



Supporting Information for “Integrating Bayesian groundwater mixing modeling with on-site helium analysis to identify unknown water sources”

Andrea L. Popp^{1,2}, Andreas Scheidegger³, Christian Moeck¹, Matthias S.

Brennwald¹, Rolf Kipfer^{1,2,4}

¹Department Water Resources and Drinking Water, Eawag, Swiss Federal Institute of Aquatic Science and Technology, 8600

Dübendorf, Switzerland

²Department of Environmental Systems Science, ETH Zurich, 8000 Zurich, Switzerland

³Department Systems Analysis, Integrated Assessment and Modelling, Eawag, 8600 Dübendorf, Switzerland

⁴Department of Earth Sciences, ETH Zurich, 8000 Zurich, Switzerland

Contents of this file

1. Texts S1 to S3
2. Figures S1 to S8
3. Tables S1 to S2

Additional Supporting Information (files uploaded separately)

1. Datasets S1 to S2
2. Source Code S1

Corresponding author: A. L. Popp (andrea.popp@eawag.ch)

Introduction

This Supplementary Material includes

- Text S1 explaining why we did not use the other obtained dissolved gas data in this study.

- Text S2 describing on-site and lab-based ^4He analyses.

- Text S3 explaining the likelihood function and Bayesian inference of the model we present.

- Figure S1 showing examples of bivariate tracer-tracer plots to graphically support the description of the Introduction (manuscript).

- Figure S2 showing a simplified geological map of the study site with the main bedrock units.

- Figure S3 illustrating the estimated fractions and uncertainties of end-members based on the model results (Dataset S2) using tracer set *S7* (Table 1).

- Figure S4 showing the results of the sensitivity test with labels on all data points.

- Figures S5 to S7 showing estimated mixing ratios at all investigated wells calculated for tracer sets 1, 4 and 7.

- Figure 8 comparing model-estimated and measured tracer concentrations.

- Table S1 containing the data on which Figure 4 (manuscript) is based on.

- Table S2 describing the instruments used, parameter units, LOQ and measurement uncertainties for the major ion data and other hydro-chemical parameters.

Datasets S1 (containing information about sampling locations and all tracer input data) and S2 (model results) as well as the model Source Code S1 are available online: <https://doi.org/10.25678/000183>.

Text S1. Other (noble) gases analyzed with the GE-MIMS could not be used to estimate mixing ratios since they were either not conservative (N_2 , O_2) or did not yield distinctive differences in their concentrations between end-members (^{40}Ar , ^{84}Kr).

Text S2.1: On-Site ^4He Analysis

With the GE-MIMS system operated in the field, we analyzed partial pressures of ^4He together with a set of other (noble) gases (N_2 , O_2 , Ar , Kr , not further discussed, see Text S1). The abstracted water was pumped through a membrane module (3M Liqui-Cel, 2017), where the dissolved gases are extracted into a head space until a gas-equilibrium between the dissolved and the free gas phase is established (Brennwald et al., 2016). The solubility equilibrium guarantees that the partial pressures of the analyzed gas species are proportional to the concentrations in water. The module is connected via a small capillary to a Quadrupole Mass Spectrometer for final gas analysis. Air-water equilibrium within the membrane module is reached after ~ 15 minutes with a determined water flow rate of about 1.5 to 1.8 L/min.

With a second capillary we sampled for ambient air, which we used as standard. Gas calibration was then performed by comparing peak heights between ambient air and the sampled gases. Each measurement cycle (including all aforementioned gas species) for either an air standard or dissolved gas measurement takes about 8–9 minutes. We conducted four cycles per well and two air standards (as first and last measurement), which results in an overall sampling time with the GE-MIMS of approximately 60 minutes at

each well (including about 2 minutes waiting time to purge the capillary after switching from a water sample to an air sample).

Water temperature was continuously recorded before the membrane module (temperature probe DS18B20 Maxim) allowing the conversion of the determined gas partial pressures into dissolved gas concentrations according to Henry's law solubility constants at the respective water temperature assuming zero salinity (Kipfer et al., 2002). For a more comprehensive description of gas analysis using the GE-MIMS system, see Brennwald et al. (2016).

Text S2.2: Laboratory-Based ^4He Analysis

After the GE-MIMS analysis we collected samples for laboratory-based noble gas analysis as water samples in copper tubes sealed with pinch-off clamps following standard procedure (Beyerle et al., 2000).

We analyzed ^4He , ^3He and other gases (not further discussed) at the ETH noble gas lab in Zurich, as described by Beyerle et al. (1999). The gas free copper tubes were stored to allow for ^3He in-growth for later tritium analysis (Tolstikhin & Kamenskiy, 1969). Tritium, however, turned out to be an unsuitable tracer to analyze water mixing at this study site due to the unsteady tritium release from several nuclear power plants located upstream of the study area. Due to the highly variable input signal of tritium concentrations into our groundwater system, we consequently excluded tritium as a tracer.

Text S3: Likelihood function and Bayesian inference

The introduction of the observation model and the residual end-member considerably increase the number of parameters to be estimated. However, by using weak and intuitive prior assumptions, Bayesian inference still guarantees the parameter identifiability.

To derive the required likelihood function, we first introduce a vector notation. In all equations below we assume that the tracer substances are indexed with $t = 1 \dots T$, the selected end-members with $m = 1 \dots M$, and the wells with $w = 1 \dots W$. Let $\mathbf{r}[w] = (r_1[w], \dots, r_{M+1}[w])$ be the vector of all mixing ratios for wells w , and $\mathbf{r} = (\mathbf{r}[1], \dots, \mathbf{r}[W])$ the set of all mixing ratios. Similarly, the end-member concentrations are $\mathbf{C}_E = (\mathbf{C}_E[1], \dots, \mathbf{C}_E[T])$ with $\mathbf{C}_E[t] = (C_{E_1}[t], \dots, C_{E_M}[t])$ and the residual end-member concentrations are $\mathbf{C}_{\text{res}} = (\mathbf{C}_{\text{res}}[1], \dots, \mathbf{C}_{\text{res}}[W])$ with $\mathbf{C}_{\text{res}}[w] = (C_{E_{\text{res}}}[w, 1], \dots, C_{E_{\text{res}}}[w, T])$ at wells w . The concentrations of the non-end-member wells (i.e., mixtures) are summarized as $\mathbf{C} = (\mathbf{C}[1], \dots, \mathbf{C}[W])$ with $\mathbf{C}[w] = (C[w, 1], \dots, C[w, T])$.

The observation model and the mixing model (Equation 2) specify the likelihood function for the mixture concentrations. The parameters are the mixing ratios, but also the (true) end-member concentrations, and the residual end-member concentrations:

$$p(C^{\text{obs}}[w, t] \mid C[w, t]) = p(C^{\text{obs}}[w, t] \mid \mathbf{C}_E[t], \mathbf{C}_{\text{res}}[w, t], \mathbf{r}[w]).$$

If we further assume that the observation errors are independent, we can write

$$p(\mathbf{C}^{\text{obs}} \mid \mathbf{C}_E, \mathbf{C}_{\text{res}}, \mathbf{r}) = \prod_{w=1}^W \prod_{t=1}^T p(C^{\text{obs}}[t, w] \mid \mathbf{C}_E[t], \mathbf{C}_{\text{res}}[t, w], \mathbf{r}[w]).$$

By combining this result with the observation model for end-members we obtain the complete likelihood function:

$$p(\mathbf{C}^{\text{obs}}, \mathbf{C}_E^{\text{obs}} \mid \mathbf{C}) = p(\mathbf{C}^{\text{obs}} \mid \mathbf{C}_E, \mathbf{C}_{\text{res}}, \mathbf{r}) p(\mathbf{C}_E^{\text{obs}} \mid \mathbf{C}_E).$$

The parameters $\{\mathbf{r}, \mathbf{C}_E, \mathbf{C}_{\text{res}}\}$ are not identifiable in a frequentist maximum likelihood setting. Therefore, we define the following prior distributions: (i) a non-informative, flat prior $p(\mathbf{C}_E) = U(0, \infty)$ for the true end-member concentrations, (ii) a Dirichlet distribution $p(\mathbf{r}[w]) = \text{Dirichlet}(1, \dots, 1)$ for the mixing ratios, which defines an uninformative

simplex that guarantees $\sum_{m=1}^{M+1} r_m = 1$ and $0 \leq r_m$, $m = 1, \dots, M + 1$ (the same choice was made by Delsman, Oude Essink, Beven, and Stuyfzand (2013)), and (iii) an informative prior for the residual end-members. For the latter we used uniform distributions with lower and upper limits selected so that they are 20% less/more extreme than any observed tracer concentrations.

With these prior distributions and the likelihood function, we define the posterior distribution as

$$p(\mathbf{C}_E, \mathbf{C}_{\text{res}}, \mathbf{r} \mid \mathbf{C}^{\text{obs}}, \mathbf{C}_E^{\text{obs}}) \propto p(\mathbf{C}^{\text{obs}}, \mathbf{C}_E^{\text{obs}} \mid \mathbf{C}_E, \mathbf{C}_{\text{res}}, \mathbf{r}) p(\mathbf{C}_E) p(\mathbf{r}) p(\mathbf{C}_{\text{res}}). \quad (1)$$

The proportional relationship is sufficient to sample from this distribution, for example with Monte Carlo Markov Chain methods (Kruschke, 2015).

References

- 3M Liqui-Cel. (2017). *A high level of consistency and control*. Wuppertal, Germany.
- Beyerle, U., Aeschbach-Hertig, W., Hofer, M., Imboden, D. M., Baur, H., & Kipfer, R. (1999). Infiltration of river water to a shallow aquifer investigated with $^3\text{H}/^3\text{He}$, noble gases and CFCs. *J. Hydrol.*, *220*(3-4), 169–185. doi: 10.1016/S0022-1694(99)00069-4
- Beyerle, U., Aeschbach-Hertig, W., Imboden, D., Baur, H., Graf, T., & Kipfer, R. (2000). A Mass Spectrometric System for the Analysis of Noble Gases and Tritium from Water Samples. *Environ. Sci. Technol.*, *34*(10), 2042–2050. doi: 10.1021/es990840h
- Brennwald, M. S., Schmidt, M., Oser, J., & Kipfer, R. (2016). A Portable and Autonomous Mass Spectrometric System for On-Site Environmental Gas Analysis. *Environ. Sci. Technol.*, *50*(24), 13455–13463. doi: 10.1021/acs.est.6b03669
- Delsman, J. R., Oude Essink, G. H., Beven, K. J., & Stuyfzand, P. J. (2013). Uncertainty

- estimation of end-member mixing using generalized likelihood uncertainty estimation (GLUE), applied in a lowland catchment. *Water Resour. Res.*, *49*(8), 4792–4806. doi: 10.1002/wrcr.20341
- Kipfer, R., Aeschbach-Hertig, W., Peeters, F., & Stute, M. (2002). Noble Gases in Lakes and Ground Waters. *Rev. Miner. Geochemistry*, *47*(1), 615–700. doi: 10.2138/rmg.2002.47.14
- Kruschke, J. K. (2015). Markov Chain Monte Carlo. In *Doing Bayesian Data Analysis* (2nd ed., pp. 143–191). Academic Press. doi: 10.1016/B978-0-12-405888-0.00007-6
- Tolstikhin, I., & Kamenskiy, I. (1969). Determination of Groundwater Ages by the T-³He Method. *Geochemistry Int.*, *6*, 810–811.

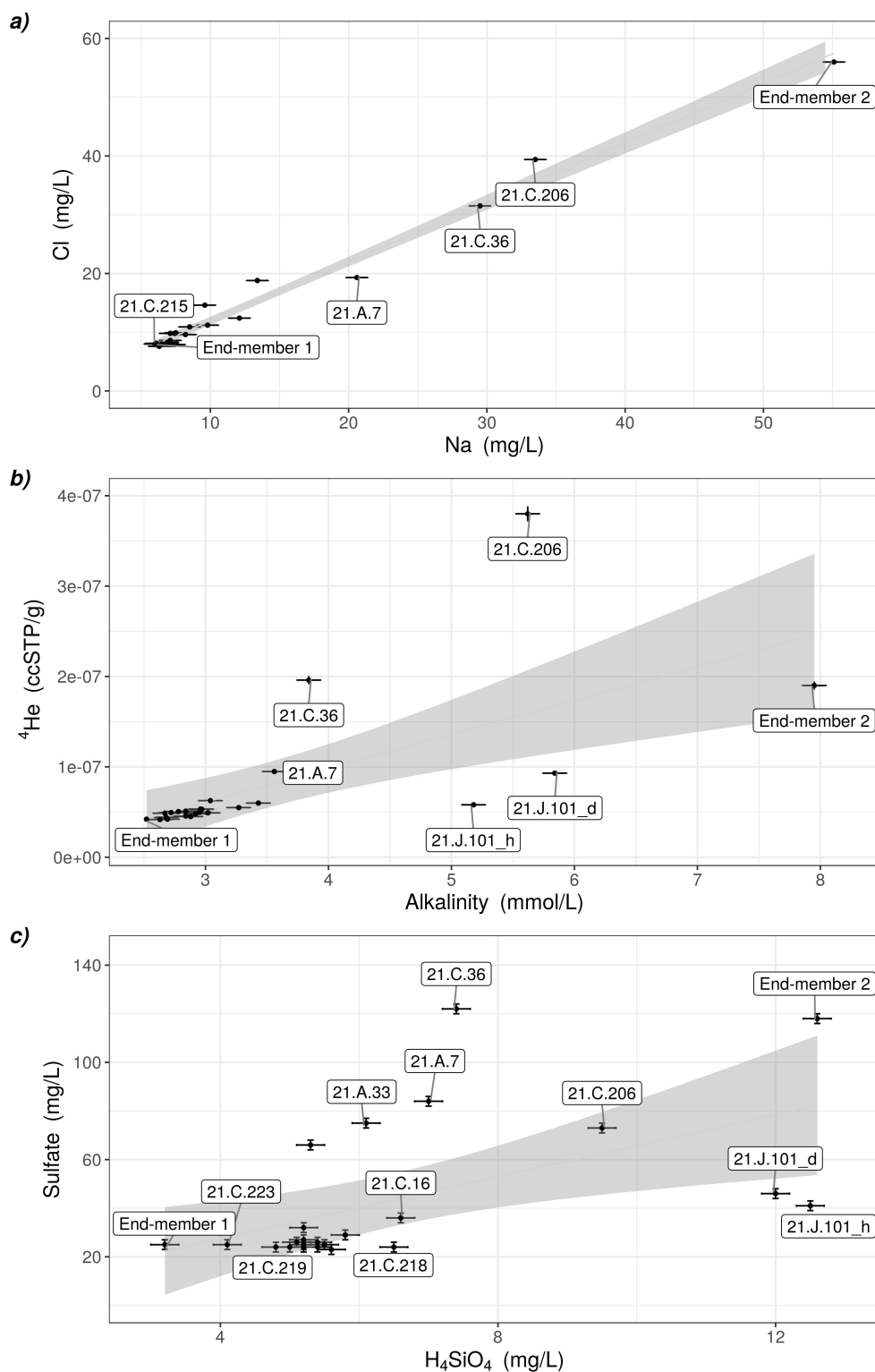


Figure S1. Exemplary bivariate tracer-tracer plot with a) showing sodium (Na) vs. chloride (Cl), b) showing alkalinity vs. helium (^4He) and c) showing silica (H_4SiO_4) vs. sulfate for the two pre-selected end-members and the sampled mixtures. The gray bands indicate the 95% confidence intervals of the linear regressions; error bars represent analytical uncertainties.

November 22, 2019, 12:53pm

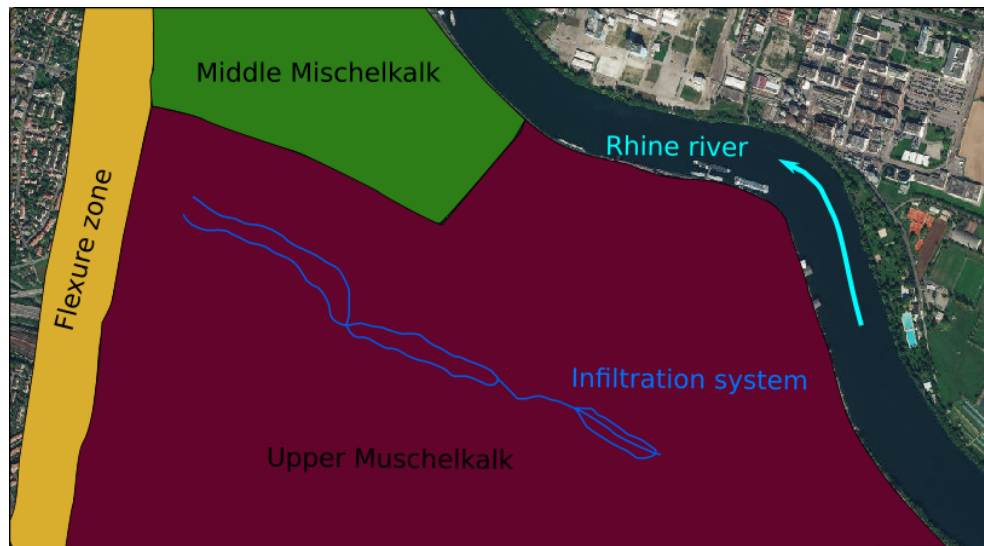


Figure S2. Simplified geological map showing the main known bedrock units at the study site. The overlying Quaternary deposits are not shown.

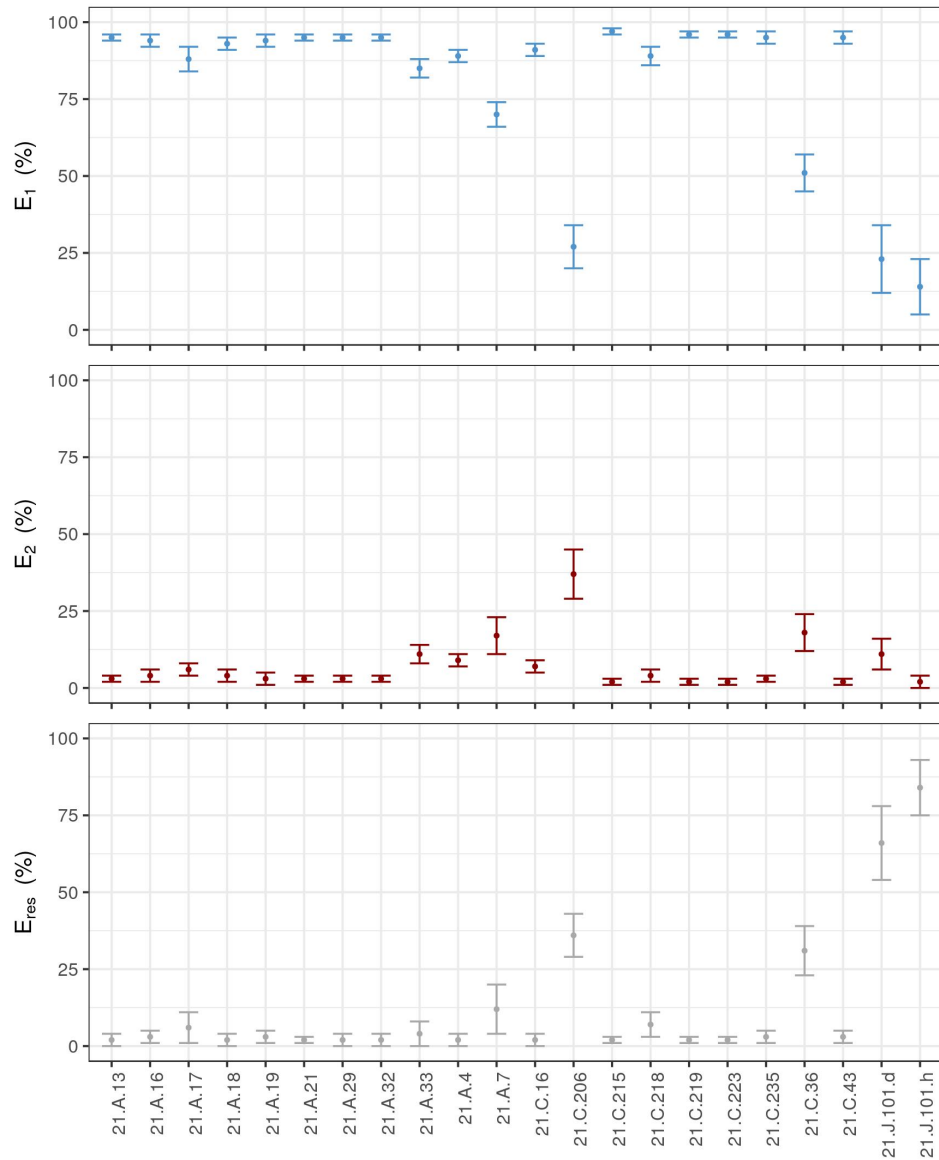


Figure S3. Estimated fractions of end-members (E_1 , E_2 and E_{res}) and their uncertainties (using TS7 and Source Code S1).

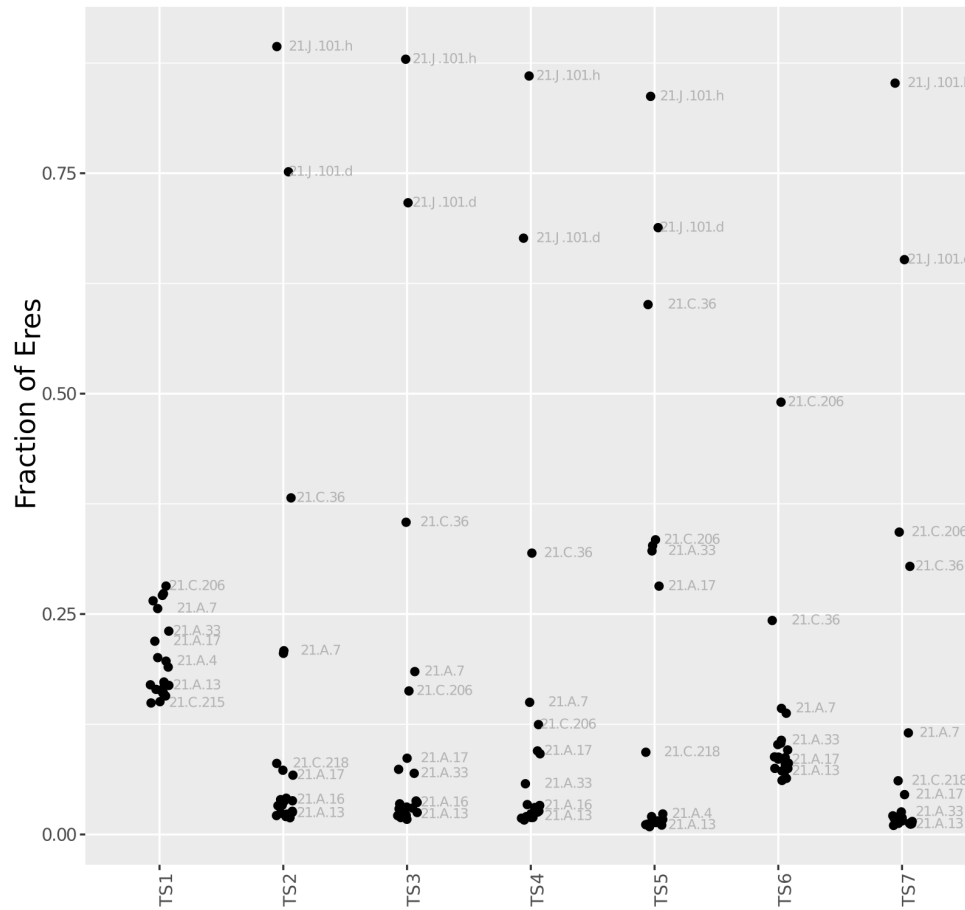


Figure S4. Sensitivity test of different tracer sets (see Table 1).

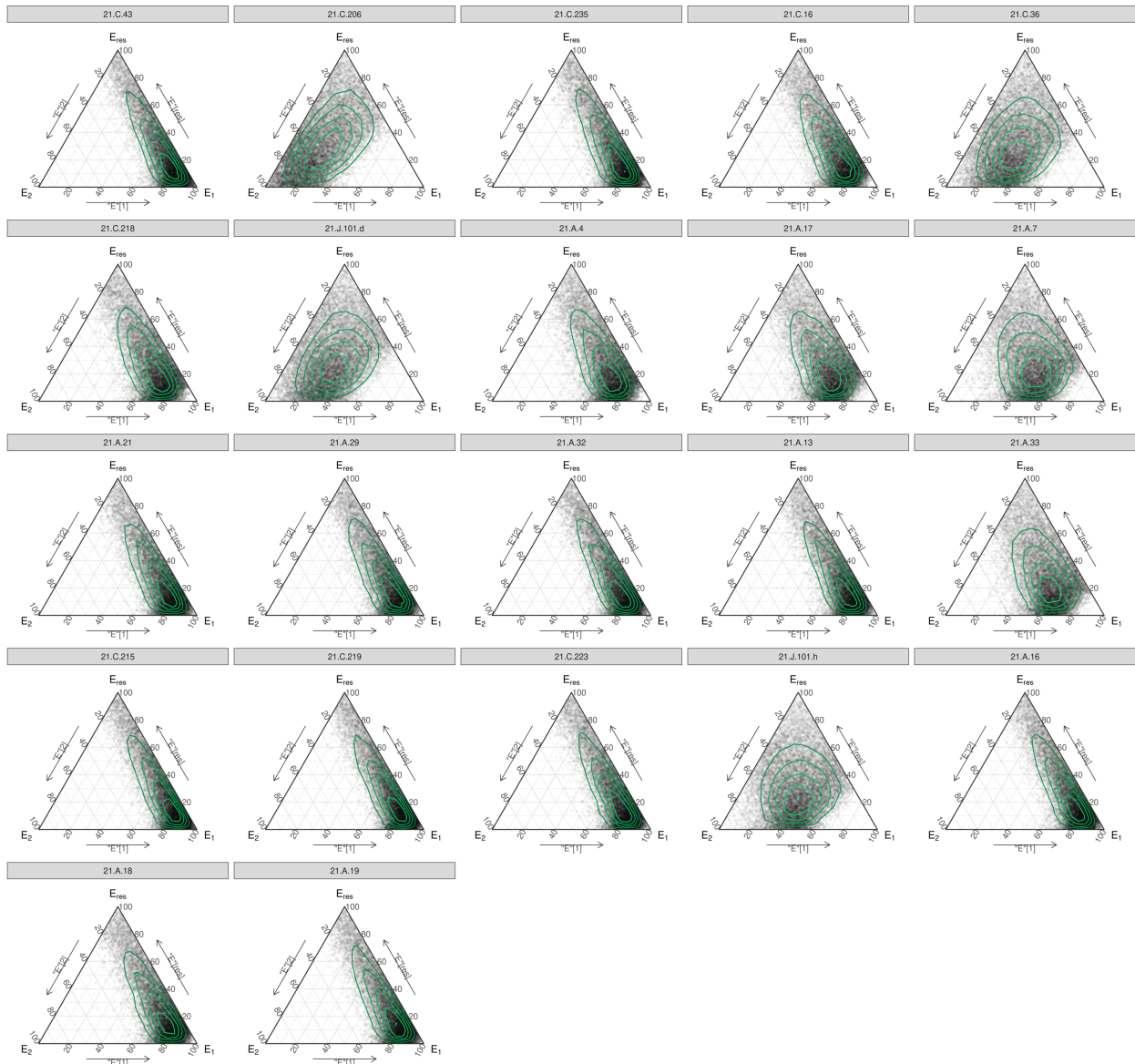


Figure S5. Ternary diagrams showing mixing ratios of all wells using tracer set 1.

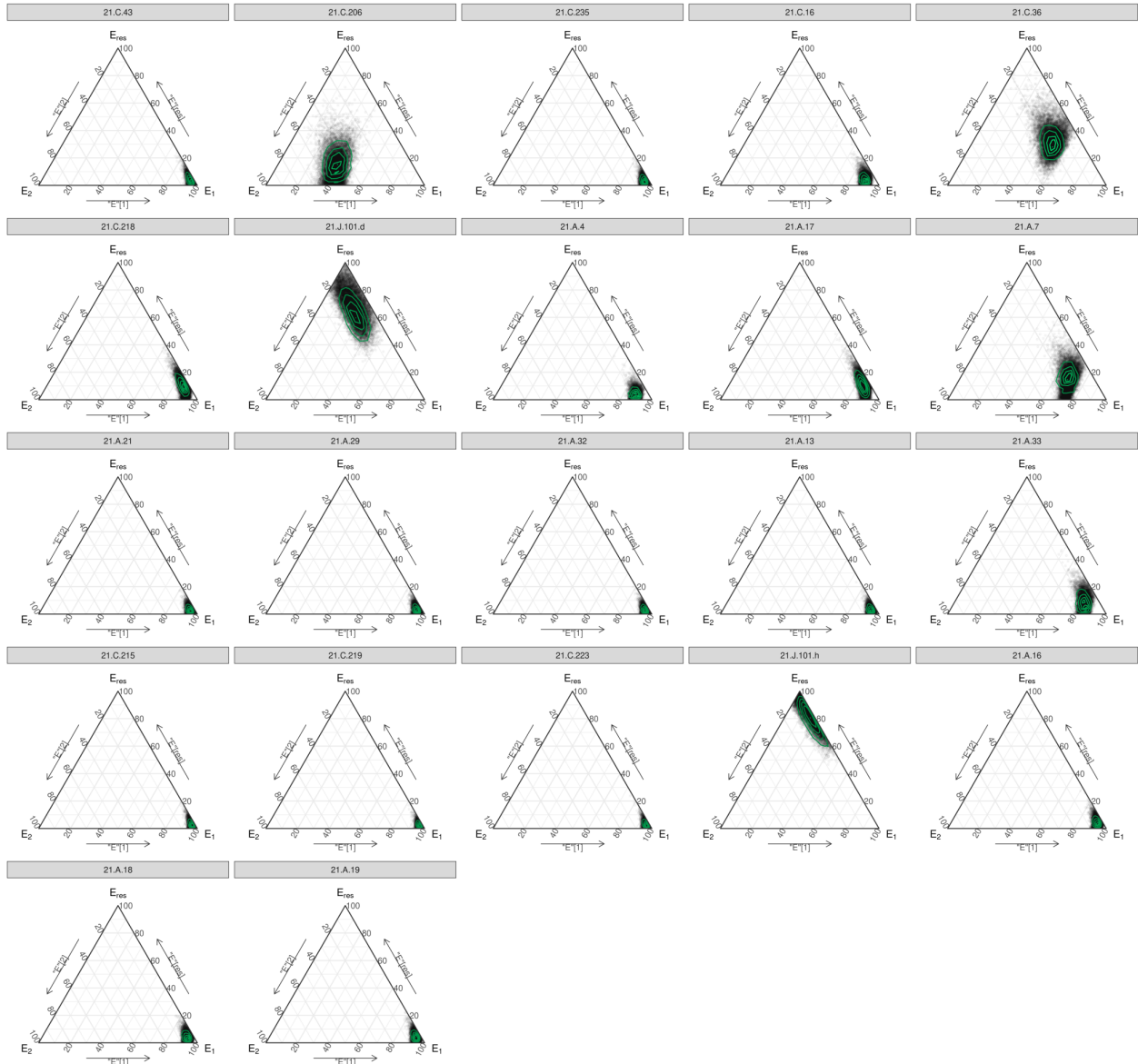


Figure S6. Ternary diagrams showing mixing ratios of all wells using tracer set 4.

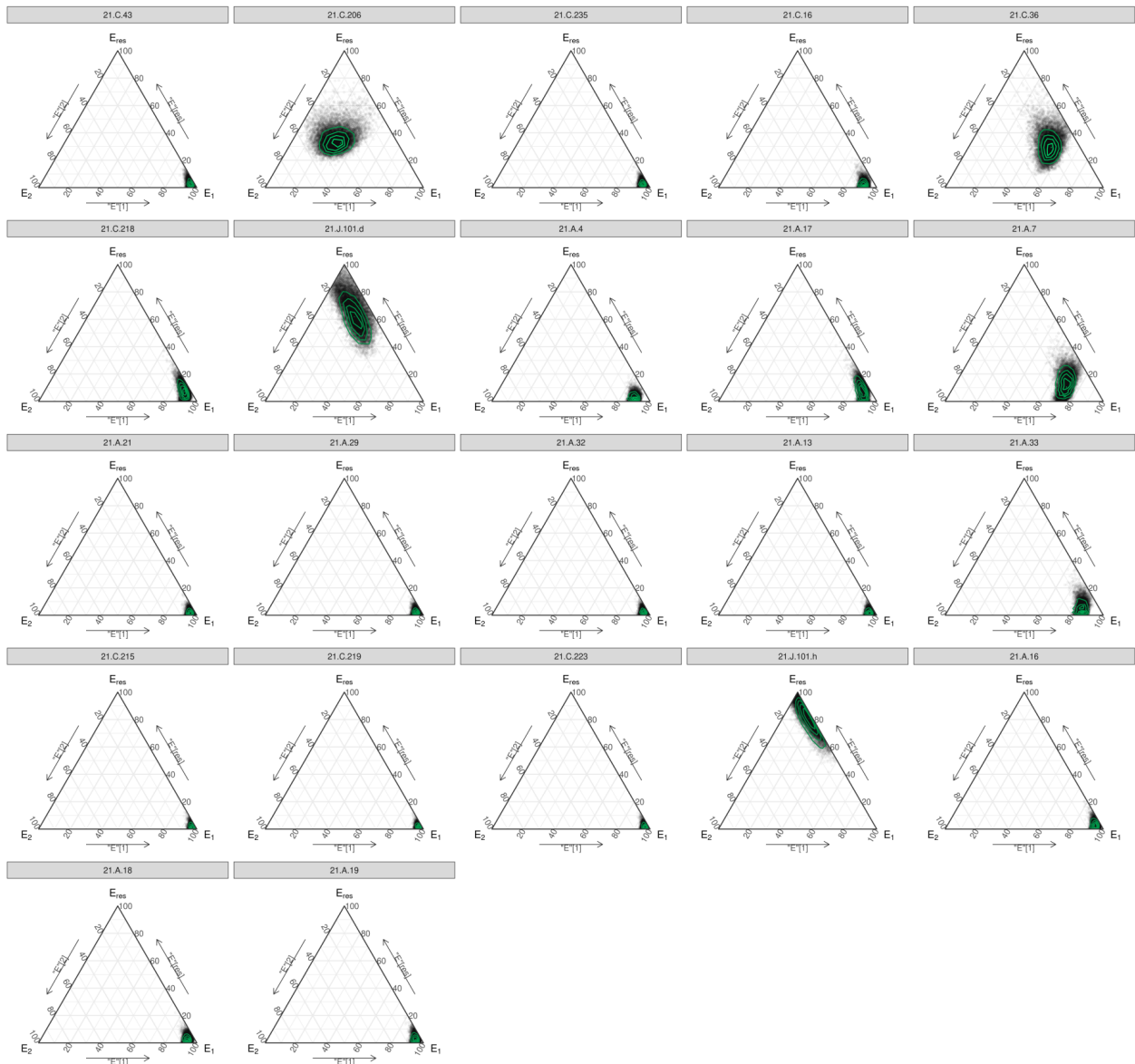


Figure S7. Ternary diagrams showing mixing ratios of all wells using tracer set 7.

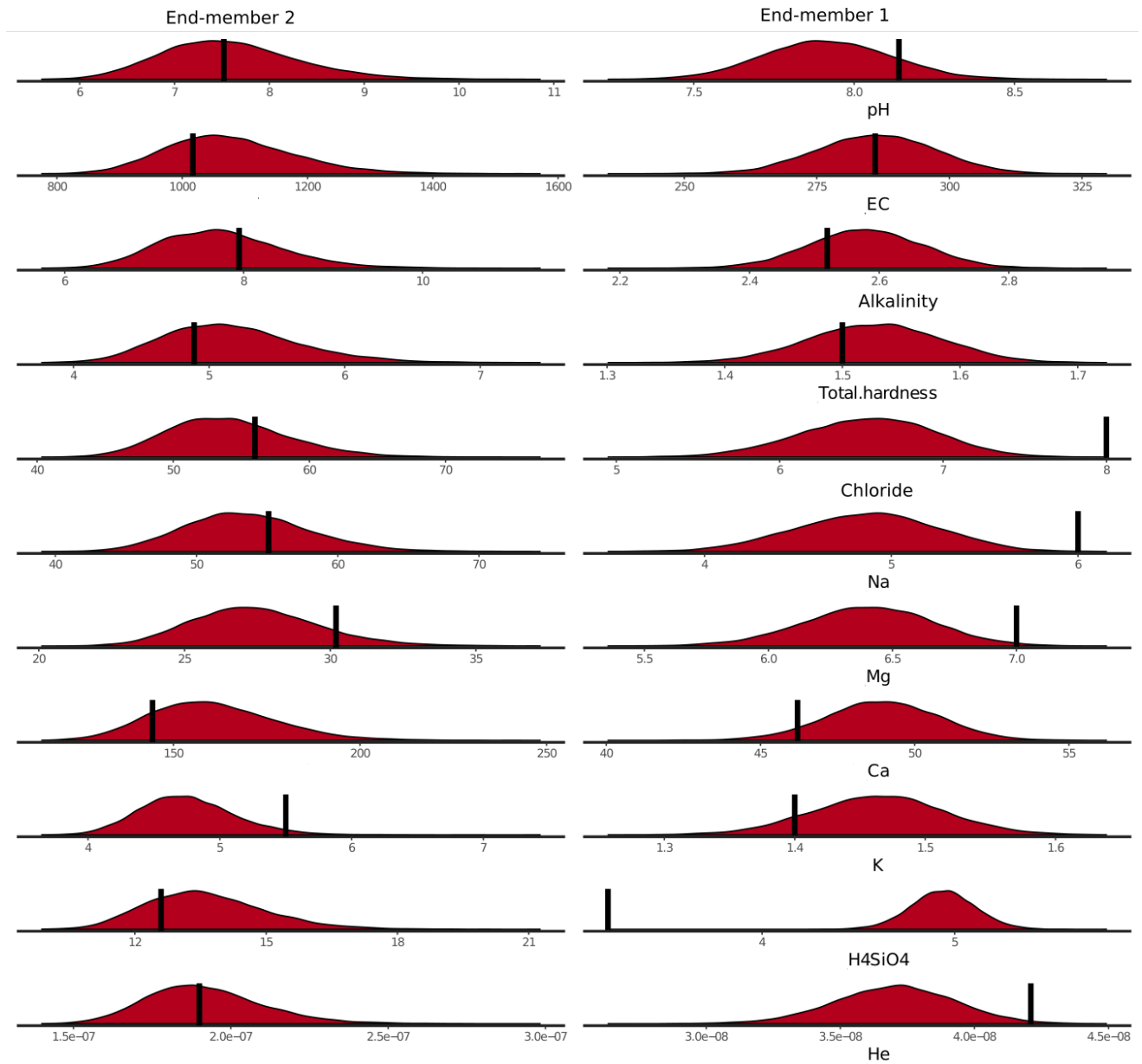


Figure S8. Comparison between model-estimated (red) and measured end-member (black) concentrations of end-members 1 and 2.

Table S1. Data for comparison between on-site and lab-based ^4He concentrations (ccSTP/g)(including analytical errors) as well as EC concentrations ($\mu\text{S}/\text{cm}$) of the respective wells studied.

Well ID	^4He (lab)	Err ^4He (lab)	^4He (on-site)	Err ^4He (on-site)	EC ($\mu\text{S}/\text{cm}$)
21.C.206	3.43E-07	1.72E-08	3.80E-07	3.80E-08	723
21.C.36	2.03E-07	1.01E-08	2.00E-07	2.00E-08	627
21.A.4	6.27E-08	3.14E-09	5.50E-08	5.50E-09	365
21.A.17	6.03E-08	3.02E-09	4.70E-08	4.70E-09	404
21.A.7	1.08E-07	5.40E-09	9.50E-08	9.50E-09	510
21.A.29	5.42E-08	2.71E-09	5.00E-08	5.00E-09	323
21.A.32	5.38E-08	2.69E-09	5.10E-08	5.10E-09	315
21.A.13	5.24E-08	2.62E-09	4.50E-08	4.50E-09	322
21.A.33	7.40E-08	3.70E-09	6.30E-08	6.30E-09	433
21.J.100_deep	2.63E-07	1.31E-08	1.90E-07	1.90E-08	1017
21.J.101_high	6.06E-08	3.03E-09	7.50E-08	7.50E-09	566
21.C.36	2.14E-07	1.07E-08	1.80E-07	1.80E-08	627
21.C.206	4.73E-07	2.37E-08	4.18E-07	4.18E-08	723
21.A.17	5.49E-08	2.74E-09	6.45E-08	6.45E-09	371.9
21.A.18	4.89E-08	2.45E-09	4.94E-08	4.94E-09	324.7
21.A.16	4.80E-08	2.40E-09	4.21E-08	4.21E-09	311.4
21.A.19	5.55E-08	2.78E-09	4.42E-08	4.42E-09	311.8

Table S2. Information about instruments used, parameter units, LOQ and measurement uncertainties for major ions and other physico-chemical parameters. LOQ and measurement uncertainties are given in the respective parameter unit.

Parameter	Instrument	Method	Units	LOQ	Uncertainty \pm
Total hardness	Metrohm 809 Titrand	Titration with Titriplex III (0.1 mol/L); Metrohm Ion Selective Electrode	mmol/L	0.2	0.1
Alkalinity	Metrohm 809 Titrand	Titration with HCl (0.1 mol/L); Metrohm pH-electrode	mmol/L	0.2	0.1
pH	Metrohm 809 Titrand	Metrohm pH-electrode	–	–	0.05
EC	Metrohm 712 Conductometer	Metrohm conductivity-measuring flow-through cell, combined with Pt 1000 temperature detection	μ S/cm	–	2
NO ₃ -N	Metrohm 761 Compact IC with chem. suppression	Column: Metrohm Metrosep A Supp 5 100/4 mm	mg/L	0.25	0.1
H ₄ SiO ₄	Skalar San++ Autoanalyzer	Spectrophotometric determination after the reaction to silicon molybdenum blue complex	mg/L	0.5	0.2
SO ₄	Metrohm 761 Compact IC with chem. suppression	Column: Metrohm Metrosep A Supp 5 100/4mm	mg/L	5	2
Na	Metrohm 761 Compact IC	Column: Metrohm C4-100/4.0	mg/L	2.5	0.8
K	Metrohm 761 Compact IC	Column: Metrohm C4-100/4.0	mg/L	1	0.3
Mg	Metrohm 761 Compact IC	Column: Metrohm C4-100/4.0	mg/L	2.5	0.8
Ca	Metrohm 761 Compact IC	Column: Metrohm C4-100/4.0	mg/L	5	1.7
Cl	Metrohm 761 Compact IC with chem. suppression	Column: Metrohm Metrosep A Supp 5 100/4mm	mg/L	0.5	0.2