version 1.0.1; Julian and Helsel [2021]), as the NH_3-N dataset includes censored values (i.e. values reported below the detection limit). Since the maximum censoring level is 24%, median, MAD and percentiles were estimated using Regression on Order Statistics (ROS) models (Helsel et al. 2020).

3.0 RESULTS AND DISCUSSION

The cleaned matched-pairs dataset is included as a digital attachment to this report (GNS SR2021-27 Data Output.xlsx). Provided information consists of the following attributes: batch and sample ID, site ID, sample collection date, sample arrival date, sample analysis date, sample arrival temperature, NH_3-N concentrations for unpreserved samples and NH_3-N concentrations for acid-preserved samples.

3.1 Data Overview and Graphical Examination of NH_3-N Matched Pairs

The matched-pairs dataset consists of 251 pairs of NH_3-N concentrations (unpreserved and acid-preserved samples) collected at 112 NGMP sites nationwide (Figure 3.1). Within this dataset, sites were either sampled in one (13 sites), two (58 sites) or three rounds (40 sites). The fact that some sites have been sampled more than others does not affect the design of the test applied, as the aim of the comparison is to assess changes in concentrations due to different preservation methods rather than between sites.

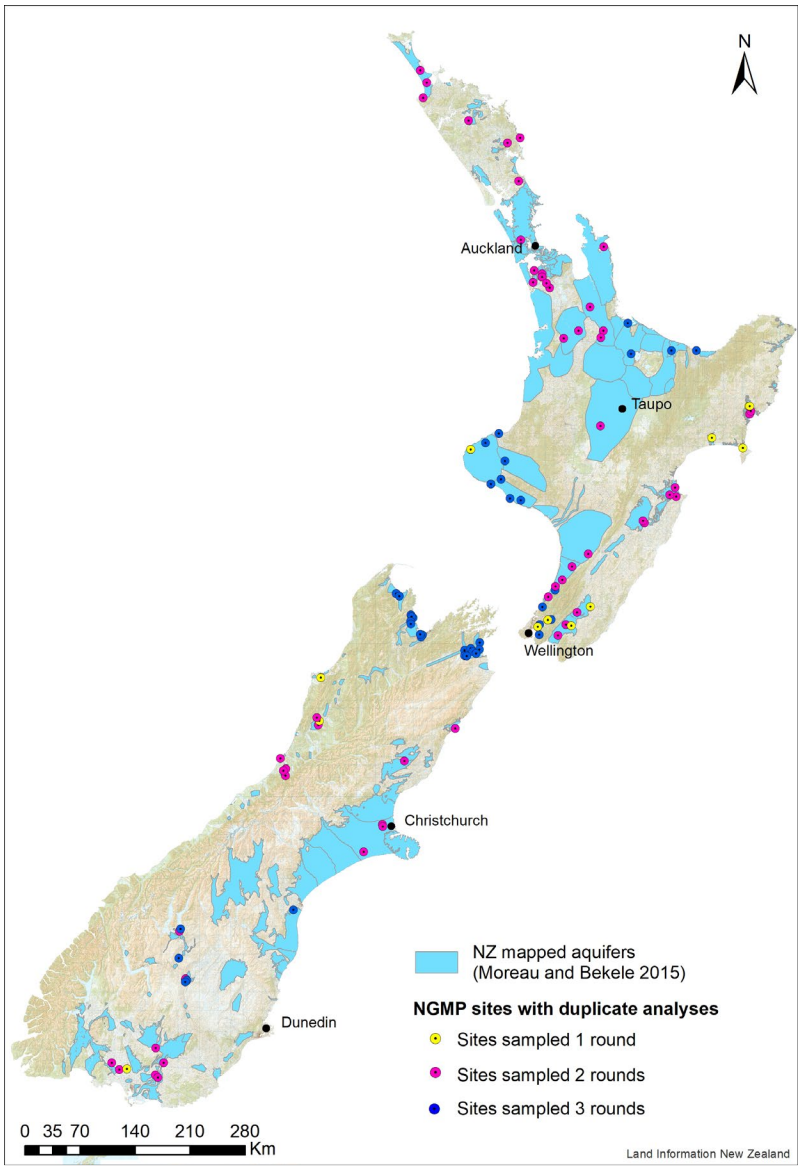

Figure 3.1 Matched-pairs sample locations. The different colours show the number of rounds that each site was sampled (yellow: one round; pink: two rounds; dark blue: three rounds). Sample locations follow the distribution of New Zealand mapped aquifers, represented by the light blue areas (Moreau and Bekele 2017).

Sample arrival temperature and sample holding time can have an influence on results of chemical analyses due to physico-chemical processes that may occur in the sampling bottles. For NH_3-N samples, aside from the adjunction of acid, the current *Standard Methods* (Baird et al. 2017) includes criteria regarding sample dispatch (i.e. samples are to be kept in the dark and below 6°C) and sample holding time (i.e. analyses should be completed within 28 days). The *Standard Methods* criteria are slightly more stringent than the NEMS for Discrete Water Quality (Milne et al. 2019). The latter includes criteria for arrival temperature (below 10°C) but not on sampling holding time.

For the paired dataset, the maximum arrival temperature and sample holding time were 13°C and 56 days, respectively. Following the *Standard Methods* criteria, of the 251 matched pairs, 185 (73.7%) met the sample arrival temperature and holding time. This subset is referred to as 'valid matched pairs' for the remainder of this report (Table 3.1). If only considering the NEMS valid arrival temperature criteria, the dataset will consist of 224 matched pairs, compared to 205 matched pairs when considering the *Standard Methods* valid arrival temperature criteria.

Both the full dataset and valid subset exhibited a similar range of NH_3-N concentrations (ranging from below the detection limit to 13.7 mg/L) and statistical distribution (Table 3.1). For example, NH_3-N concentration median and MAD were 0.007 ± 0.006 mg/L, respectively, for the unpreserved and acid-preserved samples for both the matched pairs and the valid matched-pairs subset (Table 3.1). Tied values represented 30% and 34% of the matched-pairs dataset and the valid matched pairs subset, respectively.

Table 3.1 Descriptive statistics for NH3-N concentrations (mg/L) for the matched pairs and valid matched pairs.

	All Matched Pairs		Valid Matched Pairs	
	Unpreserved Samples	**Acid-Preserved Samples**	**Unpreserved Samples**	**Acid-Preserved Samples**
% censoring	20.72	22.31	24.32	23.78
n	251		185	
Tied Values	75 (30%)		62 (34%)	
Average (mean)	0.36	0.34	0.33	0.31
Standard Deviation	1.42	1.37	1.28	1.23
Minimum	<0.003	<0.003	<0.003	<0.003
5th Percentile	0.0001*	0.00009*	0.00009*	0.00008*
25th Percentile	0.004	0.003	0.003	0.003
Median	0.007	0.007	0.007	0.007
75th Percentile	0.075	0.085	0.060	0.050
95th Percentile	2.35	2.05	1.46	1.44
Maximum	13.7	13.2	13.7	13.2
Median Absolute Deviation	0.006	0.006	0.006	0.006
Quartile Skew Value	0.90	0.90	0.86	0.83

* Values below detection limit have been modelled using an ROS model.

Graphical visualisation of the data confirmed the expected strong correlation between NH_3-N measurements in unpreserved and acid-preserved samples (R^2 = 0.97 and 0.98 for all matched pairs and the valid matched pairs subset, respectively), with a greater scatter around low concentrations close to or below the detection limit (Figures 3.2 and 3.3). This is not surprising considering the higher uncertainty of measurement for values close to the detection limit in comparison to values far above it (37.7% for a concentration range of 0.003–0.007 mg/L; 6.4% for 0.007–0.2 mg/L and 3.3% for 0.2–20 mg/L; Sanderson 2019).

Comparison of the NH_3-N matched pairs identified that, in 30% of the pairs, one of the concentrations has been reported below detection limit (i.e. a censored value). Censored values were paired with a numerical value in 54% of the pairs, with censored data for both the total dataset and valid subset. The maximum value paired with censored data (i.e. lower than 0.003 mg/L) corresponds to 0.02 mg/L. Most of the censored-value matched pairs present values in the range of 0.003–0.01 mg/L, values which are close to the detection limit when considering the uncertainty of measurement (e.g. 0.004 ± 0.001 mg/L; Sanderson 2019).

The data distribution of unpreserved and acid-preserved samples is also similar when comparing box plots (i.e. median and 75th percentile; Figure 3.3). Both datasets exhibit positive-skew distributions, consistent with the mean being larger than the median and with positive quartile skew values (Table 3.1; Helsel et al. 2020).

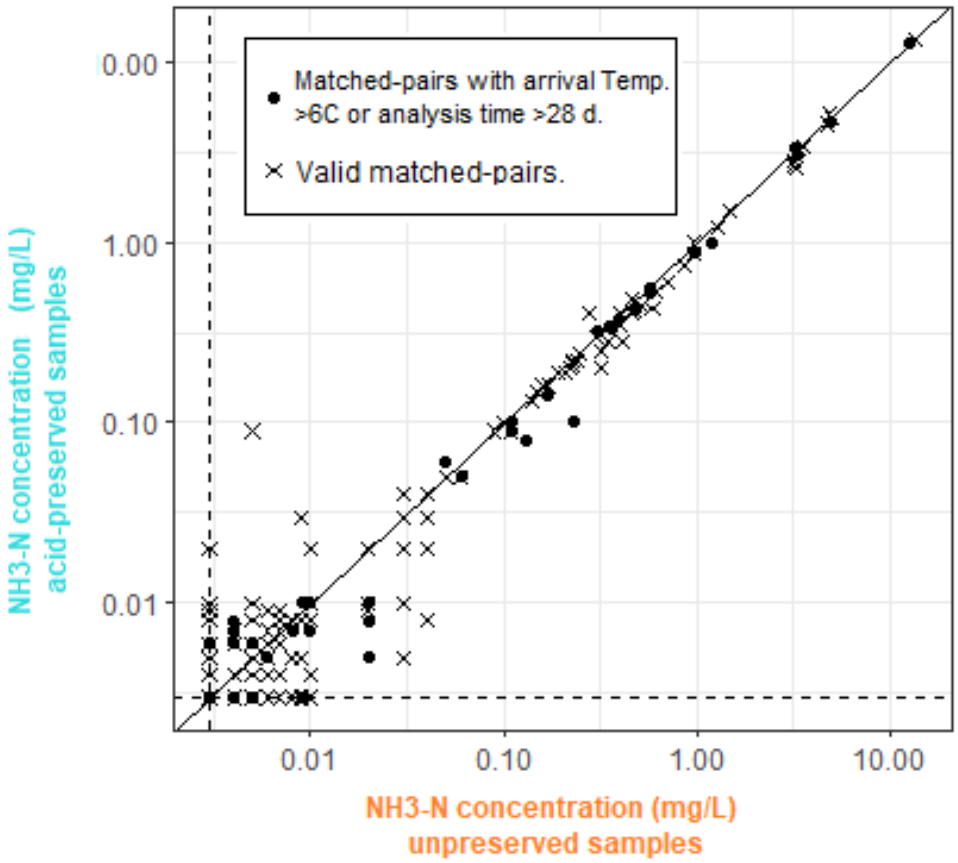

Figure 3.2 Scatter plot for matched pairs showing the relationship between unpreserved and acid-preserved samples for different arrival temperatures and analysis times. The figure shows that most samples fall close to the 1:1 line (black line). The black dashed lines represent the detection limit of NH3-N (<0.003 mg/L). A logarithmic scale was adopted to highlight measurements close to the detection limit.

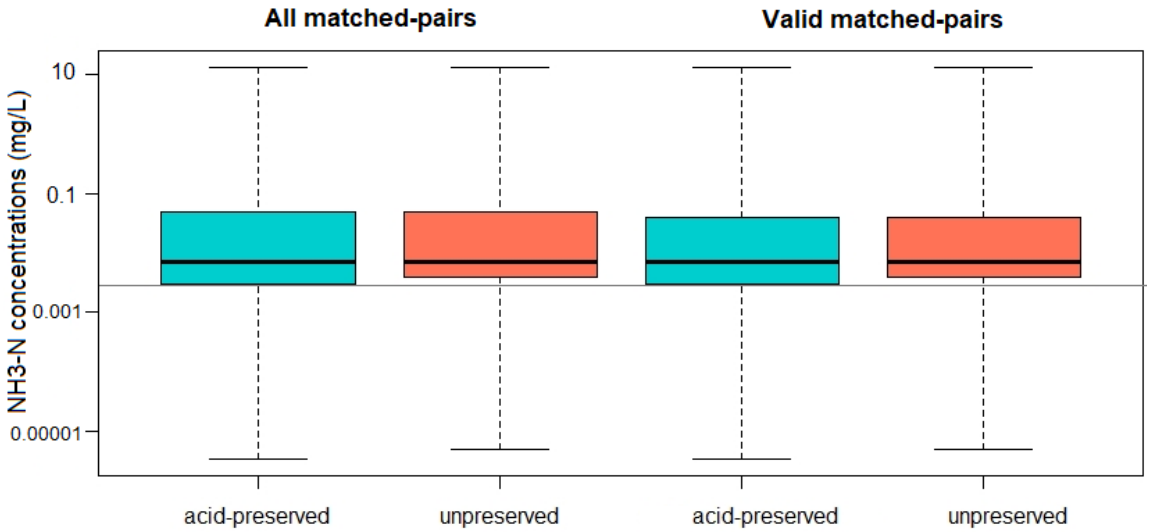

Figure 3.3 Box plots of all matched pairs for unpreserved and acid-preserved samples. The grey dashed line represents the detection limit (<0.003 mg/L). The median (50th percentile) marks the mid-point of the data and is shown by the line inside the box.

3.2 Testing the Differences between Acid-Preserved and Unpreserved Samples

3.2.1 Full Matched-Pairs Dataset

When isolating paired observations with measurement differences (Figure 3.4), the NH_3-N concentration medians for the unpreserved samples was greater than that of acid-preserved samples (Wilcoxon p-value = 0.001; median concentrations of 0.009 and 0.008 mg/L, respectively). However, the magnitude of this difference fell within the measurement uncertainty, and the Wilcoxon effect size was small (r = 0.29) suggesting that, despite the statistical test results, the difference in NH_3-N concentration was insignificant. The latter is consistent with descriptive statistics and with the median differences being equal to zero when including tied values (Table 3.1).

3.2.2 Valid Matched-Pairs Subset

When isolating paired observations with measurement differences in the valid subset (Figure 3.4), the NH_3-N concentration median for the unpreserved samples was also greater than that of acid-preserved samples (Wilcoxon p-value = 0.001; median concentrations of 0.009 and 0.008 mg/L, respectively). However, the magnitude of this difference fell within the measurement uncertainty and the Wilcoxon effect size was small (r = 0.29). This suggests that the acid-preservative effect was not significant, despite the statistical test results. The small effect size is consistent with the strong similarity observed between the unpreserved and acid-preserved sample descriptive statistics when including tied values (Table 3.1). In addition, the median of the differences between unpreserved and acid-preserved samples is equal to 0, another indicator of the similarity of the valid matched pairs when including ties.

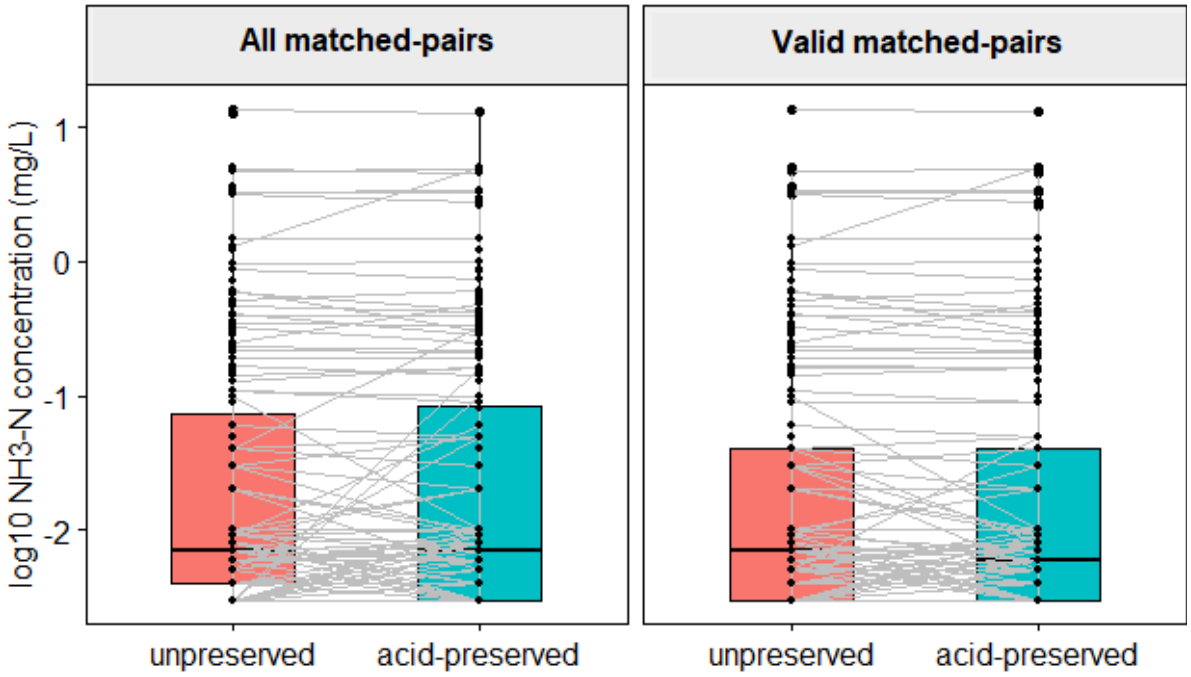

Figure 3.4 Box plots of NH3-N concentrations for paired data for the full dataset (left) and valid subset (right). Grey lines are drawn between the same sample.

3.2.3 General Observations

The graphical examination and descriptive statistics do not indicate a significant difference in NH_3-N concentrations between unpreserved and acid-preserved samples. However, the statistical test indicates a systematic difference between both matched pairs. This is not unexpected, as it reflects a change in sample preservation. For instance, NEMS states that 'step' changes may arise when a change is made to measuring methods or instruments (Milne 2019). The magnitude of this difference falls within the analytical uncertainty of the measurements, suggesting that the change in sample preservation had a negligible impact on the integrity of the long-term dataset and that no correction factor is needed.

The graphical examination (Figure 3.3), descriptive statistics and statistical test indicate that the differences between the valid subset and the matched-pairs full dataset are insignificant. This suggests that the *Standard Methods* thresholds for sample arrival temperature and holding time may not have a significant effect on NH_3-N measurements at NGMP sites.

4.0 CONCLUSIONS AND RECOMMENDATIONS FOR NATIONAL GROUNDWATER MONITORING PROGRAMME NETWORK

The aim of this study was to inform future NGMP operations regarding groundwater sample collection and operational changes within the programme. This was done through graphical examination and statistical analysis and testing of a duplicated NH_3-N concentration dataset, collected and analysed in 2016. The difference in NH_3-N concentrations observed between unpreserved and acid-preserved samples was found to be negligible. Also, the comparison between the valid subset and full matched-pairs dataset suggests that the thresholds for sample arrival temperature and holding times may not have a significant effect on NH_3-N measurements at NGMP sites.

Based on these results, the following recommendations to address operational changes within the NGMP are to:

- Accompany any change in operations with an analysis of potential impacts on the programme resources (e.g. both sampling and lab staff time). As a practical example, this study supports reverting the NGMP sampling kit to prior 2016, from acid-preserved to unpreserved samples for NH_3-N analysis. This will reduce time and cost of operations in both the field and the laboratory.
- Accompany any operational change in data collection with an assessment of its impact on the long-term dataset if data is to be considered for state and trend analysis (as is the case for the NGMP). This study suggests that the change in sample preservation in NGMP samples had a negligible impact on the integrity of the long-term dataset and therefore no correction factor is needed to account for this.
- Apply the design, methodology and statistical tests used in this study to other operational changes that may occur in the future.
- Continue recording sample arrival temperature and undertake timely quality checks to reduce sample holding time. This study highlighted a range of sample holding times and arrival temperatures in the recently collected NGMP samples. Work has already been initiated to reduce sample holding time in the form of automation of the quality assurance / quality control process and data upload.
- Consider data collection to inform on:
    - The effect of arrival temperature. This will require the collection of duplicate samples at the same site and date, with one chilled and the other not. For efficiency, this testing may be extended to the full analytical suite. The sample size should be large enough to allow for statistical testing on the differences.
    - The effect of sample holding time. This will require repeated analysis of a randomly selected batch of samples at various time intervals (e.g. 3, 18 and 56 days after the sampling date). The sample size should be large enough to allow for statistical testing on differences and, like above, the duplicate testing extended to the full analytical suite. The sample holding time assessment undertaken for nitrate concentrations (Moreau-Fournier 2010) can be used as a model.

5.0 ACKNOWLEDGMENTS

The authors would like to thank the following council staff for collecting the groundwater samples that were used in this study in particular (ordered by Council): Kristi Holland and Luke Stanley (Auckland Council); Amy Platt (Bay of Plenty Regional Council); Ross Cressy and David Evans (Environment Canterbury); Wendy Barrow and Darren May (Environment Southland); Peter Hancock (Gisborne District Council); Kasey Pitt (Greater Wellington Regional Council); Jordan Ellmers, Steve Grant, Paul Hodgkinson and Matt Parker (Hawkes Bay Regional Council); Steve Packer (Horizons Regional Council); Pete Davidson (Marlborough District Council); Ananda Hapu and Sandrine Le Gars (Northland Regional Council); Michael Anderson and Stephen Smith (Otago Regional Council); Thomas McElroy and David Olson (Taranaki Regional Council); Brenda Clapp, Joseph Thomas, Kieran Scott and Rob Merrilees (Tasman District Council); John Hughey (Waikato Regional Council); Ashton Eaves and Jonny Horrox (West Coast Regional Council) and anonymous regional council samplers. Moya Appleby and Ann Noddings (NZGAL) for providing analysis. Frederika Mourot (Groundwater Scientist, GNS Science), Matthew Coble (Isotope Hydrogeologist, GNS Science) and Stewart Cameron (Hydrogeologist; Hydrogeology & Geophysics & Modelling Team Leader, GNS Science) are also acknowledged for their review of this report.

This project was funded by the Strategic Science Investment – Infrastructure funds (Nationally Significant Databases and Collections) from the Ministry for Business, Innovation & Employment.

6.0 REFERENCES

[ANZECC and ARMCANZ] Australian and New Zealand Environment and Conservation Council, Agriculture and Resource Management Council of Australia and New Zealand. 2000. Australian and New Zealand guidelines for fresh and marine water quality. Volume 1: the guidelines. Canberra (AU): ANZECC, ARMCANZ. 1 vol. National Water Quality Management Strategy Paper 4.

Baird R, Eaton AD, Rice EW, Bridgewater L, editors. 2017. Standard methods for the examination of water and wastewater. 23rd ed. Washington (DC): American Public Health Association, American Water Works Association, Water Environment Federation. 1 vol.

Daughney CJ, Jones A, Baker T, Hanson C, Davidson P, Zemansky GM, Reeves RR, Thompson M. 2006. A national protocol for state of the environment groundwater sampling in New Zealand. Wellington (NZ): Ministry for the Environment. 54 p. (GNS Science miscellaneous series; 5).

[GGW Database] Geothermal and Groundwater Database. 2005–. Lower Hutt (NZ): GNS Science. https://doi.org/10.21420/9QTA-4884

Helsel DR, Hirsch RM, Ryberg KR, Archfield SA, Gilroy EJ. 2020. Statistical methods in water resources. In: *Hydrologic analysis and interpretation*. Reston (VA): U.S. Geological Survey. 458 p. (Techniques and methods; 4-A3).

Julian P, Helsel D. 2021. NADA2: data analysis for censored environmental data. Version 1.0.0; [accessed 2022 May]. https://CRAN.R-project.org/package=NADA2

Milne J. 2019. Water quality part 1 – sampling, measuring, processing and archiving of discrete groundwater quality data. [Place unknown]: National Environmental Monitoring Standards; [accessed 2022 May]. https://www.nems.org.nz/documents/water-quality-part-1-groundwater/

Ministry of Health. 2018. Drinking-water Standards for New Zealand 2005 (Revised 2018). Wellington (NZ): Ministry of Health. 120 p.

Moreau M, Bekele M. 2017. Groundwater component of the water physical stock account (WPSA). Wairakei (NZ): GNS Science. 31 p. Consultancy Report 2014/290. Prepared for the Ministry for the Environment.

Moreau M, Cameron SG. 2016. Outputs from GNS Science Hydrogeology Research programmes. Wairakei (NZ): GNS Science. 24 p. Consultancy Report 2016/150LR.

Moreau M, Daughney CJ. 2015. Update of national groundwater quality indicators: state and trends, December 2004–2013. Wairakei (NZ): GNS Science 32 p. Consultancy Report 2015/16. Prepared for Ministry for the Environment.

Moreau M, Daughney C. 2021. Defining natural baselines for rates of change in New Zealand's groundwater quality: dealing with incomplete or disparate datasets, accounting for impacted sites, and merging into state of the-environment reporting. *Science of The Total Environment*. 755(2):143292. doi:10.1016/j.scitotenv.2020.143292.

Moreau-Fournier M. 2010. Comparison of nitrate analysis results between two laboratories and effects of sample holding time on nitrate concentration analyses. Lower Hutt (NZ): GNS Science. 26 p. (GNS Science report; 2010/14).

Moreau-Fournier M, Daughney CJ. 2010. Procedure for checking laboratory water chemistry results prior upload to the Geothermal-Groundwater database. Wairakei (NZ): GNS Science. 62 p. Internal Report 2010/06.

Moreau-Fournier M, Daughney CJ. 2012. Dynamic groundwater monitoring networks: a manageable method for reviewing sampling frequency. *Journal of Environmental Monitoring*. 14(12):3129–3136. doi:10.1039/c2em30624g.

Rice EW, Baird RB, Eaton AD, Clesceri LS, editors. 2012. Standard methods for the examination of water and wastewater. 22nd ed. Washington (DC): American Public Health Association, American Water Works Association, Water Environment Federation. 1 vol.

Sanderson S. 2019. Personal communication. Laboratory Manager, New Zealand Geothermal Analytical Laboratory. Wairakei, NZ.

Tomczak M, Tomczak E. 2014. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*. 1(21):19–25.

www.gns.cri.nz

Principal Location

1 Fairway Drive, Avalon
Lower Hutt 5010
PO Box 30368
Lower Hutt 5040
New Zealand
T +64-4-570 1444
F +64-4-570 4600

Other Locations

Dunedin Research Centre
764 Cumberland Street
Private Bag 1930
Dunedin 9054
New Zealand
T +64-3-477 4050
F +64-3-477 5232

Wairakei Research Centre
114 Karetoto Road
Private Bag 2000
Taupo 3352
New Zealand
T +64-7-374 8211
F +64-7-374 8199

National Isotope Centre
30 Gracefield Road
PO Box 30368
Lower Hutt 5040
New Zealand
T +64-4-570 1444
F +64-4-570 4657