Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Dipl.-Phys. Michael Werner Jung
born in Heidelberg

Oral examination: 26.11.2014

Noble Gas Thermometry in Groundwater Hydrology: Development of Advanced Evaluation Methods and Review of Data Sets from the Literature

Referees:

Prof. Dr. Werner Aeschbach-Hertig
Prof. Dr. Kurt Roth

Zusammenfassung

Das Closed-System-Equilibration-Modell (CE-Modell) zur Beschreibung der Konzentrationen von im Grundwasser gelösten Edelgasen hat sich in den vergangenen Jahren als erfolgreich erwiesen, gute Fitergebnisse für gemessene Konzentrationen zu liefern und physikalisch sinnvolle Parameterabschätzungen für eine Vielzahl von Studien zu erzielen. In manchen Fällen liefert es jedoch unrealistisch hohe Werte für die Parameter Temperatur (T) und „Excess Air“ (A) in Kombination mit hohen Temperaturfehlern. In dieser Arbeit werden der Ursprung dieses Problems sowie mögliche Lösungsansätze untersucht. Es wird aufgezeigt, dass erhöhte Argon- in Kombination mit erniedrigten Xenon-Konzentrationen das beobachtete Verhalten erzeugen können und wie mittels Monte-Carlo-Simulationen dennoch realistische Ergebnisse gewonnen werden können.

Für die praktische Anwendung dieser neuartigen Methode werden Schritt-für-Schritt-Anweisungen gegeben, die aufzeigen wann diese Monte-Carlo-Simulationen notwendig sind und wie sie ausgeführt werden sollten.

Mittels dieser neugewonnenen Erkenntnisse werden zehn publizierte Studien erneut untersucht. Die Ergebnisse der meisten Studien konnten bestätigt werden, bei einigen traten jedoch Abweichungen zu den ursprünglichen Aussagen auf.

Abschließend wird die Software *PANGA* vorgestellt, die im Rahmen dieser Arbeit entwickelt wurde. Sie erlaubt es die beschriebenen Auswertemethoden auf einfache Art und Weise auf gemessene Edelgas-Datensätze anzuwenden.

Abstract

The closed-system equilibration (CE) model for the description of dissolved noble gas concentrations in groundwater proved to be able to provide good fits to measured concentrations as well as physically reasonable parameter estimates in a variety of studies. Sometimes, however, the CE model yields unrealistically high values of the temperature (T) and “excess air” (A) parameters in combination with high temperature uncertainties. In this thesis the origin of these problems is analyzed as well as possible solutions. It is shown that increased argon concentrations in combination with decreased xenon can cause the observed problems and it is demonstrated how Monte Carlo simulations may be employed to still acquire realistic results.

Step-by-step instructions are given for the practical application of this new type of method, indicating when to use Monte Carlo simulations and showing how to carry them out.

Taking into account these new insights, a re-evaluation of ten studies from the literature is carried out. The results of most studies were confirmed. For some cases, however, deviations from the original statements were found.

The thesis concludes with a description of the software *PANGA*, which was developed to allow for an easy-to-use application of the described evaluation methods to measured noble gas data sets.

Contents

1. Introduction

The analysis of noble gases has become an important tool in groundwater hydrology. Their chemical inertness and their well-known sources and sinks make them useful tracers (Burnard 2013). Noble gases in groundwater have two major fields of application. The first field is age dating of water, which can be done applying different methods. Young groundwaters up to an age of about 50 yr may be dated using the decay of ^{3}H and the subsequent accumulation of ^{3}He (^{3}H/^{3}He method; see, e.g., Schlosser et al. 1988; Solomon and Cook 2000). This method can also be combined with other tracers like ^{85}Kr (Corcho Alvarado et al. 2007). Older groundwaters can, for instance, be dated using the radioisotopes ^{39}Ar (Corcho Alvarado et al. 2007; Ritterbusch et al. 2014) and ^{81}Kr (Sturchio et al. 2004; Buizert et al. 2014). The focus of this thesis is put on the second major field of application in groundwater, which is noble gas thermometry (Stute and Schlosser 1993; Aeschbach-Hertig and Solomon 2013).

Noble gases get into the groundwater via exchange with the atmosphere as long as the water is in contact with it. The entrapment of air bubbles in the soil matrix during infiltration, however, leads to increased noble gas concentrations when compared to air-equilibrated water. This additional component is known as "excess air" (Heaton and Vogel 1981). Its size depends on the hydrostatic pressure and therefore on the amount of water table fluctuations (Ingram et al. 2007). The amount of excess air in itself may contain interesting climate information (Aeschbach-Hertig et al. 2002b).

As the accurate description of the excess air component is important to be able to infer paleo temperatures, a number of different models have been developed (Kipfer et al. 2002; Aeschbach-Hertig and Solomon 2013). An inverse method for the estimation of model parameters was developed by two groups independently (Aeschbach-Hertig et al. 1999; Ballentine and Hall 1999). In this method, parameter values are determined by the minimization of the error-weighted deviations between observed noble gas concentrations and the concentrations derived from the models.

When the inverse method was developed, only the unfractionated air (UA), partial re-equilibration (PR) and partial de-gassing (PD) models were available. The closed-system equilibration (CE) model was introduced shortly after by Aeschbach-Hertig et al. (2000). The CE model proved to be quite successful in describing noble gas concentrations in groundwater and became widely used in different studies (e.g. Beyerle et al. 2003; Kreuzer et al. 2009; Blaser et al. 2010). Despite its success, the CE model also has its drawbacks. Aeschbach-Hertig et al. (2002b), for instance, noted that fitting the CE model to one specific sample (MD6.2) out of a set of 23 wells led to unrealistically large values of the parameter A, which describes the concentration of entrapped air, in combination with rather high estimates of the noble gas temperature and large temperature uncertainties.

In the last decade, additional models have been presented in the literature, most notably the oxygen-depletion (OD, Hall et al. 2005) and gas diffusion relaxation (GR, Sun et al. 2008) models. Furthermore, the reliability and numeric stability of the different models were critically examined (Sun et al. 2010).

In this thesis, the above-mentioned problematic CE model cases are closer examined. Chapter 3 analyzes their origin and shows how these problems can be dealt with. Based on these insights, chapter 4 gives detailed step-by-step instructions on how to evaluate groundwater noble gas data sets using the CE model, also for special cases. In chapter 5 these methods are used to re-evaluate ten data sets from the literature. In chapter 6 the software *PANGA* is introduced, which was developed as a successor to *NOBLE* by Peeters et al. (2003) and provides a set of tools for the easy implementation of the advanced evaluation methods from chapters 3 and 4.

2. Theory

2.1. Noble gases in groundwater

Noble gases in groundwater are of three major sources:

Atmospheric noble gases Water exchanges gases with the free atmosphere before infiltration into the ground where it continues to exchange gases with the soil air in the unsaturated and the quasi-saturated zones. The atmospheric noble gases can be divided into two groups: the equilibrium component, which describes the gas concentrations in water in atmospheric equilibrium, and the so-called "excess air" component, which is an additional component caused by water table fluctuations in the quasi-saturated zone.

Terrigenic noble gases These are noble gases originating from different reservoirs of the Earth like, e.g., the mantle. He, the most mobile noble gas, can ascend from deeper layers, up into the groundwater.

Radiogenic noble gases This group includes all noble gases generated by radioactive decay: ^{3}He, which is the product of the decay of tritium, ^{4}He, which is emitted in α-decays of minerals containing elements like uranium or thorium. Occasionally, radiogenic ^{40}Ar may be found in groundwaters and, very rarely, radiogenic ^{21}Ne.

The amount of atmospheric noble gases present in a sample depends, among other things, on the temperature which was prevalent at the time of infiltration of the water. Thus, the atmospheric component can be used to determine paleotemperatures. He is strongly affected by noble gases of terrigenic or radiogenic origin, making it difficult to quantify the atmospheric component. Ne, Ar, Kr and Xe, however, rarely show anything but gases of atmospheric origin, rendering them ideal candidates for the determination of noble gas temperatures.

2.1.1. Equilibrium component

Gas solubility in water is described by Henry's law

$$c_i^{\text{gas}} = H_i(T, S) \cdot c_i^{\text{water}}, \tag{2.1}$$

where c_i^{gas} and c_i^{water} are the concentrations of gas i in the gas and in the water phase. Both the amount of gas i and the amount of the solvent can be expressed in volumetric, gravimetric or molar units. Two important combinations of units are mol/l and cm^3 STP/g. Volumetric units not expressed in terms of STP, i.e., standard temperature

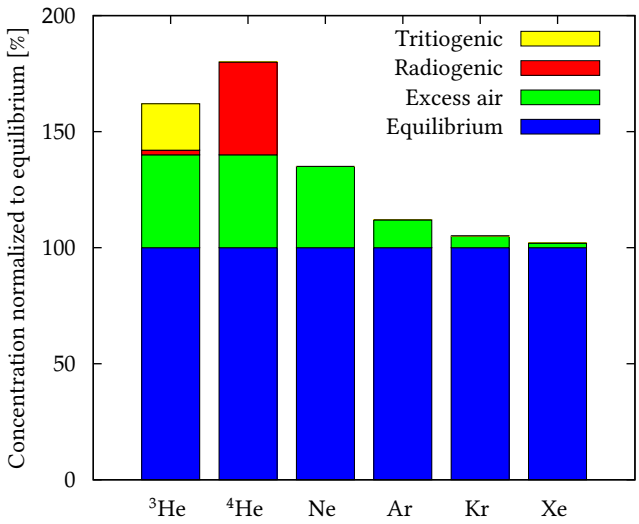


Figure 2.1.: Noble gas composition in groundwater. Adapted from Kipfer et al. (2002).

and pressure, like for instance the l in mol/l, have the disadvantage of being dependent on surrounding conditions like temperature, pressure and salinity of the water. $H_i(T, S)$ is the Henry constant, whose unit depends on the choice of units for the concentrations. If the concentrations of both the gas and the water phase are given in the same units, the Henry constant is dimensionless, but its value still depends on the choice of units. The Henry constant depends on the temperature T and the salinity S.

Henry's law may also be written in terms of the partial pressure p_i of the gas i above the water phase:

$$p_i = H_i(T, S) \cdot c_i^{\text{water}}. \tag{2.2}$$

The solubilities of noble gases were determined experimentally by a number of research groups (Weiss 1970, 1971; Weiss and Kyser 1978; Benson and Krause 1976; Clever 1979a,b, 1980; Smith and Kennedy 1983). The different noble gas solubilities were compared to gas concentrations measured in air-equilibrated water by Beyerle et al. (2000) and Aeschbach-Hertig et al. (1999). In this work, solubilities by Weiss are used for He, Ne, Ar and Kr, whereas solubilities by Clever are used for Xe, in accordance with Aeschbach-Hertig et al. (1999). Their dependence on temperature and salinity is depicted in Fig. 2.2. ^{3}He concentrations are calculated from the total He amount using the empirical equations for the ^{3}He/^{4}He ratio R_{eq} determined by Benson and Krause (1980). For a summary of the equations, refer to appendix A.

2.1.2. Excess air component

Compared to air-equilibrated water, groundwater shows an increased amount of dissolved noble gases. This gas surplus is called "excess air" and it originates from the dissolution of air bubbles which are entrapped in the groundwater or the soil matrix during infiltration. Ne concentrations exceeding the atmospheric equilibrium were first

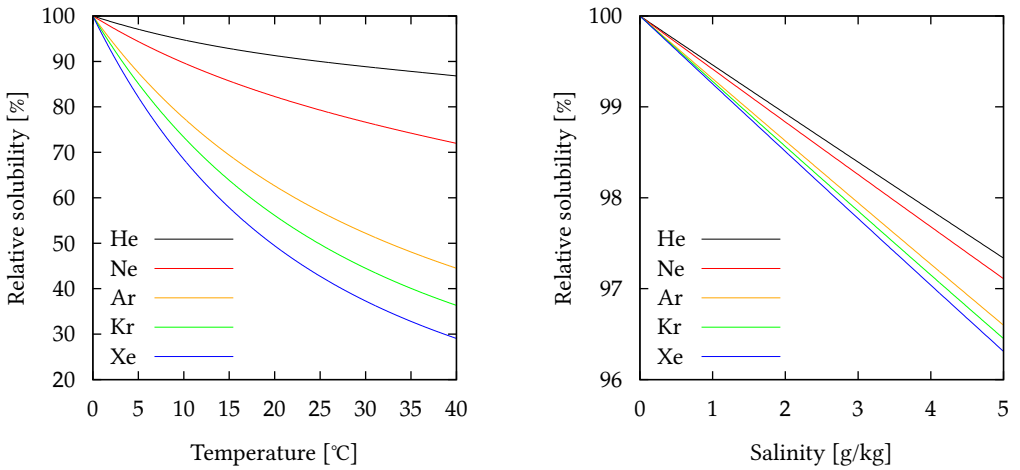


Figure 2.2.: Temperature and salinity dependencies of the equilibrium concentrations of noble gases in water. The temperature dependencies were calculated for pure water, i.e., $S = 0$. For the salinity plot, a temperature of 15 °C was assumed.

found by Andrews and Lee (1979) in a sandstone aquifer in England. The term "excess air" appeared for the first time in the study of Heaton and Vogel (1981). In this section, the most important excess air models will be introduced. A more detailed overview may be found in Aeschbach-Hertig and Solomon (2013).

The concentrations in the following equations are not expressed in molar units even though they lead to more elegant equations. Instead, cm^3 STP/g is used for concentrations in water, because this is the unit which will be used in most practical cases and that is also expected by the software *PANGA*. Unlike mol/l, it also does not depend on surrounding conditions. The conversion from mol/l to cm^3 STP/g and vice versa is given by

$$c_i[\mathrm{cm^3\ STP/g}] = c_i[\mathrm{mol/l}] \cdot \frac{V_{\mathrm{mol},i}}{\rho(T,S,P)} \tag{2.3}$$

with $V_{\mathrm{mol},i}$, the molar volume of the gas i, and the water density ρ (cf. equation A.11).

In the model equations, A denotes the amount of excess air per water mass in units of cm^3 STP/g. z_i stands for the volume fraction of the specific noble gas in dry air.

Unfractionated excess air (UA) model

The unfractionated excess air model is the simplest excess air model. It assumes complete dissolution of entrapped gas bubbles and was first used by Andrews and Lee (1979). At that time, the parameters have not yet been determined with the inverse method (which will be introduced in section 2.2), but by optimizing the agreement between the equilibration temperatures indicated by the different noble gases. The UA model remained the standard model for the description of excess air for over a decade. Its

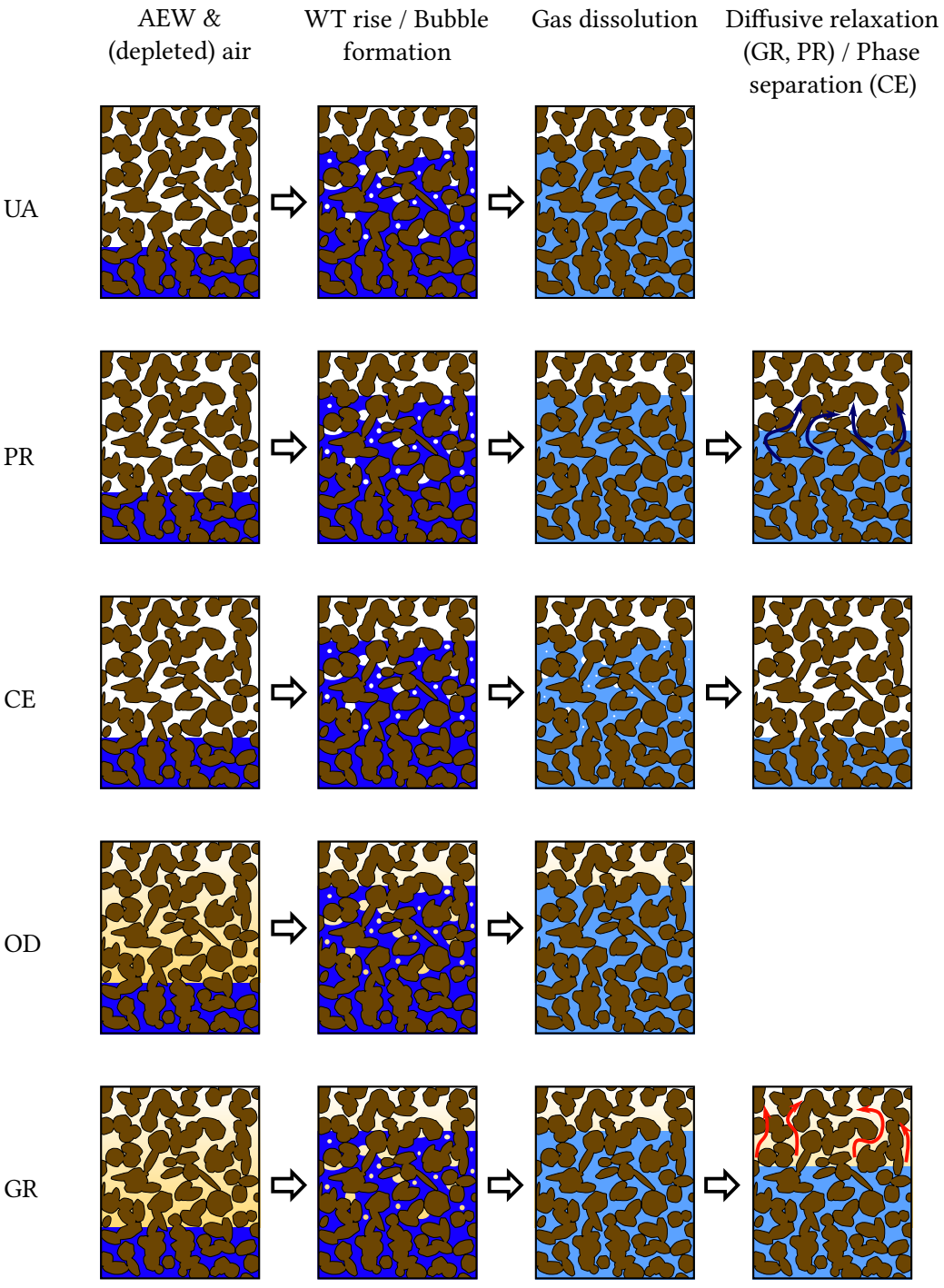


Figure 2.3.: Schematic overview of the most important excess air models. Adapted from Aeschbach-Hertig and Solomon (2013).

name is derived from the fact that its excess air component has atmospheric composition, i.e., it is unfractionated. According to the UA model, the total dissolved noble gas concentrations may be described by the equation

$$c_i = c_i^{\mathrm{eq}} + Az_i. \tag{2.4}$$

Partial re-equilibration (PR) model

After it became increasingly clear that excess air tends to contain a relatively higher amount of the heavier noble gases than atmospheric air, i.e., that the UA model often cannot provide an adequate description, Stute et al. (1995) formulated the PR model. It assumes that the water re-equilibrates again partially after the dissolution of the air bubbles. It can be expressed by

$$c_i = c_i^{\mathrm{eq}} + Az_i \cdot \exp\left[-F_{\mathrm{PR}} \cdot \left(\frac{D_i}{D_{\mathrm{Ne}}}\right)^{\beta}\right] \tag{2.5}$$

with the dimensionless excess air loss due to re-equilibration F_{PR} and the dimensionless gas transfer parameter β. The D_i are the diffusion coefficients of the noble gases in water. In *PANGA*, they are calculated following Jähne et al. (1987). Details are given in appendix A.3.

Partial degassing (PD) model

The PD model describes gas loss due to diffusion-controlled degassing into a noble-gas-free gas phase. It is mathematically similar to the PR model but the diffusive gas loss affects both the equilibrium and the excess air components (Lippmann et al. 2003):

$$c_i = \left(c_i^{\mathrm{eq}} + Az_i\right) \cdot \exp\left[-F_{\mathrm{PD}} \cdot \left(\frac{D_i}{D_{\mathrm{Ne}}}\right)^{\beta}\right] \tag{2.6}$$

Like in the PR model, the D_i denote the diffusion coefficients of the noble gases in water. The concept of gas loss relative to solubility equilibrium was already used by Zartman et al. (1961).

Closed-system equilibration (CE) model

According to the closed-system equilibration model developed by Aeschbach-Hertig et al. (2000), excess air originates from gas bubbles, which are entrapped in the soil matrix and reach a new equilibrium with the water under higher hydrostatic pressure, after the water has been shut off from the atmosphere. The noble gas concentrations may be expressed through

$$c_i = c_i^{\mathrm{eq}} + \frac{(1-F) \cdot Az_i}{1 + FAz_i/c_i^{\mathrm{eq}}}. \tag{2.7}$$

Here, the parameter A has a different meaning than in the other models: It describes the initial amount of entrapped air per unit mass of water and is measured in cm^3 STP/g.

F is the dimensionless fractionation factor, which describes the reduction of the gas volume by partial dissolution and compression. Note that the parameter A becomes a dimensionless ratio of two volumes if all concentrations are expressed in molar units, as used in the most recent representations of the CE model given by Aeschbach-Hertig et al. (2008) and Aeschbach-Hertig and Solomon (2013). The two representations can be converted into each other using the following formula:

$$A[\mathrm{cm^3\,STP/g}] = \frac{A[-]}{\rho(T,S,P)[\mathrm{g/cm^3}]} \cdot \frac{P - e_w(T)}{P_0} \cdot \frac{T_0}{T}. \tag{2.8}$$

$T_0 = 273.15\,\mathrm{K}$ and $P_0 = 1\,\mathrm{atm}$ are the standard temperature and pressure, T and P are the temperature and pressure governing atmospheric equilibrium, e_w is the saturation vapor pressure (cf. equation A.8) and ρ the density of water (cf. equation A.11).

The CE model became widely applied for the description of dissolved noble gases in groundwater due to its success in providing good fits to measured noble gas data and its ability to provide realistic estimates of physically meaningful parameters (cf. chapter 5). In addition to its normal behavior, it includes simpler models as limit cases and is also able to describe degassing (Aeschbach-Hertig et al. 2008).

Oxygen depletion (OD) model

The OD model was first introduced by Hall et al. (2005). It is based on the fact that oxygen is depleted in the soil due to biologic activity (Stute and Schlosser 1993). If the consumed gas is replaced by another gas with, for instance, a higher solubility, its partial pressure might be less than that of oxygen. In this case the partial pressures of the other gases, including the noble gases, would be increased. This effect can be described by

$$c_i = c_i^{\mathrm{eq}} \cdot P_{\mathrm{OD}} + Az_i \tag{2.9}$$

with the dimensionless overpressure factor P_{OD}. Freundt et al. (2013) confirmed the existence of the oxygen depletion effect in a study documenting the evolution of soil air composition over a period of 17 months. They found out that the sum of O_2 and CO_2 concentrations varies strongly and amounts to a fraction of 16.5 to 24.5 % of the soil air. The highest observed Ar concentrations could cause an underestimation of the NGT by 1.5 ℃. They concluded, however, that due to the annual variation, the effect is not large enough to cause overpressure factors like, e.g., $P_{OD} = 1.14$ as required by Castro et al. (2007). Hall et al. (2012) observed in a study over a period of one year that the noble gas data for early spring and mid to late summer was consistent with the typical values of $P_{OD} \approx 1.1$ found by Hall et al. (2005).

Gas diffusion relaxation (GR) model

The OD model was further developed to the GR model by Sun et al. (2008). The GR model assumes that the partial pressures of the gases are increased by different amounts because of different diffusivities. In its mathematical representation a similar term as

for the PR model is added:

$$c_i = c_i^{\mathrm{eq}} \cdot P_{\mathrm{OD}} + Az_i \cdot \exp\left[-F_{\mathrm{GR}} \cdot D_i^{\beta}\right] \tag{2.10}$$

Here, however, the D_i are the diffusitvities of gases in air, as compared to gases in water in the PR model. In *PANGA* they are calculated according to Benítez (1948).

2.2. Inverse modeling

2.2.1. Definitions

The following section contains a summary of basic definitions used throughout the rest of this chapter. The definitions are compiled from Bard (1974) and Barlow (1989). For the symbols in the upcoming equations, the following conventions are used: scalars are denoted by lowercase letters, vectors by lowercase letters in boldface and matrices by uppercase letters in boldface. Random variables are written in Greek letters whereas realizations of these variables are denoted by Latin letters.

Probability density function The probability density function $f_\xi(x)$ of a random variable ξ is defined by:

$$\text{Probability (realization } x \text{ lies between } a \text{ and } b) = \int_a^b f_\xi(x)\mathrm{d}x \tag{2.11}$$

This can be generalized to the *joint probability density function* of multiple random variables $\boldsymbol{\xi} = (\xi_1, \dots, \xi_k)^{\mathrm{T}}$:

$$\text{Probability (set of realizations } \boldsymbol{x} \text{ lies in the domain } W \subseteq \Omega) = \int_W f_{\boldsymbol{\xi}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{2.12}$$

Ω is the *sample space*, which is the set of all possible values of $\boldsymbol{\xi}$.

Expectation value The expectation value of a function $g(\boldsymbol{x})$ is given by:

$$\langle g(\boldsymbol{x})\rangle = \int_\Omega g(\boldsymbol{x}) f_{\boldsymbol{\xi}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{2.13}$$

The expectation value of one of the random variables is called μ_{x_i}:

$$\mu_{x_i} = \langle x_i\rangle = \int_\Omega x_i f_{\boldsymbol{\xi}}(\boldsymbol{x})\mathrm{d}\boldsymbol{x} \tag{2.14}$$

Variance Its variance $\sigma_{x_i}^2$, the square root of which is called *standard deviation* σ_{x_i}, is defined by:

$$\sigma_{x_i}^2 = \mathrm{var}(x_i) = \langle (x_i - \mu_{x_i})^2\rangle = \langle x_i^2\rangle - \mu_{x_i}^2 \tag{2.15}$$

Covariance The covariance between two variables is given by:

$$\mathrm{cov}(x_i, x_j) = \langle (x_i - \mu_{x_i})(x_j - \mu_{x_j})\rangle = \langle x_i x_j\rangle - \mu_{x_i}\mu_{x_j} \tag{2.16}$$

Covariance matrix The covariances can be summarized in the covariance matrix $\boldsymbol{V}$, where

$$\boldsymbol{V}_{ij} = \mathrm{cov}(x_i, x_j). \tag{2.17}$$

Note that the diagonal elements of $\boldsymbol{V}$ are again the variances.

Correlation matrix The correlation matrix with elements lying between −1 and +1 is given by

$$\boldsymbol{P}_{ij} = \frac{\mathrm{cov}(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}}. \tag{2.18}$$

Its diagonal elements are 1. The off-diagonal elements of the correlation matrix are called the *correlation coefficients*.

Data vector The *data vector* $\boldsymbol{c} = (c_1, \ldots, c_n)^{\mathrm{T}}$ is the vector of all values measured in a specific physical experiment. Here, they are the measured noble gas concentrations in groundwater. The $\boldsymbol{c}$ are realizations of the underlying random variables $\boldsymbol{\gamma}$.

Parameter vector The *parameter vector* $\boldsymbol{p} = (p_1, \ldots, p_m)^{\mathrm{T}}$ is the vector of parameters that govern a given physical system. Here, it consists of the parameters temperature, salinity, pressure and the different excess air model parameters. The true value $\hat{\boldsymbol{p}}$ of the parameter vector (if it exists) is generally unknown. The goal of the inverse method is to find estimates p_i^* for the parameters p_i.

Sampling distribution A procedure to find a single best estimate $\boldsymbol{p}^*$, a so-called *point estimate*, for the parameter vector $\boldsymbol{p}$ given the data vector $\boldsymbol{c}$ is called a *point estimation method*. An example for such a procedure is the *maximum likelihood* method, which will be explained in this chapter. The $\boldsymbol{p}^*$ depend on the $\boldsymbol{c}$, which are different realizations of the random variables $\boldsymbol{\gamma}$. The $\boldsymbol{p}^*$ therefore also follow a random distribution. This distribution is called the *sampling distribution*. It can give valuable information about the reliability of the estimates. If, for instance, a parameter is ill-determined, i.e., its estimated value is affected strongly by small variations in the data, the sampling distribution will show that the parameter has a large variance.

2.2.2. Parameter estimation

In order to obtain estimates of the parameters that govern a given physical system, a certain procedure—an estimation method—needs to be applied to the measured data. In this section, the maximum likelihood method will be used to derive an equation for least squares fitting, following Barlow (1989).

$f_{\boldsymbol{\gamma}|\boldsymbol{p}}(\boldsymbol{c})$ is the joint probability density function of the measured noble gas concentrations $\boldsymbol{c}$, which are realizations of the random variables $\boldsymbol{\gamma}$. The possible values of the concentrations depend on the parameters of the system, which is indicated by the $\boldsymbol{p}$ in

the subscript. The joint probability density function can also be considered as a function of $\boldsymbol{p}$:

$$L_{\boldsymbol{c}}(\boldsymbol{p}) \equiv f_{\boldsymbol{\gamma}|\boldsymbol{p}}(\boldsymbol{c}). \tag{2.19}$$

$L_{\boldsymbol{c}}(\boldsymbol{p})$ is called the *likelihood function*. In contrast to a "real" probability density function, which is generally unknown, because the true values of the underlying parameters are unknown, the likelihood function is itself a random variable, because it depends on the realizations $\boldsymbol{c}$ of the random variables $\boldsymbol{\gamma}$.

In the maximum likelihood method, the likelihood function is maximized with respect to the parameters $\boldsymbol{p}$, i.e., we look for the set of parameters $\boldsymbol{p}^*$, that has the highest probability of creating a set of data values like the ones measured.

For the following considerations, we assume the measured noble gas concentrations c_i to be drawn from a Gaussian probability distribution. This can be justified if we imagine that the total error of a measurement is the sum of many small, independent errors with arbitrary probability distributions. The central limit theorem states that the probability distribution of the sum of k independent random variables approximates a Gaussian distribution for large values of k (a proof can, e.g., be found in Barlow 1989). Therefore, the measured concentrations can be assumed to follow Gaussian distributions.

Under this assumption, the probability density function for the concentration of a single gas may be written as

$$f_{\gamma_i|\boldsymbol{p}}(c_i) = \frac{1}{\sqrt{2\pi\sigma_{c_i}^2}} \cdot \exp\left[-\frac{\left(c_i - c_i^{\mathrm{mod}}(\boldsymbol{p})\right)^2}{2\sigma_{c_i}^2}\right]. \tag{2.20}$$

with the standard deviation σ_{c_i} and the vector of model parameters $\boldsymbol{p}$. $c_i^{\mathrm{mod}}(\boldsymbol{p})$ is the modeled concentration of gas i and the presumed center of the distribution of γ_i. The likelihood function now becomes:

$$L_{\boldsymbol{c}}(\boldsymbol{p}) = f_{\boldsymbol{\gamma}|\boldsymbol{p}}(\boldsymbol{c}) = \prod_{i=1}^{n} f_{\gamma_i|\boldsymbol{p}}(c_i). \tag{2.21}$$

In order to find the set of model parameters $\boldsymbol{p}^*$ maximizing $L_{\boldsymbol{c}}(\boldsymbol{p})$, we use the logarithm on both sides of (2.21) and get:

$$\ln L_{\boldsymbol{c}}(\boldsymbol{p}) = -\frac{1}{2}\sum_{i=1}^{n}\left[\frac{c_i - c_i^{\mathrm{mod}}(\boldsymbol{p})}{\sigma_{c_i}}\right]^2 - \sum_{i=1}^{n}\ln\sqrt{2\pi\sigma_{c_i}^2}. \tag{2.22}$$

As the last term does not depend on $\boldsymbol{p}$, $\ln L_{\boldsymbol{c}}(\boldsymbol{p})$, and therefore also $L_{\boldsymbol{c}}(\boldsymbol{p})$, becomes maximal when

$$\chi^2 = \sum_{i=1}^{n}\left[\frac{c_i - c_i^{\mathrm{mod}}(\boldsymbol{p})}{\sigma_{c_i}}\right]^2 = \sum_{i=1}^{n}\frac{r_i^2(\boldsymbol{p})}{\sigma_{c_i}^2} \tag{2.23}$$

is minimal. This is the quantity being minimized in the inverse modeling process in order to find the unknown model parameters. The r_i, as defined on the right side, are called

residuals. Their derivatives with respect to the model parameters p_j are the elements of the *Jacobian matrix* $\boldsymbol{J}$:

$$\boldsymbol{J}_{ij}(\boldsymbol{p}) = \frac{\partial r_i(\boldsymbol{p})}{\partial p_j} \tag{2.24}$$

2.2.3. Least squares fitting

The process of minimizing χ^2 as defined in (2.23) is known as least squares fitting. The following sections give a short overview of the different methods that can be applied in order to find the minimal value of χ^2, in particular the gradient methods and the Gauss-Newton algorithm and how the advantages of both can be combined in the Levenberg-Marquardt algorithm. For a more detailed description of these algorithms refer to Marquardt (1963) or Nocedal and Wright (2006).

Gradient methods

The gradient methods are a group of methods where the solution, i.e., the set of parameters $\boldsymbol{p}$ leading to minimal χ^2, is found by stepping into the direction of the negative gradient of χ^2:

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k - \lambda \cdot \boldsymbol{\nabla}\chi^2(\boldsymbol{p}_k) \tag{2.25}$$

with a certain λ specifying the step size. The main problem with these methods is that they converge only very slowly in many cases.

The Gauss-Newton algorithm

In the Newton method χ^2 is approximated locally around the current estimate $\boldsymbol{p}_k$ by a quadratic function by expanding $\boldsymbol{\nabla}\chi^2(\boldsymbol{p})$ in a Taylor series:

$$\boldsymbol{\nabla}\chi^2(\boldsymbol{p}) \approx \boldsymbol{\nabla}\chi^2(\boldsymbol{p}_k) + \boldsymbol{H}(\chi^2(\boldsymbol{p}_k)) \cdot (\boldsymbol{p} - \boldsymbol{p}_k). \tag{2.26}$$

$\boldsymbol{H}$ is the *Hessian matrix* of $\chi^2(\boldsymbol{p})$. The position of the (supposed) minimum of this approximation, where the gradient is zero, is then taken as the next estimate:

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k - \boldsymbol{H}^{-1}(\chi^2(\boldsymbol{p}_k)) \cdot \boldsymbol{\nabla}\chi^2(\boldsymbol{p}_k). \tag{2.27}$$

With the covariance matrix of the measured noble gas concentrations

$$\boldsymbol{V}_c = \begin{pmatrix} \sigma^2_{c_1} & & 0 \\ & \ddots & \\ 0 & & \sigma^2_{c_n} \end{pmatrix},$$

the gradient and Hessian of $\chi^2(\boldsymbol{p})$ may be written as follows:

$$\nabla\chi^2(\boldsymbol{p}) = 2\sum_{i=1}^{n}\sigma_{c_i}^{-2}r_i(p)\nabla r_i(p) = 2\boldsymbol{J}^{\mathrm{T}}(\boldsymbol{p})\boldsymbol{V}_c^{-1}\boldsymbol{r}(\boldsymbol{p}) \tag{2.28}$$

$$\begin{aligned}\boldsymbol{H}(\chi^2(\boldsymbol{p}_k)) &= 2\sum_{i=1}^{n}\sigma_{c_i}^{-2}\nabla r_i(p)\nabla r_i(p)^{\mathrm{T}} + 2\sum_{i=1}^{n}\sigma_{c_i}^{-2}r_i(p)\nabla^2 r_i(p)\\ &= 2\boldsymbol{J}^{\mathrm{T}}(\boldsymbol{p})\boldsymbol{V}_c^{-1}\boldsymbol{J}(\boldsymbol{p}) + 2\sum_{i=1}^{n}\sigma_{c_i}^{-2}r_i(p)\nabla^2 r_i(p)\end{aligned} \tag{2.29}$$

For the Gauss-Newton algorithm the second term in (2.29), which is small compared to the first term in many cases, is neglected. So the algorithm is given by

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k - \left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{V}_c^{-1}\boldsymbol{J}\right)^{-1}\boldsymbol{J}^{\mathrm{T}}\boldsymbol{V}_c^{-1}\boldsymbol{r}(\boldsymbol{p}_k). \tag{2.30}$$

This algorithm converges very fast locally but it may not converge at all if the initial parameter guess is not good enough.

The Levenberg-Marquardt algorithm

The Levenberg-Marquardt algorithm combines the advantages of both of these algorithms while avoiding their problems. New estimates are calculated according to

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k - \left(\boldsymbol{J}^{\mathrm{T}}\boldsymbol{V}_c^{-1}\boldsymbol{J} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{J}^{\mathrm{T}}\boldsymbol{V}_c^{-1}\boldsymbol{r}(\boldsymbol{p}_k), \tag{2.31}$$

with a factor λ which basically blends between the two aforementioned algorithms. It is adjusted at each iteration. If the new χ^2 is smaller than the old one, λ is decreased for the next step by a factor $\nu > 0$. If it is greater, however, the new $\boldsymbol{p}$ estimate is not accepted, λ is increased by a factor ν and another estimate is calculated. This means that the gradient descent component is given more weight if (2.31) did not lead to a better estimate. Conversely, the Gauss-Newton component is strengthened if the new χ^2 is smaller than the last one, which leads to bigger steps for the next iteration.

The identity matrix in (2.31) is sometimes replaced with the diagonal of the Hessian $\boldsymbol{J}^{\mathrm{T}}\boldsymbol{V}_c^{-1}\boldsymbol{J}$ as suggested by Seber and Wild (2003). This makes the algorithm approximately invariant under rescaling of the p_i.

Some of the more modern implementations of the Levenberg-Marquardt algorithm use a full trust-region approach to update λ. The open source library *MINPACK*, which is used by *PANGA*, implements the suggestions of Moré (1978).

2.2.4. Uncertainties of the parameter estimates

An important feature of the sampling distribution is its covariance matrix $\boldsymbol{V}_p$, which gives information about the uncertainties of the parameters. Based on a single sample and without Monte Carlo simulations, only a rough approximation of the covariance matrix may be determined. For the calculation of the covariance matrix the expectation

values of the sampling distribution are needed. The estimates $\boldsymbol{p}^*$ from the least squares fit are used as an approximation of the expectation value. The covariance matrix then becomes

$$\boldsymbol{V}_p \approx \langle (\boldsymbol{p} - \boldsymbol{p}^*)(\boldsymbol{p} - \boldsymbol{p}^*)^\mathrm{T} \rangle. \tag{2.32}$$

An expression for $\boldsymbol{V}_p$ shall now be derived, roughly following Bard (1974).

The estimate $\boldsymbol{p}^*$ is the unconstrained minimum of $\chi^2(\boldsymbol{p})$. Because χ^2 also depends on the measured noble gas concentrations $\boldsymbol{c}$, it will be denoted by $\chi^2(\boldsymbol{p}, \boldsymbol{c})$ in the following. At the minimum, we have

$$\nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c}) = 0. \tag{2.33}$$

Now suppose we vary $\boldsymbol{c}$ by $\delta\boldsymbol{c}$. The minimum will then shift from $\boldsymbol{p}^*$ to $\boldsymbol{p}^* + \delta\boldsymbol{p}$:

$$\nabla_p \chi^2(\boldsymbol{p}^* + \delta\boldsymbol{p}, \boldsymbol{c} + \delta\boldsymbol{c}) = 0. \tag{2.34}$$

This can be expanded into a Taylor series, keeping only the zeroth and first order terms:

$$\underbrace{\nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c})}_{=0} + \left(\nabla_p^2 \chi^2(\boldsymbol{p}^*, \boldsymbol{c})\right) \cdot \delta\boldsymbol{p} + \left(\nabla_c \nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c})\right) \cdot \delta\boldsymbol{c} \approx 0. \tag{2.35}$$

Solved for $\delta\boldsymbol{p}$ and with $\boldsymbol{H}^* = \nabla_p^2 \chi^2(\boldsymbol{p}^*, \boldsymbol{c})$ we get

$$\delta\boldsymbol{p} = -\boldsymbol{H}^{*-1} \cdot \left(\nabla_c \nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c})\right) \cdot \delta\boldsymbol{c}. \tag{2.36}$$

With $\nabla_c \nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c}) = \nabla_p \nabla_c \chi^2(\boldsymbol{p}^*, \boldsymbol{c})$ and $\nabla_c \chi^2(\boldsymbol{p}^*, \boldsymbol{c}) = 2\boldsymbol{r}^\mathrm{T}(\boldsymbol{p}^*, \boldsymbol{c}) \boldsymbol{V}_c^{-1}$, the term in parentheses may be rewritten:

$$\nabla_c \nabla_p \chi^2(\boldsymbol{p}^*, \boldsymbol{c}) = 2\boldsymbol{J}^\mathrm{T}(\boldsymbol{p}^*, \boldsymbol{c}) \boldsymbol{V}_c^{-1}. \tag{2.37}$$

Using this result and $\boldsymbol{H}^* \approx 2\boldsymbol{J}^\mathrm{T}(\boldsymbol{p}^*, \boldsymbol{c}) \boldsymbol{V}_c^{-1} \boldsymbol{J}(\boldsymbol{p}^*, \boldsymbol{c})$ (cf. equation 2.29), equation 2.36 may be rewritten (leaving out the dependencies for brevity):

$$\delta\boldsymbol{p} = -\left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1} \cdot \boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \cdot \delta\boldsymbol{c}. \tag{2.38}$$

This can now be used to find an expression for $\boldsymbol{V}_p$:

$$\begin{aligned} \boldsymbol{V}_p &\approx \langle (\boldsymbol{p} - \boldsymbol{p}^*)(\boldsymbol{p} - \boldsymbol{p}^*)^\mathrm{T} \rangle \\ &\approx \langle \left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1} \cdot \boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \cdot \delta\boldsymbol{c} \cdot \delta\boldsymbol{c}^\mathrm{T} \cdot \boldsymbol{V}_c^{-1} \boldsymbol{J} \cdot \left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1^\mathrm{T}} \rangle. \end{aligned} \tag{2.39}$$

As the Jacobians are all evaluated at $\boldsymbol{p}^*$, they may be taken out of the expectation value:

$$\boldsymbol{V}_p \approx \left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1} \cdot \boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \cdot \langle \delta\boldsymbol{c} \cdot \delta\boldsymbol{c}^\mathrm{T} \rangle \cdot \boldsymbol{V}_c^{-1} \boldsymbol{J} \cdot \left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1} \tag{2.40}$$

Because $\langle \delta\boldsymbol{c} \cdot \delta\boldsymbol{c}^\mathrm{T} \rangle$ is the covariance matrix $\boldsymbol{V}_c$ of the measurement data (if the measured noble gas concentrations $\boldsymbol{c}$ are used as an approximation of their expectation values), (2.40) can be reduced to:

$$\boldsymbol{V}_p \approx \left(\boldsymbol{J}^\mathrm{T} \boldsymbol{V}_c^{-1} \boldsymbol{J}\right)^{-1} \tag{2.41}$$

2.2.5. Goodness of fit

The sum of the squares of n independent, standard Gaussian distributed random variables, like χ^2 as defined in (2.23), are distributed according to the χ^2 distribution

$$P_n(\chi^2) = \frac{2^{-n/2}}{\Gamma(n/2)} \chi^{n-2} e^{-\chi^2/2} \tag{2.42}$$

with the standard gamma function $\Gamma(x)$. The proof for this statement may, e.g., be found in Barlow (1989). The χ^2 probability is defined by the integral

$$\mathrm{Prob}_n(\chi^{*2}) = \int_{\chi^{*2}}^{\infty} P_n(\chi'^2) \mathrm{d}\chi'^2. \tag{2.43}$$

It represents the probability to measure a data set with $\chi^2 \geqslant \chi^{*2}$ under the assumption that the used model function precisely describes the physical processes underlying the observed system. If the χ^2 probability is high, it means, that the model function describes the data well. If it is too small, the function is not a good representation of the underlying effects. If the probability is close to 100 % and χ^2 therefore very low, it can also be a clue that something may be wrong, because normal statistical fluctuations should lead to a higher χ^2. One reason for this could be overestimated measurement errors. Conversely, high χ^2 values may also be caused by underestimated errors.

In the situation described in this thesis, the functions' m parameters are determined by minimizing χ^2. This means, however, that χ^2 will be smaller than expected. The distribution is no longer described by $P_n(\chi^2)$ but rather by $P_k(\chi^2)$ with the degrees of freedom $k = n - m$.

2.2.6. Error propagation

In this section an equation shall be derived, for the general case of error propagation of correlated variables. The derivation roughly follows Brandt (2013). Let $y_i(\boldsymbol{x})$ with $i = 1, \dots, m$ be a set of m functions depending on n random variables $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathrm{T}}$. These functions can be expanded in a Taylor series around $\boldsymbol{\mu_x} = \langle \boldsymbol{x} \rangle$:

$$y_i(\boldsymbol{x}) \approx y_i(\boldsymbol{\mu_x}) + \sum_{k=1}^{n} \left(\frac{\partial y_i(\boldsymbol{x})}{\partial x_k} \right)_{\boldsymbol{\mu}} (x_k - \mu_{x_k}) \tag{2.44}$$

In matrix notation this may be written as (with $\boldsymbol{a} = \boldsymbol{y}(\boldsymbol{\mu_x})$):

$$\boldsymbol{y}(\boldsymbol{x}) \approx \boldsymbol{J}\boldsymbol{x} + \boldsymbol{a} \tag{2.45}$$

This representation can now be used to calculate the covariance matrix (leaving out the $\boldsymbol{x}$ dependencies for brevity):

$$\begin{aligned}
\boldsymbol{V}_y &= \langle (\boldsymbol{y} - \langle \boldsymbol{y} \rangle)(\boldsymbol{y} - \langle \boldsymbol{y} \rangle)^{\mathrm{T}} \rangle \\
&\approx \langle (\boldsymbol{J}\boldsymbol{x} + \boldsymbol{a} - \boldsymbol{J}\langle \boldsymbol{x} \rangle - \boldsymbol{a})(\boldsymbol{J}\boldsymbol{x} + \boldsymbol{a} - \boldsymbol{J}\langle \boldsymbol{x} \rangle - \boldsymbol{a})^{\mathrm{T}} \rangle \\
&= \langle \boldsymbol{J}(\boldsymbol{x} - \boldsymbol{\mu_x})(\boldsymbol{x} - \boldsymbol{\mu_x})^{\mathrm{T}} \boldsymbol{J}^{\mathrm{T}} \rangle \\
&= \boldsymbol{J} \langle (\boldsymbol{x} - \boldsymbol{\mu_x})(\boldsymbol{x} - \boldsymbol{\mu_x})^{\mathrm{T}} \rangle \boldsymbol{J}^{\mathrm{T}} \\
\boldsymbol{V}_y &\approx \boldsymbol{J}\boldsymbol{V}_x\boldsymbol{J}^{\mathrm{T}}
\end{aligned} \tag{2.46}$$

This is the law of error propagation in its general form. It is needed if dependent quantities are to be derived from the parameter estimates $\boldsymbol{p}^*$ resulting from the χ^2 fit.

2.2.7. Monte Carlo simulations

In sections 2.2.2, 2.2.3 and 2.2.4, methods were given to obtain information about the sampling distribution of the model parameters $\boldsymbol{p}$. Because several assumptions and approximations were used to derive (2.41), the covariance matrices it produces are sometimes incorrect, though, in many cases, they are good approximations of the "true" covariance matrices. Likewise, the maximum likelihood estimate, resulting from the minimization of χ^2 (equation 2.23), often is a good approximation of the expectation value of the model parameter vector $\boldsymbol{p}$, but may also show significant deviations.

The results can be improved if the experiment is carried out repeatedly, yielding a set of random samples of the sample distribution. This set of samples may then be used for a statistical analysis. Creating sufficiently large sets of samples, however, is not feasible in most cases, as the accompanying expenses and efforts would not be acceptable.

A solution to this problem can be Monte Carlo (MC) simulations. For the Monte Carlo method, repeated measurements are simulated, creating a large set of samples, which can then be used to study the sample distribution. The procedure consists of the following steps (Bard 1974):

1. The individual noble gas concentrations are assigned a normal distribution with a mean value equal to the original concentration and a standard deviation equal to the presumed measurement error.

2. Sets of simulated measurement data are created by randomly drawing data from the noble gases' presumed probability distributions. The random numbers are created using a pseudorandom number generator. *PANGA* creates its pseudorandom numbers using a Mersenne Twister 19937, seeded with the system time of the computer (Matsumoto and Nishimura 1998).

3. The least squares method, as described above, is applied to every sample of the simulated data set, yielding distributions of each estimated value.

4. The resulting distributions are subjected to a statistical evaluation, giving estimates for the mean of the sampling distribution and its covariance matrix:

$$\overline{\boldsymbol{p}^*} = \frac{1}{k} \sum_{i=1}^{k} \boldsymbol{p}_i^* \tag{2.47}$$

$$\boldsymbol{V}_p^* = \frac{1}{k-1} \sum_{i=1}^{k} (\boldsymbol{p}_i^* - \overline{\boldsymbol{p}_i^*})(\boldsymbol{p}_i^* - \overline{\boldsymbol{p}_i^*})^{\mathrm{T}}, \tag{2.48}$$

with the number of Monte Carlo realizations k and the estimate from the i-th experiment $\boldsymbol{p}_i^*$.

3. Properties of the closed-system equilibration model

The results in the following chapter have already been published in Jung et al. (2013). Large parts of the upcoming text are almost identical to parts of the article.

3.1. Samples

In this chapter, four pairs of groundwater samples are used to illustrate typical fitting behaviors of the CE model and to test new approaches to improve the data evaluation. The samples were selected to show pairs of a "normal" and a "bad" sample for different study areas with various environmental conditions. In particular, the selected study areas cover mean annual air temperatures (and thus expected NGTs) in the range of about 10 to 28 ℃. Two pairs were taken from published noble gas temperature records, the other two from recent studies performed in the noble gas laboratory at Heidelberg University. The pairs were selected to represent as similar recharge conditions as possible, thus very similar results for the derived parameters of the two samples are expected, at least for the equilibration temperature. In two cases the pairs are duplicate samples from one well, thus strictly identical parameters can be expected.

The coolest recharge conditions investigated here are represented by two samples from rather shallow wells from Ireland. The samples MAS_21 and MAS_24 originate from an agricultural site near the city of Wexford in southeastern Ireland. The mean annual air temperature in the area is about 10 ℃. The recharge altitude is estimated to be 61 m above sea level (asl) for MAS_21 and 40 m for MAS_24 according to the land elevation at the wells.

A pair of samples with slightly warmer expected temperatures was taken from the study of Aeschbach-Hertig et al. (2002b). The samples MD6.2 and MD9.1 were collected in 1995 from wells in the outcrop and recharge area of the Aquia aquifer in southern Maryland, USA. These samples represent modern groundwater based on ^{14}C, ^{3}H and ^{4}He data. The recharge altitude in this coastal region is about 50 m asl and the mean annual air temperature is 12.8 ℃.

A pair of duplicate samples representing rather warm recharge conditions was taken from the study of Castro et al. (2007). The samples TX46.2 and TX46.3 are two out of three replicates from a well just downstream of the recharge area of the Carrizo Aquifer in Texas. According to the evaluation of Castro et al. (2007), the He model age of groundwater at this well is 960 yr and the NGT estimates derived by the UA model for the two samples are 17.9 ℃ and 17.4 ℃, respectively. These NGTs are lower than the mean annual air temperature of 20.6 ℃ in the area and even much lower than current

ground temperatures which were estimated to be 24.5 ℃. On average 4.7 ℃ higher NGTs were obtained by applying the OD model with an overpressure factor of 1.14 to the record from Texas, reconciling modern NGTs and ground temperatures. The recharge altitude for the Carrizo Aquifer is 200 m asl.

Another pair of duplicate samples representing even warmer recharge conditions was taken from a recent study in India. The samples IND32_1 and IND32_3 were both taken from the same well in a sedimentary aquifer system in the North Cambay Basin, Northwest India (Wieser 2011). Based on ^{14}C data, the water was estimated to be (1.7 ± 0.5) kyr old. The mean annual air temperature in this region (city of Ahmedabad) is 27.5 ℃, the recharge altitude is about 130 m asl.

Table 3.1.: Concentrations of dissolved noble gases for the samples used to analyze the CE model behavior

Sample	Ne [10^{-7} cm^3 STP/g]	Ar [10^{-4} cm^3 STP/g]	Kr [10^{-8} cm^3 STP/g]	Xe [10^{-9} cm^3 STP/g]
MAS_24	2.726 ± 0.014	4.270 ± 0.015	9.663 ± 0.118	12.812 ± 0.312
MAS_21	2.172 ± 0.008	3.963 ± 0.015	9.262 ± 0.110	12.057 ± 0.291
MD9.1	2.493 ± 0.020	4.020 ± 0.028	8.986 ± 0.081	12.216 ± 0.098
MD6.2	2.145 ± 0.017	3.836 ± 0.027	8.757 ± 0.079	12.156 ± 0.097
TX46.3	2.167 ± 0.028	3.418 ± 0.044	7.534 ± 0.113	10.011 ± 0.220
TX46.2	2.031 ± 0.026	3.341 ± 0.043	7.323 ± 0.110	9.789 ± 0.215
IND32_3	1.884 ± 0.011	2.759 ± 0.022	6.054 ± 0.119	7.620 ± 0.113
IND32_1	1.869 ± 0.011	2.787 ± 0.022	5.900 ± 0.122	7.412 ± 0.107

Table 3.2.: Prescribed pressures, fitting results and Monte Carlo statistics

		Fitting results				Monte Carlo statistics	
Sample	P [atm]	χ^2	T [℃]	A [cm^3 STP/g]	F	T [℃]	A [cm^3 STP/g]
MAS_24	0.9952	3.94	10.89 ± 0.72	0.012 ± 0.004	0.490 ± 0.078	10.93 ± 0.76	0.013 ± 0.005
MAS_21	0.9927	7.98	12.71 ± 3.35	0.214 ± 0.900	0.893 ± 0.009	12.67 ± 0.58	0.245 ± 0.133
MD9.1	0.9940	1.28	13.23 ± 0.42	0.020 ± 0.006	0.663 ± 0.038	13.27 ± 0.47	0.021 ± 0.011
MD6.2	0.9940	0.025	14.43 ± 2.53	0.292 ± 0.891	0.894 ± 0.015	13.58 ± 0.81	0.159 ± 0.164
TX46.3	0.9762	0.41	18.40 ± 1.17	0.017 ± 0.018	0.755 ± 0.088	19.03 ± 1.89	0.066 ± 0.134
TX46.2	0.9762	0.0015	20.56 ± 12.93	0.127 ± 1.413	0.876 ± 0.016	19.83 ± 1.31	0.108 ± 0.121
IND32_3	0.9904	2.56	27.43 ± 0.74	0.0008 ± 0.0088	0.001 ± 9.957	27.91 ± 0.98	0.028 ± 0.093
IND32_1	0.9904	0.87	30.27 ± 16.32	0.115 ± 2.010	0.903 ± 0.023	30.04 ± 0.84	0.111 ± 0.070

The samples' noble gas concentrations are summarized in Table 3.1. The prescribed pressures based on the estimated recharge altitudes as used for the fit are listed in Table 3.2 together with the fit results. The samples MAS_24, MD9.1, TX46.3 and IND32_3 are of the usual well-behaved type: the CE model provides good fits to these data sets with reasonable parameter values and low uncertainties. In contrast, the samples MAS_21, MD6.2, TX46.2 and IND32_1 show the typical problems which sometimes can occur when applying the CE model: the temperatures derived from these samples are all higher than for the corresponding normal samples and temperature uncertainties are

large. The values and uncertainties of the A parameter are also much higher than for the normal samples.

It is interesting to check the high estimates of A typical for problematic samples for their physical plausibility. On the basis of field data of Fayer and Hillel (1986) on the volume fraction of entrapped air, A may be expected to range from 0.02 to 0.18 cm^3 STP/g. However, the higher values of A most likely include entrapped air that is mobile and may escape over time. Presumably only the immobile entrapped air is relevant for the formation of excess air, and according to the perhaps most reliable data from laboratory experiments, the fraction of the pore space that contains persistently trapped immobile air is likely smaller than 5 % (Faybishenko 1995). Thus a realistic expectation is that the parameter A should be smaller than 0.05 cm^3 STP/g. Such values of A have indeed been observed in a variety of study areas (Aeschbach-Hertig et al. 2002a, 2000).

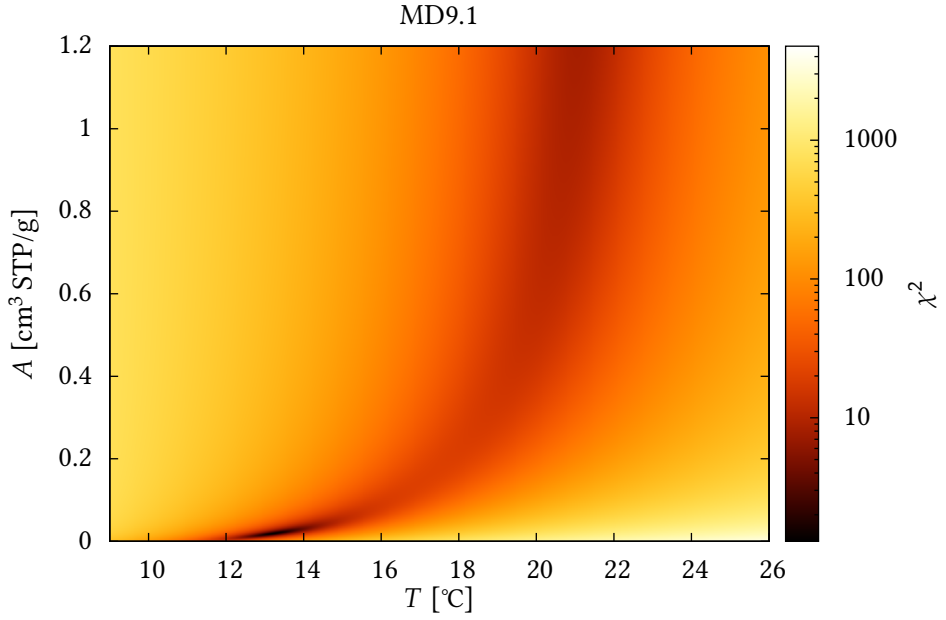
The problematic samples all have values of A higher than 0.1 cm^3 STP/g, clearly above the realistically expected range. At least the highest A values of 0.2 to 0.3 cm^3 STP/g are definitely unrealistic, as they imply that about 20 % or more of the pore space was initially occupied by entrapped air. It is hardly possible that such large amounts of air could be trapped. The experience of our group with applying the CE model to many data sets suggests that A values above 0.1 cm^3 STP/g lead to large uncertainties of both A and T and are characteristic of the type of problematic fits that are investigated here.

3.2. Analysis of the χ^2 space

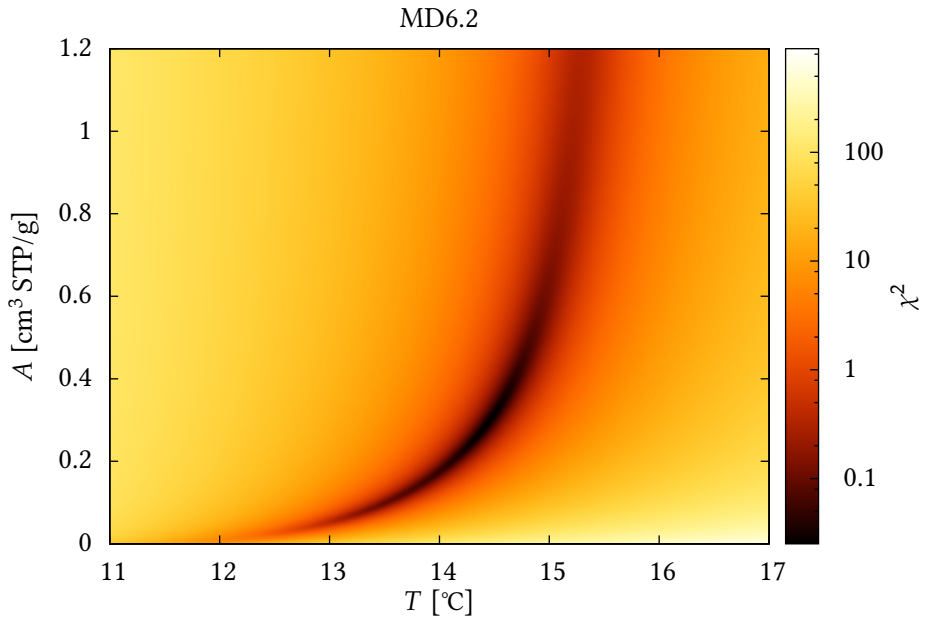
One possible problem of parameter estimation for non-linear models is that multiple minima of the χ^2 surface in the parameter space may occur, making the solution found by the fitting routine non-unique. Such non-uniqueness might lead to numerical instability and a dependence of the solution from the initial values of the parameters used at the start of the search for the minimum. These aspects were studied for the CE model by von Oehsen (2008), who found that for synthetic samples a second local minimum exists in the χ^2 surface in the three-dimensional parameter space (T, A, F), which is characterized by too high values of A and T. The fit results then depend on the choice of initial values for the fit parameters, as too high initial values for A and T lead to the fitter finding the "wrong" minimum.

In order to further investigate the phenomenon of multiple local minima in the χ^2 surface and its possible connection to the poor fitting behavior observed for some samples, 2D plots of χ^2 were produced, i.e., projections of the χ^2 surface onto the subspace spanned by two fit parameters. Fig. 3.1a shows χ^2 for the sample MD9.1 as a function of A and T, as an example of such a plot for a normal sample. One can see two local minima, one at realistic parameter values ($A = 0.033\,cm^3$ STP/g, $T = 13.35\,°C$), which is the solution found by the fit, and another one at unrealistically high values for A and T. In fact, if the fitting routine is forced into finding this second minimum by appropriately high initial parameter values, for many samples such as MD9.1, it yields extremely high values of A, many orders of magnitude above the physical range.

This pattern with two local minima is observed for all well-fitting samples. Even



(a) χ^2 of a well-behaved sample



(b) χ^2 of a problematic sample

Figure 3.1.: χ^2 as a function of A and T for a well-behaved (a) and a problematic (b) sample. The parameter F was fitted for minimal χ^2. The MD9.1 sample shows the characteristic pattern of two local minima, one at realistic parameter values, the other one at unrealistically high values. For MD6.2 one can only see a single minimum at high but not impossible values.

though the second minimum at high values of A and T can in some cases be the global minimum (i.e., have a lower χ^2 value than the first minimum), it is not found by the fit routine if low initial values for A and T are used, except for a few rare cases which are easily identifiable. Hence, the existence of two minima does not seem to lead to numerical instability or to otherwise be a problem for the interpretation of most samples.

In fact, the problematic samples, where the fits converge to solutions with large A, all show a single minimum only and their χ^2 surface looks like that of sample MD6.2 shown in Fig. 3.1b. Since there is only one minimum, the fit result cannot depend on the initial parameter values in these cases, thus the problematic behavior can neither be explained by bad parameter initials nor by an instability due to multiple minima and thus a non-unique solution. It rather seems that problems arise when the two well-separated minima of normal samples, one of them clearly being physically unrealistic, merge into one broad minimum, which now occurs at values of the parameter A that are high but not completely impossible. Such solutions cannot simply be ruled out and may indeed be correct within the large parameter uncertainties that typically occur, but they are practically useless due to the large uncertainties. They are probably also biased towards too high values of A and T.

Note that for both well-behaved and problematic samples, the region of low χ^2 values in the plot of A versus T (Fig. 3.1) has the shape of an elongated, curved valley, which turns upwards (toward high values of A) at higher temperatures. The finding that this valley becomes almost vertically aligned for large A can be explained by the fact that the CE model equation becomes independent of A for high values of A (von Oehsen 2008):

$$\begin{aligned} \lim_{A\to\infty} \frac{(1-F)Az_i}{1+FAz_i/c_i^{\mathrm{eq}}} &= \frac{1-F}{F}c_i^{\mathrm{eq}} = \frac{1}{F}c_i^{\mathrm{eq}} - c_i^{\mathrm{eq}} \\ \Rightarrow c_i^{\mathrm{mod}} &= \frac{1}{F}c_i^{\mathrm{eq}} \end{aligned} \tag{3.1}$$

Thus, in the limit of large A, the CE model describes a uniform increase of all noble gas concentrations by the factor $1/F$, which is the same as the effect of a pressure increase (Aeschbach-Hertig et al. 2002a, 2000). In this limit, the CE model is equivalent to a pure OD model (only excess pressure, no excess air), with $1/F$ instead of the OD pressure factor. However, the CE model contains an additional but in this case useless parameter, as the model equations become completely independent of A. It is therefore clear that the parameter A cannot be determined by the model inversion in the limit of large A.

The second minima of normal samples occur at very large values of A, falling into the limit where A is indeterminable. In contrast, the problematic samples have single minima at A values that are above the realistic range ($A > 0.1\,\mathrm{cm^3\,STP/g}$), but still small enough to have a certain influence on the modeled concentrations. In this range, the parameters T and A exhibit a high degree of correlation (the correlation coefficient can reach values of about 0.999 or even higher), which explains the large parameter uncertainties. In particular, the parameter A is approaching the limit of its identifiability. Thus, similar to the case of a model with free pressure parameter, the occurrence of high parameter correlation is part of the explanation of the problems with the CE model.

However, in contrast to a model with unknown pressure, these high correlations only occur for certain samples, which yield solutions in the critical range of A values.

3.3. Monte Carlo analysis

The analysis of the χ^2 surface in the parameter space shows characteristic differences between normal and problematic samples, but it does not lead to a solution of the problem in terms of local minima that should be avoided. As a next step to study the behavior of the fits, and in particular the reliability of the parameter error estimates, Monte Carlo analyses were performed for the test samples. Fig. 3.2 shows one- and two-dimensional histograms of the parameters A and T, where a 1D histogram is equivalent to the projection of the 2D histogram onto the respective axis. For the preparation of these plots, as well as the other Monte Carlo plots in this chapter, one million random data sets were created and evaluated as described in section 2.2.7. For the creation of the 1D histograms, the domain of the parameter values was divided into 500 bins. For the 2D histograms the two-dimensional parameter value domain was divided into $500 \cdot 500$ bins. Most of the normal samples shown in the left column of Fig. 3.2 more or less exhibit the expected behavior, where the solutions for the Monte Carlo realizations of a sample are approximately normally distributed around the solution for the original sample. For the two cooler samples MAS_24 and MD9.1, the standard deviations of the 1D histograms for A and T correspond closely to the parameter errors from the covariance matrix estimated in the fit procedure (cf. Table 3.2).

The two warmer samples, TX46.3 and IND32_3 already show some deviations, with a significant fraction of the Monte Carlo realizations leading to solutions with unrealistically large A and increased T. These "bad" realizations do not occur for MAS_24 and are very rare for MD9.1. For TX46.3 and IND32_3, however, they are more frequent and start to form a second cluster at high A and T in the 2D histograms. In the 1D histograms, long tails or even weak secondary peaks at high parameter values become visible (under magnification). As a result, the standard deviations of these 1D distributions from the Monte Carlo analysis are larger than the parameter errors estimated from the fit.

For comparison the same Monte Carlo analysis was applied to a large number of well-fitting samples not shown here. Many histograms showed either a single cluster only or, like MD9.1, just a few results scattered in the area of high A ($A > 0.1\ \mathrm{cm^3\ STP/g}$). But several samples also had a second cluster in this area. Most of the time it was, however, not as pronounced as the clusters of the poorly-fitting samples.

The 1D histograms for the parameter A also show a different anomaly for part of the normal samples, namely a distinct peak at low values of A, superimposed on the general distribution. This anomaly is also visible in the 2D histogram for sample MAS_24. While all realizations of this sample converge at low and realistic values of A, a small but significant fraction yields a fixed, very low value of A, leading to a peak at the lower end of the A-distribution in the 1D histogram. A similar, even more pronounced peak is visible for sample TX46.3. It corresponds to a significant fraction of the Monte

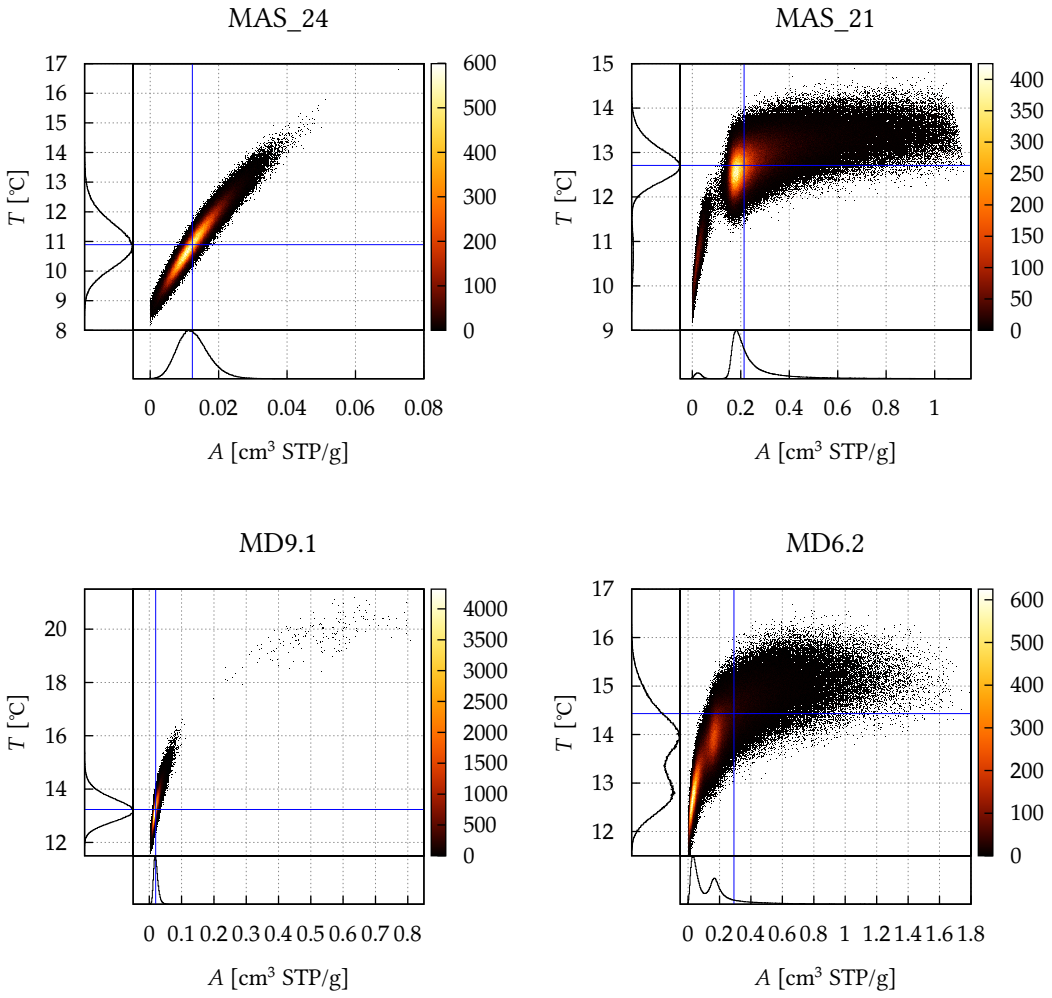


Figure 3.2.: Histograms of the Monte Carlo results of the analyzed samples (continued on the following page).

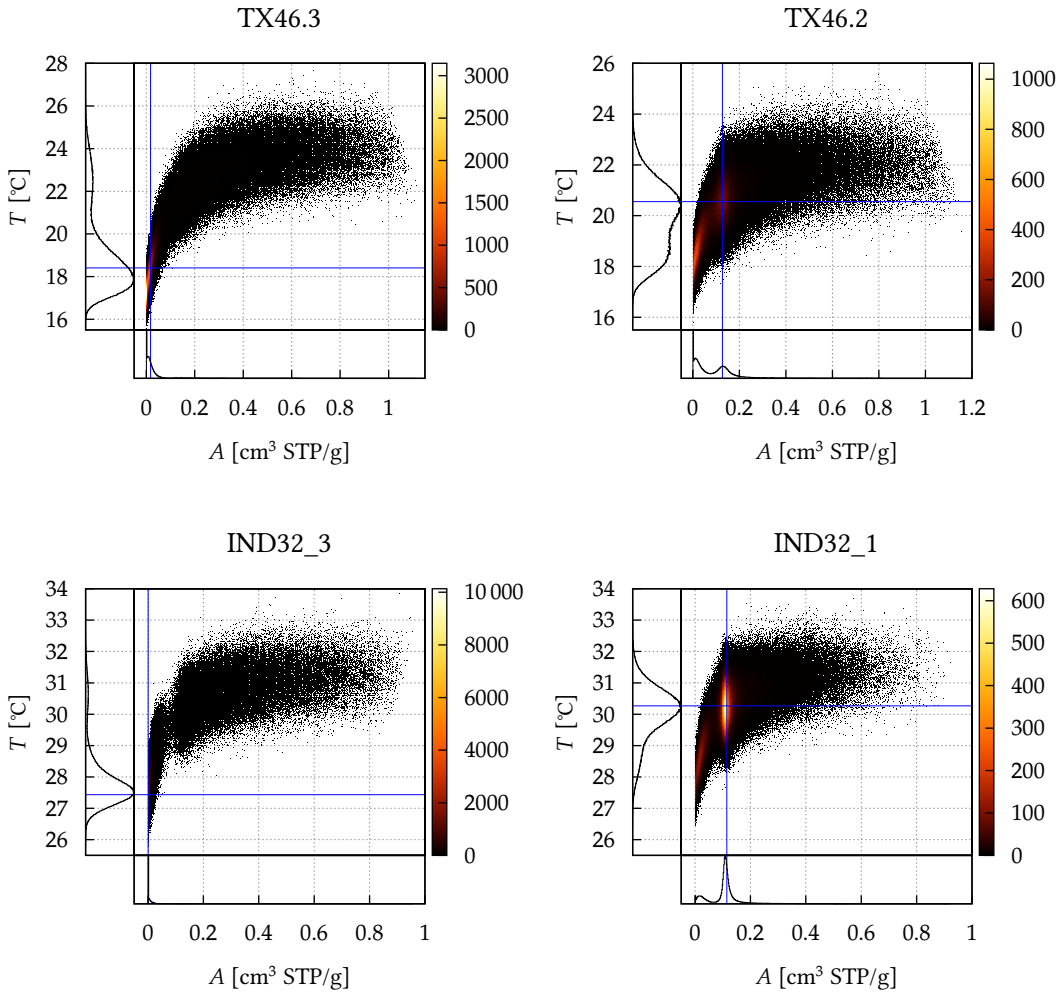


Figure 3.2 (cont.): Histograms of the Monte Carlo results of the analyzed samples, showing the number of realizations yielding specific values of the parameters A and T. The blue lines represent the original samples' fit results. The left column shows the well-behaved samples. For these, most Monte Carlo realizations concentrate around the fit results. The problematic samples, however, which are shown in the right column, exhibit a split-up into two clusters.

Carlo realizations yielding fits with low A ($< 0.01\,\mathrm{cm^3\,STP/g}$, sometimes even below $0.001\,\mathrm{cm^3\,STP/g}$) and F equal or very close to zero. These solutions with $F = 0$ correspond to unfractionated atmospheric composition of the excess air, or the UA model in the terminology of Kipfer et al. (2002). They reflect the fact that the CE model contains the simpler UA model as a limiting case (with $F = 0$) and obviously this limit provides the best solution for a significant part of the Monte Carlo realizations of such samples.

Solutions of the CE model with $F = 0$ occur quite frequently, mainly for well-fitting samples. An extreme example is IND32_3, where the "good" cluster at low A, including the fit for the sample itself, consists almost exclusively of solutions with $F = 0$ and only the realizations in the "bad" cluster at high A represent true (but problematic) CE model solutions. The separation of Monte Carlo realizations as well as single fits to real samples into a group corresponding to the UA model (with $F = 0$) and a group of true CE cases (with $F > 0$) has usually not been perceived as a problem, as the UA type solutions are well constrained, but it certainly deserves further analysis. It occurs, however, separately from the phenomenon of solutions with large A and does not seem to be related to that problem, which is the main focus of this investigation.

The phenomenon of solutions with large A is illustrated by the problematic samples in the right column of Fig. 3.2. All of them show a similar behavior, with their Monte Carlo realizations forming two more or less distinct clusters, one of them concentrating in the area of the fit result (at high A), the other one at parameter values which are much more realistic. As discussed above, such a pair of clusters also occurs for some normal samples (especially IND32_3), but for the problematic samples the clusters with large A are much more pronounced and also contain the original samples' fit results. The 1D histograms for the problematic samples typically show a two-peak structure both in A and T, corresponding to the two clusters. For example, for IND32_1, the "bad" cluster dominates in a similar way as the "good" cluster does for IND32_3. A result of the clearly non-normal distributions of the parameters in such cases is that the standard deviations in these histograms, i.e., the parameter error estimates derived from the Monte Carlo procedure, do not agree with the parameter errors from the covariance matrix estimate obtained from the fitting algorithm (cf. Table 3.2). Although the distributions derived from the Monte Carlo analysis are rather wide for the problematic samples, their standard deviations are significantly smaller (up to a factor of almost 20 for IND32_1) than the original parameter error estimates. For most samples the mean values are also slightly shifted compared to the fit results. This result shows that the usually very large parameter errors for problematic fits (see Table 3.2) are not reliable. The Monte Carlo analysis provides better error estimates, but more importantly it reveals a systematic behavior of the solutions that deserves further investigation.

In order to find an explanation for the clustering of Monte Carlo realizations in the space of the A and T parameters, an analysis was carried out to determine which noble gas concentrations led to parameter values in which cluster. The mean concentrations of both clusters were calculated for eight different samples (both well- and poorly-fitting) and compared to each other. The comparison showed that the clusters with the unrealistic values all have increased Ar (+0.2 to +1.8 %) and decreased Xe (−1.4 to −11.0 %) concentrations. The Ne concentrations were only slightly decreased (−0.1 to −0.3 %),

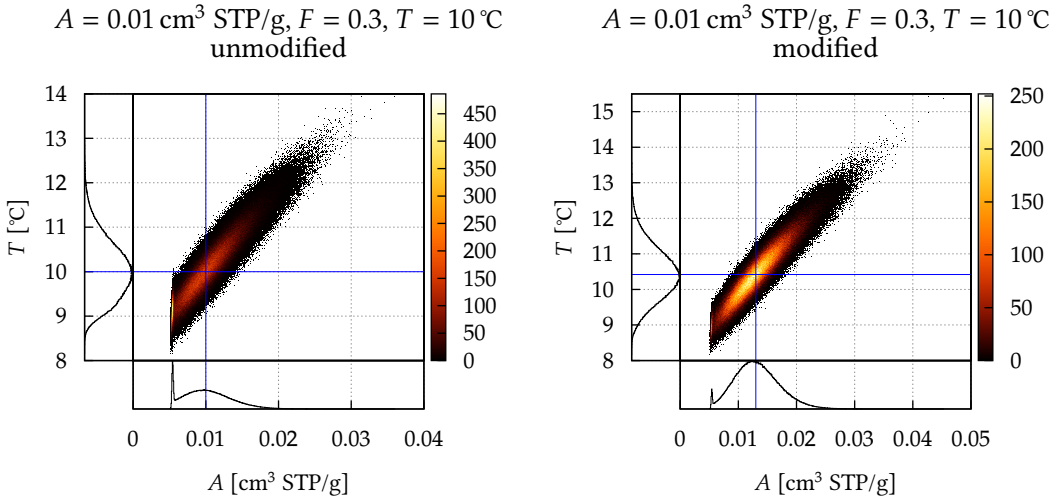
whereas Kr concentrations were both increased and decreased (–1.2 to +1.0 %). It is very interesting to note that Sun et al. (2010) described the same kind of deviations (measured Ar higher and Xe lower than modeled) as a characteristic misfit pattern of the UA model. The CE model appears to be able to fit samples with comparatively high Ar and low Xe better, but in the extreme, such deviations are leading to the occurrence of fits with large A. Since large values of A in the CE model correspond to a pressure effect, this finding also seems consistent with the effect described by Castro et al. (2007) and Sun et al. (2010), that the characteristic UA model misfit (high Ar, low Xe) was decreased in the OD model with increasing pressure parameter.

3.4. Synthetic samples

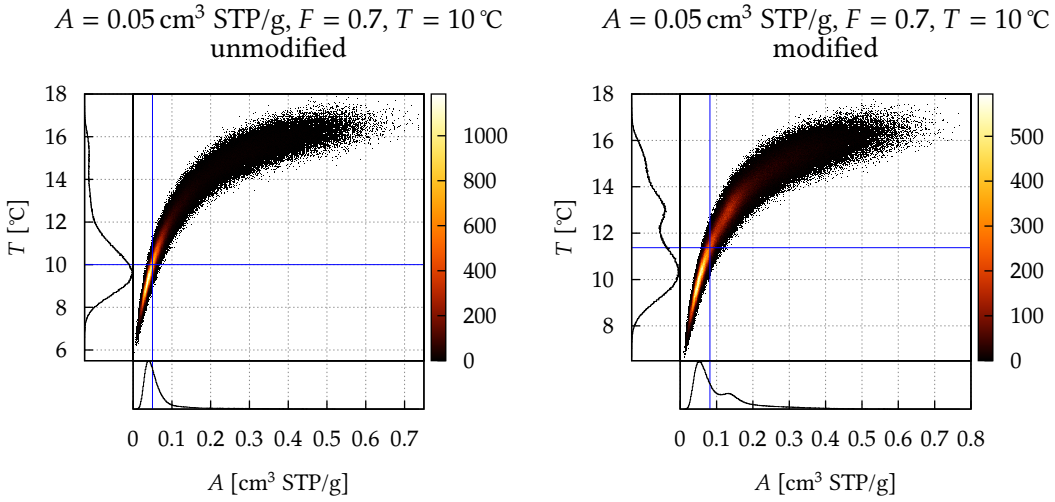
In order to simulate and possibly reproduce the behavior of the problematic samples, the Monte Carlo analysis was applied to synthetic samples. 27 samples were created for the following parameter values: A = 0.01, 0.05, 0.1 cm^3 STP/g; F = 0.3, 0.5, 0.7; T = 10, 20, 30 °C; P = 1 atm; S = 0. Note that this choice of parameter values in particular spans the range from low to rather critically high values of A. The uncertainties of the Ne, Ar, Kr and Xe concentrations were chosen to be 0.5 %, 0.5 %, 1 % and 3 % respectively for the Monte Carlo simulations, according to typical values for current noble gas analyses in our laboratory at Heidelberg University. It should be mentioned that the Xe uncertainty is relatively large and some of the following results may differ if a smaller Xe uncertainty was assumed.

The left column of Fig. 3.3 shows histograms of three representative synthetic samples, going from low values of all parameters (Fig. 3.3a) to higher values (Fig. 3.3b and c). The first example shows a well constrained range of variability in A and T, superimposed by a group of Monte Carlo realizations exhibiting the clustering at $F = 0$ and very low A. Apart from this group, the Monte Carlo solutions are normally distributed around the original parameter values (which are correctly retrieved by fitting the original synthetic samples in all cases). It is noteworthy, however, that the solutions with $F = 0$ (UA model) yield too low estimates of A and T. This may have been expected for the parameter A, which has a different physical meaning in the case of the UA model (finally dissolved air rather than initially entrapped air), but it is a potential problem with regard to the estimation of T. It implies that care should be taken in combining solutions with $F = 0$ with true CE model solutions in real data sets.

More important for the main theme of this investigation is the behavior of the Monte Carlo realizations for synthetic samples with higher values of A, F and T, as shown in Fig. 3.3b and c. Here, similar to the finding for most of the well-behaved real samples, a significant number of Monte Carlo realizations provides solutions in the problematic area with high values of A. Correspondingly, the 1D histograms for A and T deviate from normal distributions with a tail or even a weakly expressed second peak on the side of high values of A and T. As a result, the mean estimated temperature from the Monte Carlo analysis deviates from the true temperature of the original synthetic samples, with a bias towards higher T. This bias is very low for the sample shown in Fig. 3.3a

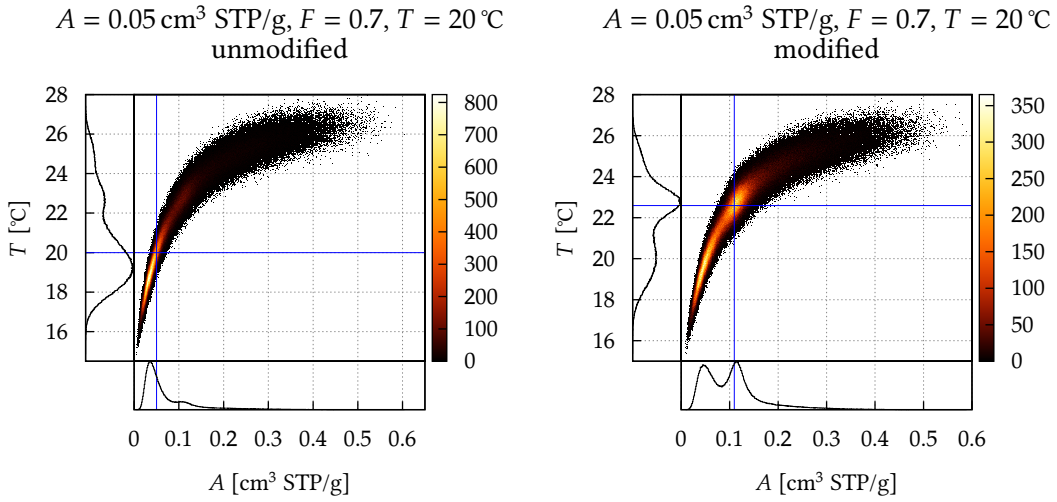


(a) Low excess air, fractionation and temperature.



(b) High excess air and fractionation, low temperature.

Figure 3.3.: Histograms of the Monte Carlo analyses of the synthetic samples (continued on the following page).



(c) High excess air, fractionation and temperature.

Figure 3.3 (cont.): Histograms of the Monte Carlo analyses of the synthetic samples. The left column shows unmodified samples, the right column shows the effect of increasing Ar concentrations by 0.5 % and decreasing Xe concentrations by 1.5 %. The blue lines represent the samples' fit results without Monte Carlo simulations. For sample a, one can see that the Monte Carlo results are normally distributed around the fit result for both the unmodified and the modified sample (except for a few cases with low A and $F = 0$). If A and F are increased (sample b), a few Monte Carlo realizations appear in the problematic domain. With the modifications applied they even form a small distinct peak. The fit result, however, still does not completely move into the problematic area. If the temperature is increased, too (sample c), the second cluster becomes more dominating and, in the modified case, also contains the fit result.

(mean T = 10.03 ℃ instead of exactly 10 ℃). It becomes noticeable for the synthetic samples shown in Fig. 3.3b (mean T = 10.43 ℃ instead of 10 ℃) and Fig. 3.3c (mean T = 20.52 ℃ instead of 20 ℃). This analysis confirms the bias of the CE model as a temperature estimator as described by Sun et al. (2010), although their bias is more than twice as high than in this analysis, which is likely due to the fact that different errors were assumed for the noble gas concentrations. It also confirms the expectation that the bias comes from the tendency of the CE model to produce solutions with unrealistically high A and correspondingly overestimated T for some specific patterns of noble gas concentrations.

As deduced from the analysis of problematic real samples, the solutions with high A occur for high concentrations of Ar combined with low concentrations of Xe. In order to study the effect of such systematic deviations, each synthetic sample was also analyzed with Ar increased by 0.5 % and Xe decreased by 1.5 % (see Table 3.3). The right column of Fig. 3.3 shows histograms of the three representative synthetic samples discussed above after applying these modifications. Samples with low values of A, F and T, like in Fig. 3.3a, do not show important changes in their behavior as a result of the applied modification, except for a slight shift of the mean parameter values. The prevalence of solutions with $F = 0$ was reduced in the example shown here, but this has little effect on the mean values and errors of the parameters deduced from the Monte Carlo procedure. For higher values of A, F and T, like in Fig. 3.3b and c, where even the unmodified synthetic samples have Monte Carlo realizations in the problematic area, the modified samples tend to produce a second cluster at high A. In the modified sample shown in Fig. 3.3b one can already see a second cluster forming, but the fit result is still acceptable, although the parameter errors already increase considerably. If the temperature is increased to 20 ℃ as in Fig. 3.3c, the second cluster becomes more pronounced and now also contains the fit result, which shows the same large A and T uncertainties as the problematic physical samples.

Fig. 3.4 shows how the histograms change if the noble gas concentrations are modified step by step for the example of the synthetic sample of Fig. 3.3c. One can see the "wrong" maximum gradually grow whereas the "true" maximum becomes smaller. This result shows that the applied systematic concentration changes (increased Ar and decreased Xe) do not just gradually shift the results of the CE model fits, but lead to a transition from a "normal" behavior with well defined maxima of A and T close to the values used to produce the original synthetic samples towards a "problematic" behavior with two separate maxima, one of them occurring at unrealistically large A and much higher than original T.

It is very interesting to note, however, that the clusters and corresponding peaks with low values of A and T still fall close to the original parameter values, even after the modification of the concentrations of the synthetic samples. It appears as though this lower cluster retains the information from the original synthetic sample. This finding is intriguing in view of the behavior of the problematic physical samples. There, too, a lower cluster with potentially meaningful values of A and T occurs (Fig. 3.2), and the question arises if these "realistic" clusters may represent the "true" parameter values for these samples.

Table 3.3.: Fit results of a number of synthetic samples after increasing Ar by 0.5 % and decreasing Xe by 1.5 %

Parameters used to create sample			Fit results after modification				
A [cm^3 STP/g]	F	T [℃]	χ^2	A [cm^3 STP/g]	F	T [℃]	ΔT [℃]
0.01	0.3	10	0.08	0.013 ± 0.004	0.38 ± 0.08	10.4 ± 0.7	0.4
0.01	0.3	20	0.07	0.013 ± 0.003	0.37 ± 0.07	20.6 ± 0.9	0.6
0.01	0.3	30	0.06	0.013 ± 0.003	0.37 ± 0.07	30.8 ± 1.2	0.8
0.01	0.5	10	0.07	0.015 ± 0.006	0.58 ± 0.07	10.4 ± 0.8	0.4
0.01	0.5	20	0.07	0.014 ± 0.006	0.57 ± 0.06	20.6 ± 1.0	0.6
0.01	0.5	30	0.06	0.014 ± 0.006	0.57 ± 0.06	30.8 ± 1.3	0.8
0.01	0.7	10	0.07	0.019 ± 0.013	0.77 ± 0.05	10.5 ± 0.8	0.5
0.01	0.7	20	0.06	0.019 ± 0.014	0.77 ± 0.05	20.7 ± 1.1	0.7
0.01	0.7	30	0.05	0.020 ± 0.015	0.76 ± 0.05	30.9 ± 1.5	0.9
0.05	0.3	10	0.03	0.055 ± 0.006	0.31 ± 0.01	10.8 ± 1.2	0.8
0.05	0.3	20	0.02	0.056 ± 0.008	0.31 ± 0.01	21.3 ± 1.9	1.3
0.05	0.3	30	0.01	0.058 ± 0.011	0.31 ± 0.01	32.1 ± 3.1	2.1
0.05	0.5	10	0.01	0.061 ± 0.014	0.51 ± 0.01	11.0 ± 1.5	1.0
0.05	0.5	20	0.00	0.065 ± 0.023	0.51 ± 0.01	21.7 ± 2.6	1.7
0.05	0.5	30	0.00	0.076 ± 0.062	0.51 ± 0.02	33.3 ± 6.8	3.3
0.05	0.7	10	0.00	0.082 ± 0.062	0.71 ± 0.01	11.4 ± 2.4	1.4
0.05	0.7	20	0.03	0.109 ± 0.337	0.71 ± 0.02	22.6 ± 10.5	2.6
0.05	0.7	30	0.24	0.096 ± 0.367	0.71 ± 0.03	32.3 ± 14.1	2.3
0.1	0.3	10	0.00	0.115 ± 0.021	0.30 ± 0.00	11.8 ± 2.7	1.8
0.1	0.3	20	0.03	0.129 ± 0.064	0.30 ± 0.01	23.8 ± 7.9	3.8
0.1	0.3	30	0.32	0.127 ± 0.115	0.30 ± 0.01	33.9 ± 15.4	3.9
0.1	0.5	10	0.09	0.145 ± 0.130	0.50 ± 0.00	12.5 ± 6.4	2.5
0.1	0.5	20	0.43	0.134 ± 0.201	0.50 ± 0.01	22.1 ± 11.2	2.1
0.1	0.5	30	0.70	0.115 ± 0.202	0.50 ± 0.02	31.0 ± 14.2	1.0
0.1	0.7	10	0.44	0.151 ± 0.332	0.70 ± 0.01	11.2 ± 7.0	1.2
0.1	0.7	20	0.69	0.125 ± 0.378	0.70 ± 0.01	20.6 ± 10.4	0.6
0.1	0.7	30	0.78	0.105 ± 0.379	0.70 ± 0.02	30.0 ± 13.6	0.0

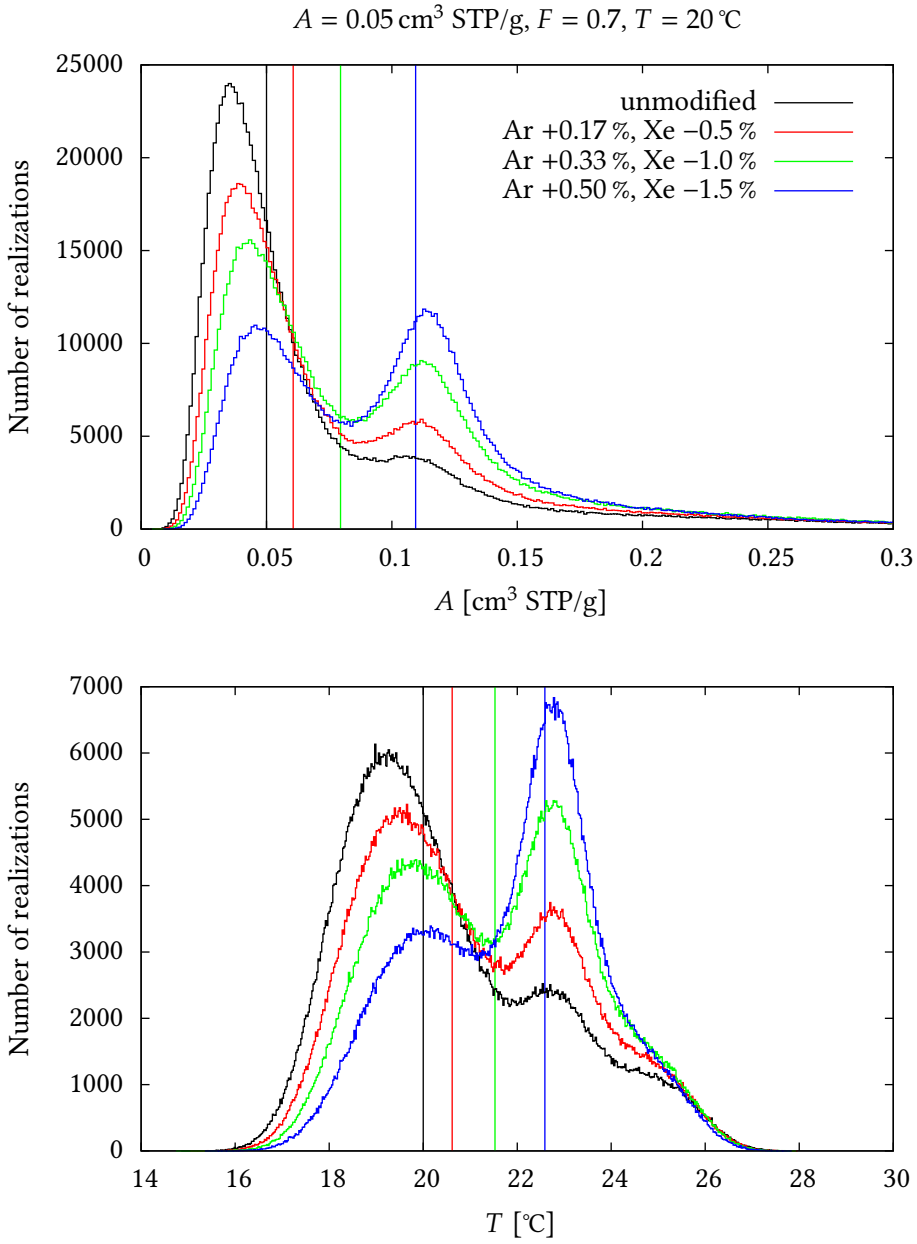


Figure 3.4.: Effects of incremental modifications to the noble gas concentrations for the last synthetic sample of Figure 3.3. The plots show histograms of the Monte Carlo results for the parameters A and T. The vertical lines represent the samples' fit results. One can see that a second peak emerges as the noble gas concentrations are altered step by step.

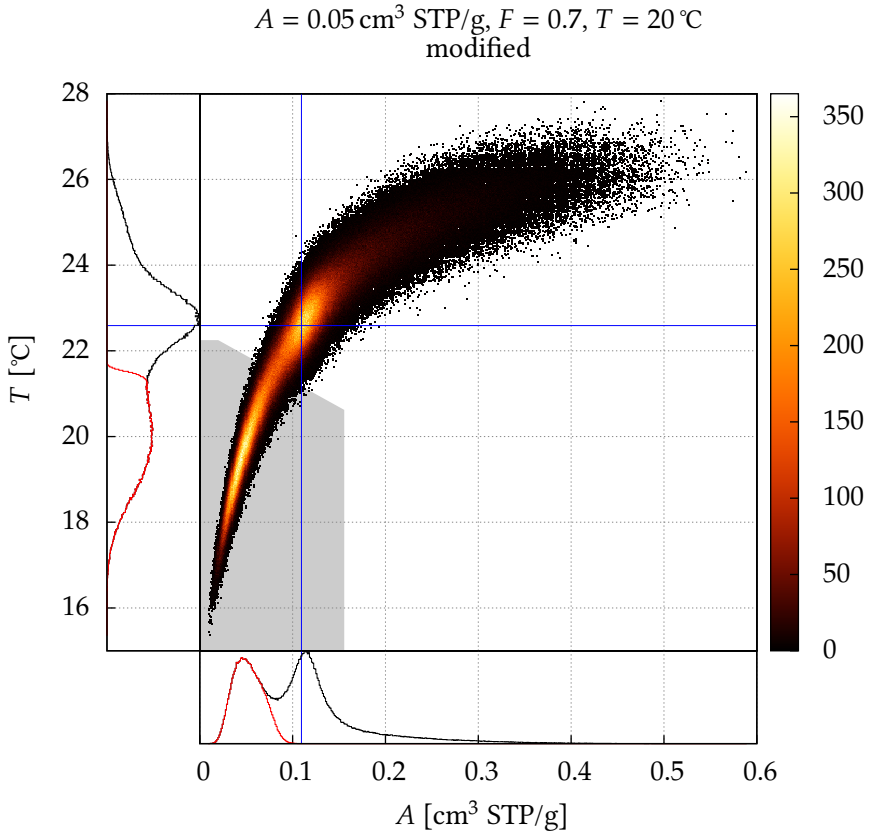


Figure 3.5.: Illustration of the proposed approach for handling poorly-fitting samples for the modified synthetic sample of Figs. 3.3c and 3.4. The Monte Carlo realizations are separated into two clusters, one with realistic and one with unrealistic parameter values. The statistical analysis is then reduced to the realistic cluster (shaded in gray here). The 1D histograms show both the original distribution (black) as well as the distribution of the remaining Monte Carlo realization (red).

In order to find out if it is permissible to restrict the analysis to the Monte Carlo realizations leading to fit results in the more realistic cluster, mean value and standard deviation of these subsets were determined for several synthetic samples. Fig. 3.5 illustrates this process for the modified sample shown in Fig. 3.3c. Based on the 2D histogram obtained from the Monte Carlo analysis, a separation of the realizations into two groups corresponding to the two clusters is performed. The separation line is drawn at the saddle point between the two maxima, orthogonally to their connection line. Only the realizations corresponding to the realistic cluster are retained for a statistical analysis of mean values and standard deviations of the parameters. The results obtained in this way agreed very well with the parameter values used to generate the synthetic samples, i.e., the deviations were less than the standard deviation, often even less than 0.2 ℃.

3.5. Application to poorly-fitting samples

The analysis of synthetic CE model samples shows that systematic modifications of the concentrations (+0.5 % for Ar and −1.5 % for Xe) can lead to the observed fitting problems. These deviations are less than typical measurement uncertainties of the noble gas laboratory in Heidelberg, which means that a few per cent of the measured samples will show these problems simply for statistical reasons. Furthermore, the problems occurred more often for higher values of the parameters A, F and T, meaning that samples taken in regions with warmer climate are more likely to be affected. The analyses of the modified synthetic samples suggest that the problematic samples can still be evaluated—if they show two separate clusters of parameters—by limiting the statistical analysis to the cluster with the more realistic values.

Table 3.4.: Results obtained by restricting the evaluation to the cluster of Monte Carlo realizations with realistic values for A and T

Sample	T [℃]	A [cm^3 STP/g]	F
MAS_21	10.68 ± 0.48	0.026 ± 0.014	0.85 ± 0.17
MD6.2	12.88 ± 0.42	0.049 ± 0.027	0.88 ± 0.05
TX46.2	18.74 ± 0.79	0.026 ± 0.020	0.74 ± 0.26
IND32_1	28.78 ± 0.65	0.025 ± 0.016	0.84 ± 0.16

Table 3.4 shows the results of this method when applied to the poorly-fitting samples MAS_21, MD6.2, TX46.2, and IND32_1. The results seem to be realistic and compare very well to the results of the corresponding normal samples MAS_24, MD9.1, TX46.3, and IND32_3, respectively (see Table 3.2). This result supports the conclusion from the analysis of synthetic samples, that it seems to be reasonable to restrict the analysis of problematic fits with the CE model to the subset of Monte Carlo realizations that lie in the range of realistic values of A. In order to investigate the potential of this approach, it was applied to a larger number of problematic samples than discussed here. In general, this yielded promising results. However, there were also some samples where this method did not work, because these samples did not show the characteristic splitting

behavior. Their Monte Carlo results did not show two clusters but were all concentrated in areas of physically unrealistic parameter values and the samples could thus not be evaluated in the way proposed here.

4. Recommended sample evaluation process

The evaluation of groundwater noble gas data by inverse modeling tools requires a certain expertise to judge the obtained results and choose a reasonable evaluation strategy. As Sun et al. (2010) have demonstrated in detail, there are systematic offsets between the noble gas temperatures estimated with different excess air models, but the temperature variation within a data set is robust with regard to the model choice. It is therefore strongly recommended to evaluate a given data set with only one model for all samples. As the new software *PANGA*, which will be introduced in chapter 6, adds additional ways to treat individual samples, even more care should be taken to obtain a consistent and appropriate evaluation of any given data set. It is therefore recommended to start with simple, traditional evaluation methods and only use the advanced models and features such as Monte Carlo simulations and ensemble fitting when clearly indicated. The following sections contain the description of a recommended step-by-step evaluation approach for groundwater samples, with a particular emphasis on the application of the CE model, which probably has been the most frequently used model in the past. This evaluation approach also takes into account the special cases, which appear regularly, like the one discussed in detail in chapter 3 or UA limit cases.

4.1. Step 1: UA model fits

In the beginning, fitting the UA model, which is the most simple and also numerically most stable excess air model, can give helpful information about the samples. For example, very small or very large values of the parameter A may indicate unusual excess air conditions (possibly due to equilibration or air contamination during sampling), whereas negative A values indicate that the samples were probably affected by degassing. The χ^2 values obtained by the UA model also provide some orientation. Generally low χ^2 values with corresponding probabilities larger than 1 % may even supersede the need for further evaluations, but this is rarely the case. Somewhat higher χ^2 values call for more complex models accounting for excess air fractionation with an additional parameter, which in general can improve on the UA-model fits. Very high χ^2 values from the UA model, however, may indicate some problem with the data that no model will be able to explain.

4.2. Step 2: CE model fits

As a next step, fitting the CE model is suggested, which can describe both excess air and degassing and contains other models as limiting cases (Aeschbach-Hertig et al. 2008). Information from fitting the UA model in the first step should be taken into account for the selection of initial fit parameters. Initial values of $A = 0.01\,\mathrm{cm^3\,STP/g}$ and $F = 0.5$ usually worked very well for samples with excess air, whereas $F = 3$ seems to be a good choice for degassed samples, in order to find the minimum with $F > 1$ corresponding to degassing. The temperature initial values should be close to the expected paleotemperatures or somewhat lower to prevent the fitter from falling into a second local χ^2 minimum at higher temperatures and unrealistically high A values which is present in most samples as described in section 3.2. For many samples, standard CE model fitting works very well and provides sound parameter and uncertainty estimates as well as reasonable χ^2 values. As a guidance, we consider the following range of CE model parameters and uncertainties as reasonable: $A < 0.05\,\mathrm{cm^3\,STP/g}$, $0.1 \leq F < 1$ and $\Delta T < 2^{\circ}\mathrm{C}$. The χ^2 fit probability (cf. equation 2.43) should be greater than 1 %.

In order to find out whether or not Monte Carlo fits would significantly improve the parameter estimates and their uncertainty in case of such samples with "well-behaved" CE model fits, around 250 samples from five different datasets were analyzed. A comparison of fit results with estimates in the standard CE model range with their respective Monte Carlo counterparts showed good agreement of the temperature parameter in almost all cases. In two cases the mean Monte Carlo temperature was about 0.7 to 0.8 ℃ lower than the fit temperature. In five cases it was about 0.4 ℃ lower. For three other cases the temperature estimates agreed well, but the Monte Carlo temperature error was about 0.4 ℃ lower. So, for most well-behaved samples, Monte Carlo analysis does not improve parameter estimation and may be omitted. In such cases, the analysis may be terminated at this point, unless there are reasons such as suspected temperature biases or oxygen depletion that motivate the exploration of further models.

4.3. Step 3: Monte Carlos fits

It is recommended to carry out Monte Carlo fits if the CE model fit results do not meet the criteria given above. Among these samples, two special patterns are quite common: The first case, which has been discussed in detail in chapter 3, is characterized by too high temperature estimates in conjunction with large temperature uncertainties. In the second case, which corresponds to a UA limit case and is discussed in more detail below, the results exhibit highly negative F values (typically between −1000 and −5000) combined with low A values (less than $10^{-6}\,\mathrm{cm^3\,STP/g}$) or, in rare cases, with interchanged signs, i.e., high positive F together with small negative A. If the fit was carried out in constrained mode, this case is characterized by $F = 0$. It is identical to the special case which already appeared in chapter 3. In both cases, Monte Carlo analyses and their appropriate evaluation, as demonstrated later in this section, can help to obtain better results.

Table 4.1.: Noble gas concentrations of the samples illustrating the different Monte Carlo cases and the pressures and salinities used for their fits.

Sample	Ne [10^{-7} cm^3 STP/g]	Ar [10^{-4} cm^3 STP/g]	Kr [10^{-8} cm^3 STP/g]	Xe [10^{-9} cm^3 STP/g]	P [atm]	S [g/kg]
BE_8	2.726 ± 0.035	4.955 ± 0.051	11.360 ± 0.172	16.630 ± 0.238	0.9976	0.596
IND_25_3	2.724 ± 0.015	3.907 ± 0.031	7.950 ± 0.167	9.688 ± 0.135	0.9904	0.0
IND_19_2	2.363 ± 0.013	3.939 ± 0.032	8.504 ± 0.121	10.664 ± 0.230	0.9868	0.0
TX48.2	2.202 ± 0.029	3.282 ± 0.043	7.127 ± 0.107	9.903 ± 0.218	0.9762	0.0
CN_16_2	1.392 ± 0.006	3.004 ± 0.014	7.638 ± 0.126	10.898 ± 0.467	1.0	0.0

It is recommended to begin with unconstrained fits to avoid any problems in the boundary area which could be caused by the constraints. Only if the results indicate it, the fits should be repeated in constrained mode.

4.4. Examples of common Monte Carlo cases

In the following sections a list of the most common patterns appearing in the evaluation process will be given together with recommendations on how to handle them. To illustrate the different cases the following set of samples will be used: BE_8 (just called 8 in the article) was taken from the Belgium dataset by Blaser et al. (2010), which will be reviewed in section 5.10. TX48.2 is a sample from the Texas data set by Castro et al. (2007) which was already used in chapter 3. IND_25_3, IND_19_2 are two samples from a sedimentary aquifer system in the North Cambay Basin in Northwest India (Wieser 2011). The last sample, CN_16_2, is from a well in the Beijing area and part of a recent study in the North China Plain (Schneider 2014).

Table 4.2.: Results of the initial UA model fit

Sample	χ^2	Probability [%]	A [10^{-3} cm^3 STP/g]
BE_8	9.3	0.9	3.31
IND_25_3	91.6	0	4.85
IND_19_2	55.4	$9.4 \cdot 10^{-11}$	2.40
TX48.2	1.1	57.6	2.09
CN_16_2	23.0	$1.0 \cdot 10^{-3}$	−2.74

Table 4.2 shows the results of an initial UA model fit. The A estimates indicate that all samples but CN_16_2 have excess air. CN_16_2, however, has negative A, meaning that degassing has probably occurred. For the samples IND_25_3, IND_19_2 and CN_16_2, the fit probabilities are very low. The χ^2 probability of BE_8 is almost acceptable. TX48.2 is one of the rare cases where the probability is very good, so that further fitting might not be necessary. Nevertheless, all samples will be further analyzed here, to show their different behaviors with respect to CE model Monte Carlo fitting. As they have excess air, all samples but the last one should be fitted with an initial F smaller than 1. The

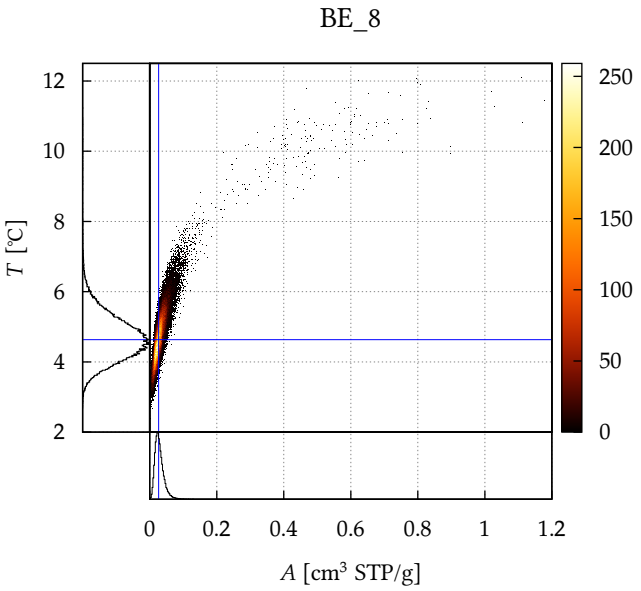


Figure 4.1.: Normal case: Monte Carlo analysis is not necessary.

remaining sample, CN_16_2, should have an initial F greater than 1. The CE model fit can now be used to distinguish between the different cases.

4.4.1. Normal case

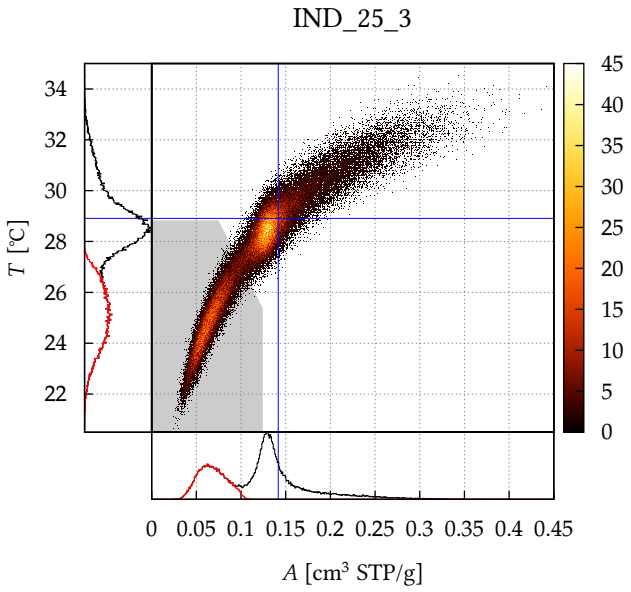
BE_8 (Fig. 4.1) shows the standard case: All parameter estimates and uncertainties look realistic and are in the above-mentioned range. Monte Carlo analysis is usually not necessary in this case and will probably only show small deviations (cf. Table 4.3).

4.4.2. High A in combination with large temperature uncertainties

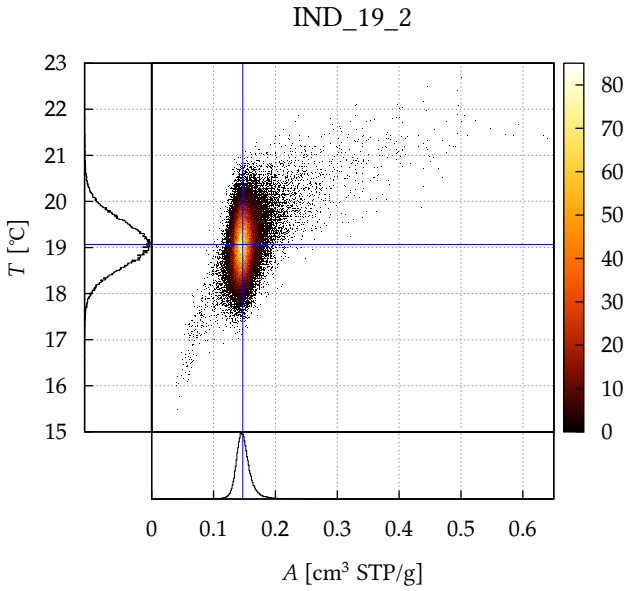
The two samples from India, IND_25_3 and IND_19_2, both show high values of A in combination with large temperature uncertainties. This case is usually accompanied by too high temperature estimates, which may, however, still be in the realistic range.

Table 4.3.: Comparison of noble gas temperatures obtained from the CE model fit and from Monte Carlo analysis

Sample	T_{fit} [°C]	T_{MC} [°C]
BE_8	4.63 ± 0.61	4.70 ± 0.67
IND_25_3	28.9 ± 13.9	25.05 ± 1.24
IND_19_2	19.1 ± 10.6	—
TX48.2	19.13 ± 0.88	19.27 ± 0.55
CN_16_2	11.89 ± 1.53	11.81 ± 1.44



(a) Most common subcase. NGT can be determined by restricting the analysis to the sector shaded in gray.



(b) Rare subcase. NGT cannot be determined, because almost all Monte Carlo results are concentrated in the area of unrealistic values.

Figure 4.2.: High A in combination with large temperature uncertainties

These types of samples need to be analyzed using Monte Carlo fits. Their histograms normally show a split-up of the A and T parameters into two peaks, or, in the two-dimensional A-T histogram, the formation of two more or less separate clusters, one of which is at unrealistic parameter values, in particular high values of A, and contains the original fit result. The other cluster shows more realistic A values and, as described in more detail in chapter 3, corresponds to the "true" parameter values. Therefore, the evaluation needs to be restricted to the realistic cluster, whereas the Monte Carlo realizations in the unrealistic cluster are dropped. In Fig. 4.2a, this process is illustrated for the sample IND_25_3. However, in rare cases like IND_19_2, which is shown in Fig. 4.2b, this approach is not feasible because all or almost all Monte Carlo results lie in the cluster at unrealistic values. For these samples it is impossible to retrieve the correct results as the realistic cluster is not visible.

4.4.3. UA limit case

The CE model includes two different UA model limit cases. The first one is $F = 0$, in which case the CE model equation simplifies to $c_i^{\mathrm{eq}} + Az_i$, which is the UA model. The other case is less obvious and appears in the limit of $|F| \gg 1$:

$$c_i = c_i^{\mathrm{eq}} + \frac{(1-F)\cdot Az_i}{1+\frac{FAz_i}{c_i^{\mathrm{eq}}}} \overset{|F|\gg 1}{\approx} c_i^{\mathrm{eq}} + \frac{-FAz_i}{1+\frac{FAz_i}{c_i^{\mathrm{eq}}}}. \tag{4.1}$$

If, additionally, $|FAz_i/c_i^{\mathrm{eq}}| \ll 1$, which generally is the case for $|FA| \ll 0.01\,\mathrm{cm^3\,STP/g}$, this can further be simplified to

$$c_i \approx c_i^{\mathrm{eq}} - FAz_i. \tag{4.2}$$

This is identical to the UA model with A replaced by $-FA$.

If unconstrained fits are used, the latter case is the one appearing most of the time. However, because the approximation in equation 4.2 is not justified in all cases, the results correspond to some sort of mathematical mixing of the UA and CE models, which does not need to have any physical meaning. It is therefore advisable to use constrained fits for these types of samples and thereby prevent the fit from falling into the cases with highly positive or highly negative F. Fig. 4.3 shows Monte Carlo results of the sample TX48.2 which is an example for the UA limit case. Monte Carlo statistics typically reproduce the fit temperature very well but lead to lower temperature uncertainties and should therefore be carried out whenever this limit case occurs. Even though this is an UA model limit case, the UA model should not be used for the statistics, as not all CE model Monte Carlo results fall into this limit case and therefore contribute to a higher temperature error (which is still lower than the CE model fit error).

4.4.4. Combination of high A and UA limit cases

The UA limit case sometimes appears together with the case of high A, i.e., the direct fit of the sample shows either behavior but the Monte Carlo histograms show clusters

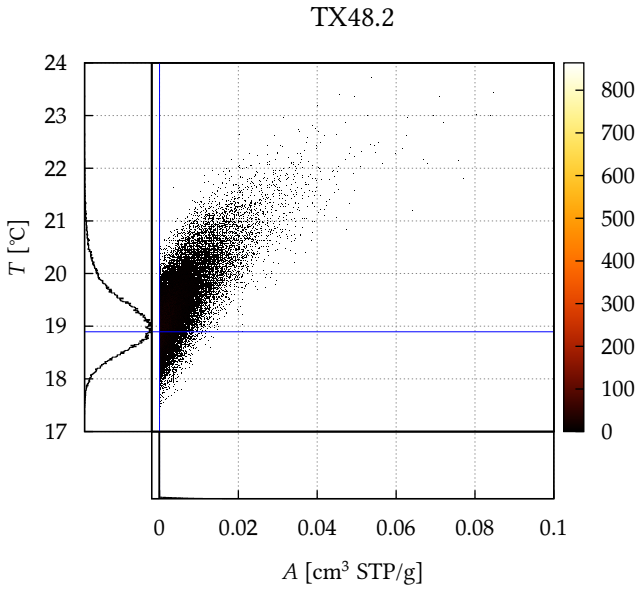


Figure 4.3.: UA limit case: a large portion of the Monte Carlo realizations is in the UA limit case with $F = 0$. This plot shows a zoomed-in view of the results of a constrained fit. A few Monte Carlo realizations lie in the area of higher A and T values, similar to sample BE_8 in Fig. 4.1.

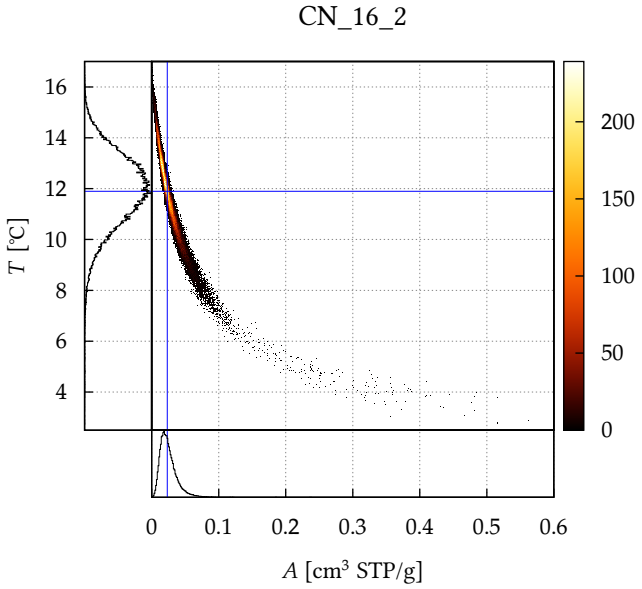


Figure 4.4.: Degassed case

at high A as well as in the UA limit case. For this type of sample the analysis should be restricted to the normal and the UA limit case clusters, dropping the cluster at high A values. A pair of examples for this case was analyzed in section 3.3: the two replicate samples IND32_3 and IND32_1 in Fig. 3.2.

4.5. Degassed samples

Degassed samples exhibit F values larger than 1 and the corresponding CE model solutions are obtained by using an initial value of $F > 1$. Besides this, they show the same special Monte Carlo patterns as samples with excess air: in addition to the normal case they can have high A values, be in the UA limit case (here normally with large, positive F) or show a combination of both. This means that the same evaluation methods as for samples with excess air may be tried. These cases were, however, not analyzed in this work. It is therefore recommended to use the more conservative approach of using all Monte Carlo realizations for the statistical analysis.

In addition, in the case of degassed samples the PD model based on diffusion-controlled degassing should also be tried. The experience of our group shows that often the CE model degassing case is able to provide better fits to degassed samples. In some cases, however, the PD model can describe the measured noble gas concentrations more accurately.

5. Review of literature data

The following sections contain a review of important groundwater noble gas data sets from the literature. All data sets were re-evaluated using the methods laid out in chapter 4.

Samples whose χ^2 values have a probability of less than 1 % (i.e., $\chi^2 \geqslant 6.64$ for one degree of freedom, which is the normal CE model case) were excluded from the analysis, because their results are considered to be too bad a fit. For samples evaluated using Monte Carlo simulations, χ^2 was determined from the mean of the distribution of $\chi = \sqrt{\chi^2}$, because this distribution normally has a symmetric shape, in contrast to the χ^2 distribution, which is typically skewed to the right.

The most important statements in the evaluated studies relate to temperature differences between Holocene samples and samples from the last glacial period, especially the Last Glacial Maximum (LGM). In order to calculate temperature differences, a method has to be chosen to determine mean temperatures for a specific time period. For this re-evaluation, the samples of one time period were averaged with the weighted arithmetic mean, which is the maximum likelihood estimator of the mean value of multiple Gaussian probability distributions sharing the same mean. The weighted arithmetic mean and its variance are calculated according to (Bard 1974):

$$\bar{T} = \frac{\sum_{i=1}^{n} T_i \cdot \sigma_{T_i}^{-2}}{\sum_{i=1}^{n} \sigma_{T_i}^{-2}} \tag{5.1}$$

$$\mathrm{var}(\bar{T}) = \frac{1}{\sum_{i=1}^{n} \sigma_{T_i}^{-2}}. \tag{5.2}$$

For this calculation the measurement errors of the individual samples need to be adapted, as they only include the statistical uncertainty of the sampling and measuring processes. The "real" paleotemperatures of the samples (as opposed to the measured ones) are, however, also distributed in some way around the mean temperature of the relevant time period. To estimate this temperature variability, the distribution of NGTs in the data sets Stute 1995 Brazil, Aeschbach-Hertig 2002 Maryland and Kreuzer 2009 China was analyzed and compared to their measurement errors with an (unweighted) least-squares fit minimizing $S = \sum_i (\sigma_{T_i}^2 + V_T - V_{tot,i})^2$. The index i traverses the samples of all three data sets in the specific time period. σ_{T_i} denotes the temperature error of a specific sample; $V_{tot,i}$ is the variance of the samples from the data set to which sample i belongs. V_T is the variance of the searched-for temperature variability. The associated standard deviation, i.e., the square root of the variance, was estimated to be about 0.8 °C for both the Holocene and the late Pleistocene. To account for the temperature variability, this value was added in square to the measurement error before plugging it into

equation (5.1). If multiple samples from one well were present, they were combined by calculating their weighted average before the 0.8 ℃ temperature variability was added.

For all plots in this chapter, the following conventions are used: data points that were accepted in the re-evaluation, based on their χ^2, are shown in black or fully saturated colors. Samples not accepted in the re-evaluation are drawn in gray or pale colors. If a sample was not accepted in the original evaluation, it is marked with an open symbol. Samples that were originally included are depicted with closed symbols. Samples excluded from both analyses are sometimes omitted from the plots. The different symbols indicate the way the samples were evaluated: circles mean that the results were determined from a normal fit. A triangle with its tip pointing upwards means that Monte Carlo simulations with constrained fits were applied, while a triangle with its tip facing downwards means that the sample was evaluated using Monte Carlo simulations with unconstrained fits.

The fit results of this re-evaluation are all summarized in appendix C.

5.1. Stute 1995 Brazil

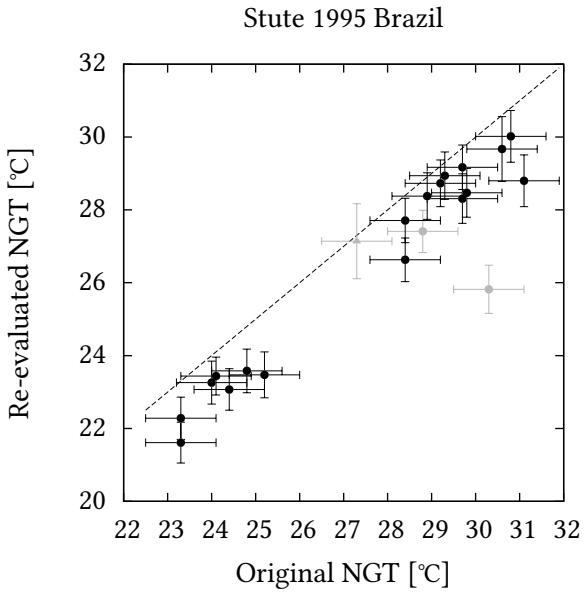


Figure 5.1.: Stute 1995 Brazil: comparison of original and re-evaluated NGTs

This set of 21 samples from the Maranhão Basin in the Piaui Province, northeastern Brazil, was published by Stute et al. (1995). It was used to introduce the PR model, as the UA model did not result in good agreement among the temperatures derived from the different noble gases. In this study, the model parameters have not been determined using the least-squares method, yet. Instead, they were estimated by minimizing the standard deviations of the temperatures derived from the individual noble gases after

correction for PR-fractionated excess air. The re-evaluation was carried out using a pressure calculated from a recharge altitude of 450 m and a salinity of zero. An evaluation with the CE model resulted in consistently lower than original noble gas temperatures (Fig. 5.1). For most samples, the deviation lies between 0.5 ℃ and 2.0 ℃. The estimated temperature for sample 17, however, is about 4.5 ℃ lower.

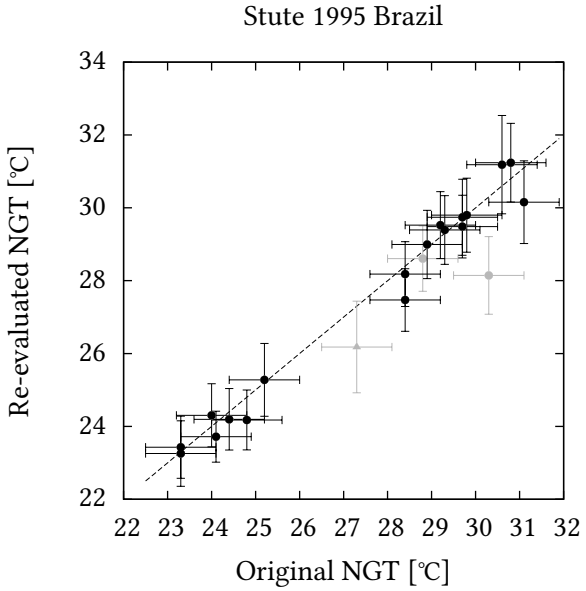


Figure 5.2.: Stute 1995 Brazil: comparison of original NGTs and the NGTs determined from a second re-evaluation with the PR model.

Fig. 5.2 shows the results of a PR χ^2 fit compared to the original results. With a few exceptions, which show about 1 to 2 ℃ lower temperatures, the original results were reproduced quite well. The deviations were probably caused by the different estimation method.

Sample H10 was excluded from the analysis because its χ^2 value was considered too high. For sample 4, no Ne concentrations were given. Therefore the fit had zero degrees of freedom and its results could not be tested with a χ^2 test. Because of this, the sample was excluded from the analysis as well. Sample 17, even though it is in the warm cluster, was not considered for the Holocene LGM temperature difference, because it probably originated from a warm period before the LGM.

The re-evaluation yielded a mean temperature of the Holocene of (28.6 ± 0.3) ℃. The mean glacial temperature was found to be (23.0 ± 0.4) ℃. Combined, these results give a temperature difference of (5.6 ± 0.5) ℃ between Holocene and LGM, which is in very good agreement with the original result of (5.6 ± 0.6) ℃.

Fig. 5.3 shows a comparison of the re-evaluated results with the re-evaluation of Aeschbach-Hertig et al. (2000), which was used to introduce the CE model. The results of both re-evaluations agree very well with the exception that the NGTs by Aeschbach-

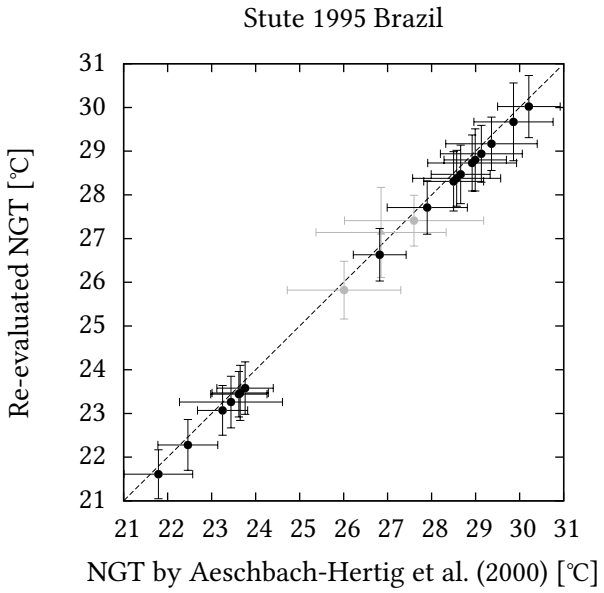


Figure 5.3.: Stute 1995 Brazil: comparison of the re-evaluated NGTs with the re-evaluation carried out by Aeschbach-Hertig et al. (2000)

Hertig et al. (2000) are consistently higher by about 0.2 °C. This is due to the fact that they were using 400 m as the presumed recharge elevation instead of the 450 m used here. Aeschbach-Hertig et al. (2000) found a temperature difference between Holocene and LGM of (5.6 ± 0.4) °C, which is almost identical to the (5.6 ± 0.5) °C found here.

5.2. Beyerle 1998 Switzerland

This data set from Switzerland was published by Beyerle et al. (1998). It consists of eight samples taken from boreholes tapping the Glatt Valley aquifer. In the original evaluation, the NGT was estimated with a least-squares fit of the UA model. The presumed recharge elevation was given as 540 m. Salinity values for the original evaluation were not given. For the re-evaluation they were assumed to be zero. The waters from boreholes 2, 4 and 6 were mixtures of younger and older components which could be separated using multiple tracers. For the re-evaluation, these three samples were used after being corrected for the older component.

Fig. 5.4 shows the results of the re-evaluation compared to the original results. For the samples 1, 2 and 6, the original results were reproduced quite well. The samples 0 and 4, however, show by 0.8 to 1.3 °C higher NGTs. The samples 3, 5 and 7 were originally included but had to be rejected in the re-evaluation due to their high χ^2 values.

Beyerle et al. (1998) concluded from their results that in the LGM the temperature was about 5 °C lower than now. The re-evaluation gave a Holocene mean temperature of (7.9 ± 0.6) °C. The two glacial samples had an average of (4.8 ± 0.6) °C, yielding a tem-

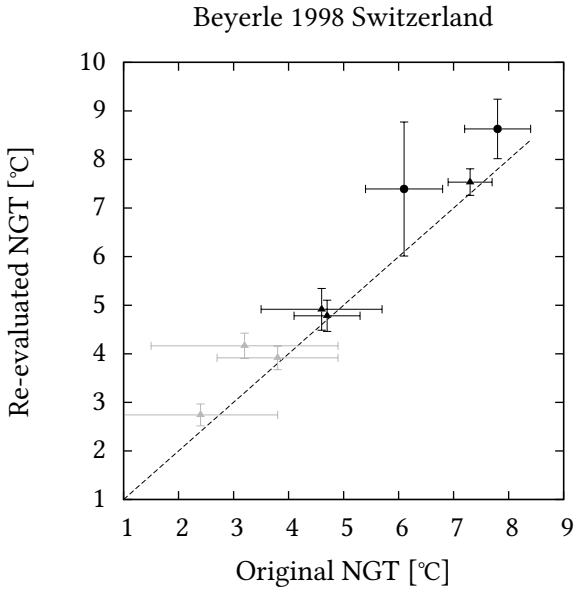


Figure 5.4.: Beyerle 1998 Switzerland: comparison of original and re-evaluated NGTs

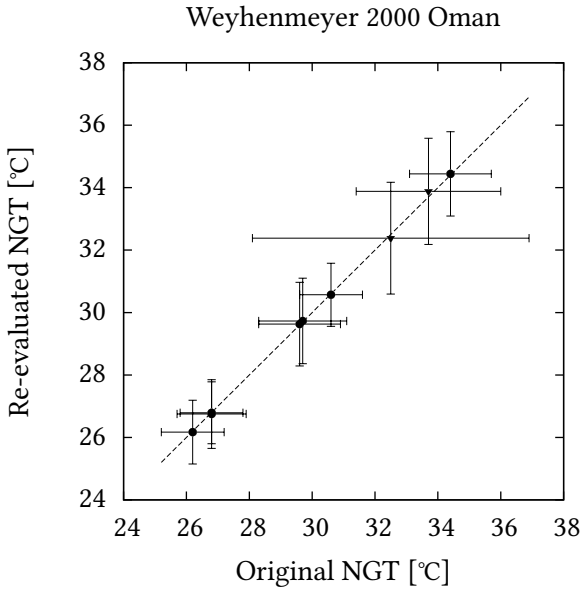


Figure 5.5.: Weyhenmeyer 2000 Oman: comparison of original and re-evaluated NGTs

perature difference of (3.0 ± 0.9) ℃. This result, however, remains questionable because the three (supposedly) coldest samples had to be rejected due to their bad fit and because of the small number of samples in general.

5.3. Weyhenmeyer 2000 Oman

This data set contains eleven samples from Oman, southeastern Arabia, and was published by Weyhenmeyer et al. (2000). The samples were taken from three aquifers in the Al Khwad Fan catchment. Nine of them were evaluated using the CE model with a pressure of 1 atm and a salinity of zero. The two remaining samples could not be described with any existing model. Fig. 5.5 shows that the results of the re-evaluation match the original ones very well. For the samples RGS-5F and RGS-2U, the errors could be reduced using Monte Carlo simulations. DP-2 and 21/6, the samples ruled out in the original evaluation, were excluded from the re-evaluation as well due to their high χ^2 values.

In the re-evaluation, the three Holocene samples showed a mean temperature of (33.7 ± 1.0) ℃. The Pleistocene samples had a mean of (26.6 ± 0.8) ℃. Combined, they give a temperature rise of (7.1 ± 1.3) ℃, which is slightly higher than the original value of (6.5 ± 0.6) ℃, but still agrees within the limits of uncertainty.

5.4. Aeschbach-Hertig 2002 Maryland

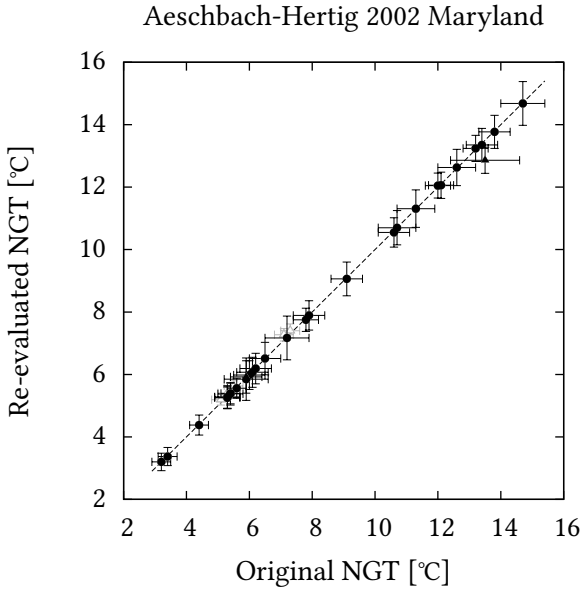


Figure 5.6.: Aeschbach-Hertig 2002 Maryland: comparison of original and re-evaluated NGTs

This data set recorded by Aeschbach-Hertig et al. (2002b) consists of 34 samples from the Aquia aquifer in southern Maryland, USA. The original evaluation was carried out using the CE model and inverse fitting. A recharge altitude of 50 m was used for the calculation of the pressure parameter; the salinity was assumed to be zero. As the sample MD6.2 yielded an unrealistically high A value (0.288 cm^3 STP/g), Aeschbach-Hertig et al. (2002b) chose to fix A in the center of realistic A values (i.e., $A = 0.1\,cm^3$ STP/g) and then selected the temperature error in such a way that it spans the results one gets when using the minimal or maximal A values. This changed the NGT estimate from (14.4 ± 2.6) ℃ to (13.5 ± 1.1) ℃. The same sample was evaluated in detail in sections 3.3 and 3.5, where it was used to analyze the behavior of the CE model. The sample's re-evaluated NGT was found to be (12.9 ± 0.4) ℃. Fig. 5.6 shows that the results of the re-evaluation are in very good agreement with the original ones, with the exception of the aforementioned sample MD6.2.

The re-evaluation gave a mean Holocene temperature of (13.6 ± 0.5) ℃ and mean temperature of (4.6 ± 0.4) ℃ for the LGM. This agrees very well with the original values of (13.7 ± 0.3) ℃ and (4.7 ± 0.5) ℃ respectively. The Holocene LGM temperature difference was found to be (9.0 ± 0.6) ℃, exactly matching the original result.

5.5. Beyerle 2003 Niger

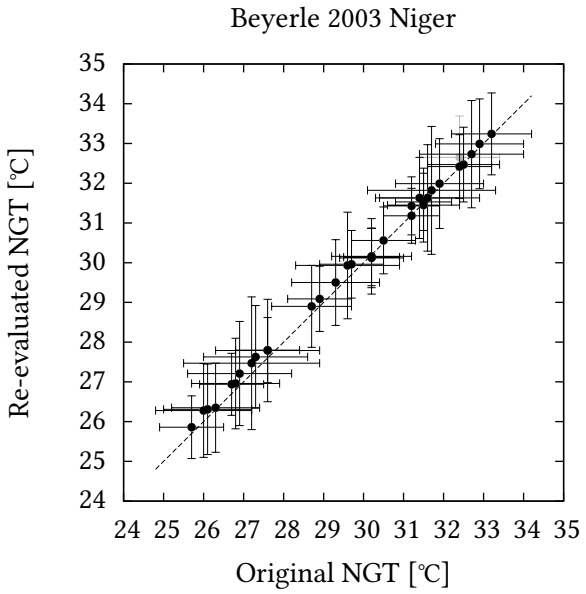


Figure 5.7.: Beyerle 2003 Niger: comparison of original and re-evaluated NGTs

This data set from southwestern Niger consists of 34 samples and was published by Beyerle et al. (2003). The samples were taken from the Continental Terminal aquifer system, which is divided into three layers. For the original evaluation, the CE model

was used. The pressure for the re-evaluation was calculated from the recharge elevation, which was estimated to be 250 m using the map included in the article in combination with *Google Earth*. The salinity was assumed to be zero. The samples were all well-behaved, no Monte Carlo simulations were necessary. Fig. 5.7 shows that the original results could be reproduced very well. Minor differences might have been caused by slightly different pressure assumptions.

In the article, different groups of samples were analyzed: the modern samples from the upper (CT3) aquifer, the samples from the recharge area of the middle aquifer (CT2), which were about 6 kyr old, the CT2 samples between 10 and 14 kyr BP (the coldest samples) and all confined samples from the middle and lower aquifers. The re-evaluation gave the following mean temperatures for these groups: (31.5 ± 0.3) ℃ for the modern CT3 samples, (30.1 ± 0.6) ℃ for the CT2 samples in the recharge area, (26.1 ± 0.7) ℃ for the cold group of CT2 samples between 10 and 14 kyr BP and (27.3 ± 0.4) ℃ for all confined CT2 and CT1 samples. The cooling between modern samples and the 6 kyr old ones was originally found to be about 2 ℃, the re-evaluation gave (1.4 ± 0.7) ℃, which agrees within the uncertainty. For the coldest samples, the temperature difference increases to (5.4 ± 0.8) ℃, which matches the original result of 5.5 ℃ very well. The temperature difference between the modern CT3 samples and the confined samples of the aquifers CT2 and CT1 was found to be (4.2 ± 0.5) ℃, which is lower than the approximately 5 ℃ given in the paper.

5.6. Kulongoski 2004 Kalahari

These are samples from an inland sandstone aquifer at an altitude of around 1 km in the Kalahari Desert in central Botswana. The data set was published by Kulongoski et al. (2004) and consists of 12 samples. For the original evaluation, different excess air models (CE, PD and PR) were mixed: for every sample, the model was chosen, which led to the best fit probability. This can, however, cause problems, because NGTs from different models may be shifted and thus cannot be compared easily. The original approach is therefore not recommended. As the paper does not include detailed information about the recharge altitude, it was estimated to be about 1000 m using *Google Earth*. The salinity was assumed to be zero. Fig. 5.8 shows that the CE model samples were reproduced very well. Both PR model samples, however, deviate significantly. All five PD model samples needed to be excluded from the re-evaluation due to too high χ^2. One of the five samples (769) was excluded from the original study as well. The original usage of the PD model for these five samples is a little surprising, as neither the UA model results (which all show positive A values) nor the Ne concentrations indicate degassing. Trying to fit the CE model in degassed mode (i.e., with an initial value of $F > 1$), either led back to the non-degassed case with $0 < F < 1$ or, if the fit did not manage to overcome the threshold at $F = 1$, gave extremely high values of χ^2.

The three remaining Holocene samples gave a mean NGT of (22.5 ± 0.8) ℃. The samples older than 24 kyr BP have an average temperature of (18.6 ± 0.6) ℃. Combined, this yields a temperature difference of (4.0 ± 1.0) ℃, which is more than 1 ℃ lower than the

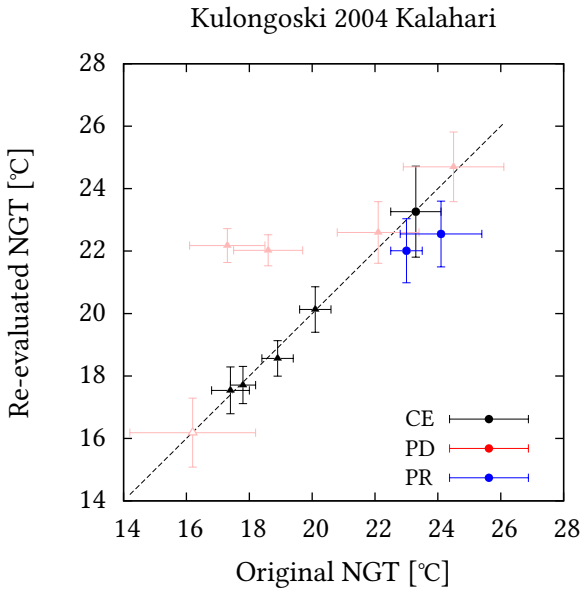


Figure 5.8.: Kulongoski 2004 Kalahari: comparison of original and re-evaluated NGTs. The colors indicate the excess air model used for the original evaluation.

original result of (5.2 ± 1.5) °C. Both results, however, still agree within their ranges of uncertainty.

5.7. Ma 2004 Michigan

This data set was published by Ma et al. (2004). The region of study is southern Michigan. The samples were taken from 13 wells in the Marshall aquifer and one well in the Saginaw aquifer. Noble gas temperatures were originally determined following Ballentine and Hall (1999), i.e., with the UA model and inverse fitting. The pressure for the re-evaluation was calculated from the presumed recharge altitude of 300 m, the salinity was set to zero. For eight out of 19 samples, the re-evaluated χ^2 values are too high to fulfill the criterion of $p \geqslant 0.01$. Therefore, they had to be omitted according to the conditions used for this re-evaluation. χ^2 for most of these samples is in the range of 11.8 to 16.0, corresponding to probabilities between 0.006 and 0.06 %. Samples 10a and 10b, however, cannot be described at all, as they show extremely high χ^2 values of 30 and 56 respectively. In the original evaluation, only the samples 10a and 10b had been excluded. The problem of high χ^2 values mainly affects the modern and LGM samples, leaving only one sample in each group. The remaining sample in the Holocene group is sample 8, which is more than 1 °C colder than samples 1, 6 and 7 (which have too high χ^2). The temperature difference between this sample and the remaining LGM sample is (3.2 ± 1.4) °C, which is significantly lower than the original difference of 5 °C. This result, however, remains questionable. In their article, Ma et al. (2004) discussed an

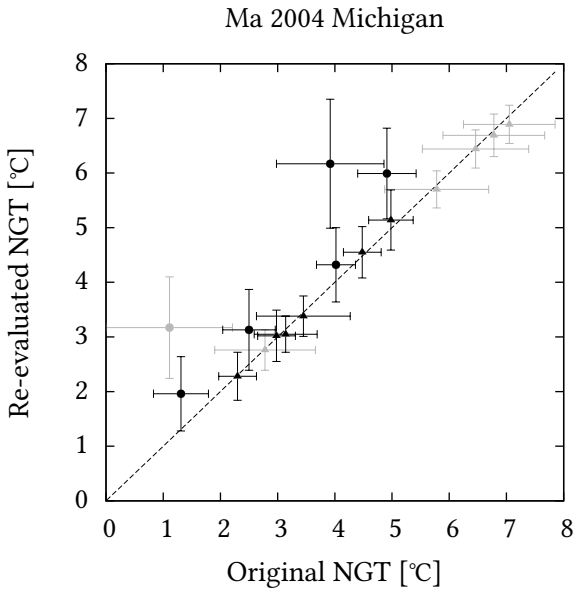


Figure 5.9.: Ma 2004 Michigan: comparison of original and re-evaluated NGTs

abrupt warming event at around 12 kyr BP with a temperature increase of about 2.4 ℃. The re-evaluation of the involved samples (12a,b and 15) yielded a significantly higher value of (3.8 ± 1.4) ℃. Ma et al. (2004) also mentioned a possible slight temperature increase of about 0.7 ℃ at around 6 kyr BP (samples 2a,b and 4a,b). This increase could not be confirmed, as the re-evaluation of this difference gave (0.1 ± 1.2) ℃.

Overall, many samples in this study cannot be described very well using existing models. The NGTs of modern samples cannot reproduce the MAAT; they are significantly lower. The OD model, which was developed by Hall et al. (2005) as a consequence of the problems in this data set, does not perform any better with respect to paleotemperatures. It even yields slightly lower temperatures, because the least χ^2 is found at $P_{OD} = 0.98$. If the possible range of P_{OD} values is constrained to the range from 1.00 to 1.26, the best match is found at $P_{OD} = 1$, reproducing the UA model results, which, in turn, have also lower NGTs than the CE model results. Whichever excess air model is used, the probability of the data set's χ^2 never exceeds $2.3 \cdot 10^{-14}$ (excluding samples 10a and 10b).

5.8. Kreuzer 2009 China

Kreuzer et al. (2009) recorded a set of 41 samples from the North China Plain. The CE model was used for the original evaluation. The recharge elevation was assumed to be 50 m for wells with an elevation of less than 50 m and assumed to be equal to the well elevation for all wells higher than 50 m. The salinity was set to zero. Fig. 5.10 shows that the re-evaluation was able to reproduce very well the original results. Most samples

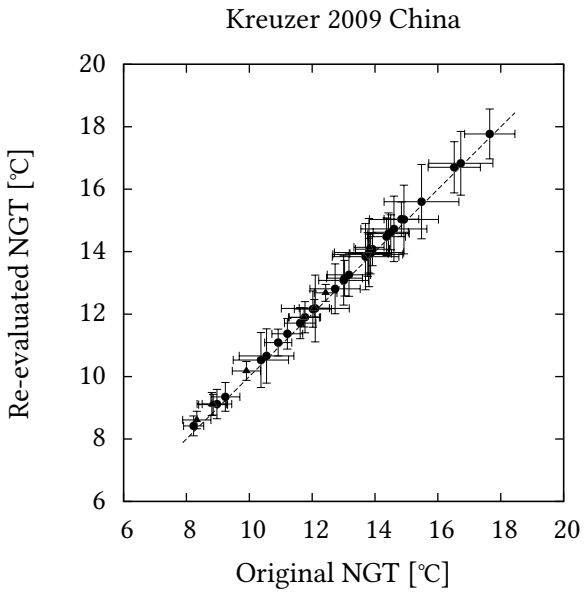


Figure 5.10.: Kreuzer 2009 China: comparison of original and re-evaluated NGTs

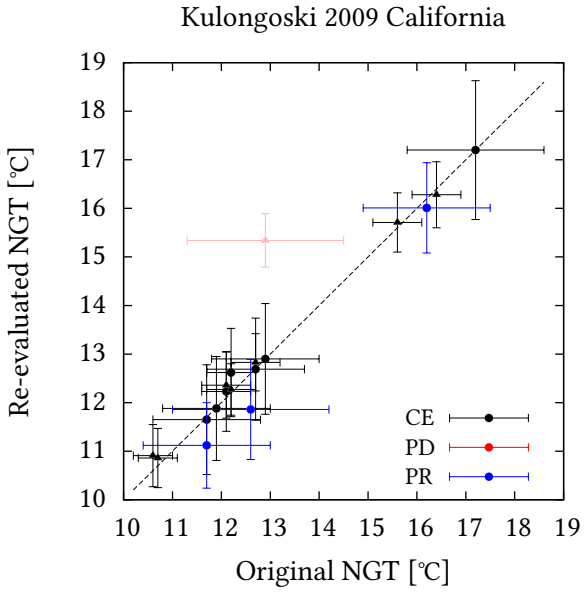


Figure 5.11.: Kulongoski 2009 California: comparison of original and re-evaluated NGTs. The colors indicate the excess air model used for the original evaluation.

were well-behaved and a Monte Carlo analysis was not necessary. For nine samples, the temperature uncertainties could be reduced using Monte Carlo calculations. For two samples, no reliable temperatures were determined, because of too high χ^2 values. As in the original study, the samples exhibiting exceptionally high NGTs, i.e., samples 15, 39, 45 and 46, were excluded from any temperature averaging.

For the modern samples, the re-evaluation yielded a temperature of (13.8 ± 0.3) ℃. All Holocene samples together gave a mean temperature of (13.4 ± 0.2) ℃. The Pleistocene samples had an average NGT of (9.7 ± 0.3) ℃ and the coldest six samples, representing the coldest sampled period, yielded (8.9 ± 0.4) ℃.

The re-evaluated Holocene Pleistocene NGT difference is (3.7 ± 0.3) ℃. This agrees very well with the original difference of (3.8 ± 1.5) ℃ but its estimated uncertainty is significantly lower, which is probably due to too conservative error estimates for the mean temperatures of the age groups in the original publication. A comparison of Holocene samples to only the coldest Pleistocene samples yields (4.4 ± 0.4) ℃, again agreeing well with the original value of (4.6 ± 1.2) ℃. As to the difference between the modern samples and the coldest Pleistocene ones, the re-evaluation gives (4.9 ± 0.4) ℃. The original result was (5.1 ± 1.1) ℃.

5.9. Kulongoski 2009 California

This data set was published by Kulongoski et al. (2009) and consists of 18 samples from the Mojave River Basin regional aquifer in the western Mojave Desert, California, USA. The samples were taken from ten different wells. They were evaluated using a recharge altitude of 1100 m and a salinity of 0.03 ‰. As in their earlier study, Kulongoski et al. (2009) evaluated the samples using different models. Unlike before, however, only CE model results were used for the evaluation of mean temperatures and temperature differences. If multiple CE-evaluated samples were available for a single well, only the one with the best fit probability was used. In the re-evaluation, all samples including replicate samples were used, except for sample 6b, which had a χ^2 slightly above the threshold.

For the 13 remaining groundwater samples older than 11.9 kyr, a mean temperature of (12.1 ± 0.4) ℃ was found. The four Holocene samples had an average NGT of (16.2 ± 0.6) ℃. Combined, this yields a Holocene Pleistocene temperature difference of (4.1 ± 0.8) ℃, which matches the originally found (4.2 ± 1.1) ℃ very well.

5.10. Blaser 2010 Belgium

This data set consists of 44 samples from the Ledo-Paniselian Aquifer in Belgium. It was published by Blaser et al. (2010) and exhibits several degassed samples as well as many of the special cases mentioned in section 4.3. The original evaluation was carried out using the CE model; the recharge elevation was assumed to be 20 m. The salinity was set to the values measured on-site for the different wells. Fig. 5.12 shows that for most samples, the re-evaluated results match the original ones quite well. For ten

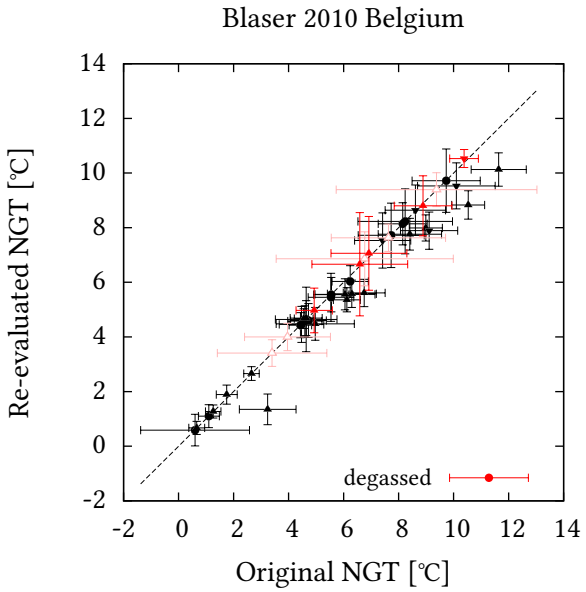


Figure 5.12.: Blaser 2010 Belgium: comparison of original and re-evaluated NGTs

samples, the re-evaluated noble gas temperature, after using Monte Carlo analysis, was lower than the original one by 0.5 °C or more. The ten degassed samples, which were also evaluated using Monte Carlo, show only small deviations of at most 0.2 °C. It is interesting to note that the two-dimensional A-T diagrams of these samples resemble those of samples with excess air, mirrored around some point on the T axis. They show the same patterns, e.g., split-ups into two clusters, or only a single cluster at too high values of A. Because these effects were not studied in detail for degassed samples, their evaluation was carried out using the more conservative approach of keeping all Monte Carlo realizations for the statistical analysis. As in the original study, five out of the ten degassed samples did not provide acceptable fits, i.e., their χ^2 test probability was less than 1 %. The re-evaluation also followed the original decision to exclude degassed samples for the calculation of mean temperatures or temperature differences within the record.

The re-evaluation gave the following results: the average temperature of the three modern non-degassed samples is (9.8 ± 0.7) °C. This is almost 1 °C lower than the original value of (10.7 ± 0.7) °C, because two of the three samples showed in Monte Carlo simulations the case of split-up clusters with fit results that are significantly shifted towards the unphysical cluster. For the coldest six samples, a mean temperature of (1.1 ± 0.5) °C was found, which compares well to the original (1.2 ± 0.2) °C. The difference between these two groups is (8.8 ± 0.8) °C, which is now a little smaller than the difference of (9.0 ± 0.6) °C in the Maryland study by Aeschbach-Hertig et al. (2002b). Originally, the NGT difference was found to be a little larger with a value of (9.5 ± 0.7) °C.

5.11. Comparison of excess air models

For the re-evaluation of all studies in this chapter, the CE model was used. In this section, this choice will be supported by a comparison of the CE, OD and PR models for all evaluated studies.

For the comparison, the three models were applied to all data sets without carrying out any Monte Carlo simulations. For every model and data set, a total χ^2 and a total number of degrees of freedom were determined either by adding up the numbers from the individual fits (CE and PR models) or by using the results of the ensemble fit (OD model). From these two values, the χ^2 probability defined in (2.43) was calculated as a measure of goodness of fit. The CE and PR model fits were carried out in the unconstrained mode, which normally does not give significantly different results from the constrained mode. For the OD model, both modes were analyzed, because small changes in P_{OD} may cause significant temperature shifts. Some of the studies contain several problematic samples with very high values of χ^2 for any model. These samples contribute very strongly to their studies' total χ^2 and lead to extremely low χ^2 probabilities. In order to limit their effect, the most problematic samples were excluded from this comparison. Samples with $\chi^2_{CE} > 20$ were regarded as belonging to this category. Degassed samples were also not considered because neither the OD nor the PR model is able to handle them.

Table 5.1 summarizes the results of this comparison. Judging from the χ^2 probabilities, the CE model gives the best results. In all but two studies it yields reasonably high values. For three studies, the PR model gives higher probabilities. Two of those—Kulongoski 2004 Kalahari and Ma 2004 Michigan—are, however, not described very well by any model. The OD model fails to give acceptable probabilities in all but two cases: in Beyerle 1998 Switzerland, where it performs a little better than the CE model but worse than the PR model, and in Kulongoski 2009 California.

The mean NGTs show that the CE and the PR model give similar temperatures in many cases, with the PR temperatures being a little higher most of the time. The OD model shows significantly higher temperatures for many studies, even up to about 10 °C, if the P_{OD} parameter is allowed to vary freely. In the unconstrained mode, P_{OD} leaves the realistic range (1.00 to 1.26) in half of the cases. In Ma 2004 Michigan, the one study where the CE and PR model NGTs of modern samples are much lower than the MAAT, the OD model also fails to provide temperatures in the expected range.

All in all, the CE model provides the best results in most cases. Except for special circumstances, it seems to be the best choice for the evaluation of noble gas data sets from groundwater.

Table 5.1.: Summary of the comparative analysis of the CE, OD and PR excess air models. The subscripts u and c refer to unconstrained and constrained fits, respectively.

	χ^2 probability [%]				Mean NGT [°C]			
Study	CE_u	OD_c	OD_u	PR_u	CE_u	OD_c	OD_u	PR_u
Stute 1995 Brazil	1.7	$5.0 \cdot 10^{-15}$	$1.4 \cdot 10^{-8}$	$5.0 \cdot 10^{-5}$	26.5	32.7	36.1	29.0
Beyerle 1998 Switzerland	2.7	3.3	3.6	8.7	6.6	6.2	5.5	6.5
Weyhenmeyer 2000 Oman	1.5	$8.6 \cdot 10^{-9}$	$2.8 \cdot 10^{-8}$	$1.1 \cdot 10^{-6}$	30.0	35.9	38.3	30.3
Aeschbach-Hertig 2002 Maryland	1.1	$8.2 \cdot 10^{-11}$	$8.2 \cdot 10^{-11}$	$3.1 \cdot 10^{-9}$	8.4	13.4	13.4	8.9
Beyerle 2003 Niger	23.8	$2.0 \cdot 10^{-31}$	$4.5 \cdot 10^{-21}$	$6.7 \cdot 10^{-3}$	29.8	35.7	40.1	31.6
Kulongoski 2004 Kalahari	$7.9 \cdot 10^{-4}$	$8.9 \cdot 10^{-10}$	$8.9 \cdot 10^{-10}$	$1.7 \cdot 10^{-3}$	20.4	24.6	24.6	21.2
Ma 2004 Michigan	$1.1 \cdot 10^{-9}$	$1.1 \cdot 10^{-12}$	$1.3 \cdot 10^{-12}$	$1.1 \cdot 10^{-4}$	4.1	3.9	3.4	5.0
Kreuzer 2009 China	5.3	$1.4 \cdot 10^{-27}$	$1.4 \cdot 10^{-27}$	$1.5 \cdot 10^{-4}$	12.8	14.4	14.4	13.0
Kulongoski 2009 California	98.8	8.7	8.7	94.7	12.9	15.5	15.5	13.2
Blaser 2010 Belgium	55.9	$9.4 \cdot 10^{-9}$	$9.4 \cdot 10^{-9}$	0.060	5.8	9.8	9.8	6.1

6. The software PANGA

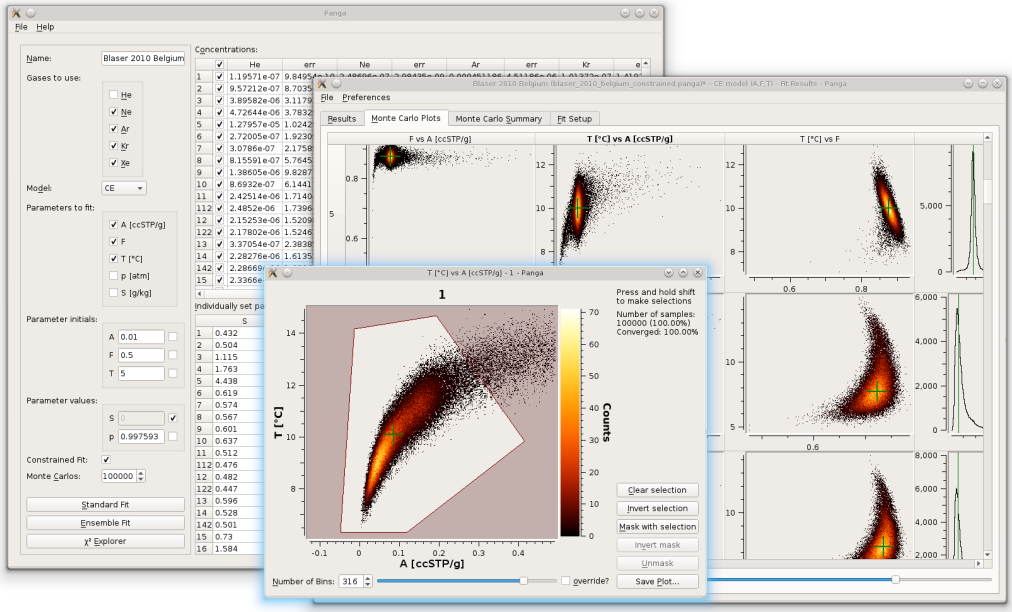


Figure 6.1.: The software *PANGA*.

With existing software solutions it was tedious, if not impossible, to carry out some of the evaluation methods detailed in chapters 3 and 4. The software *PANGA*[1] was developed in the context of this work with the goal of providing an easy-to-use tool for the application of those advanced evaluation methods.

6.1. Features

In its most basic mode of operation, *PANGA* finds a set of sample parameters in such a way that the modeled noble gas concentrations reproduce the measured sample concentrations in the best possible way (i.e., with minimal χ^2, cf. chapter 2.2). This is done for all samples in the data set independently of each other. If desired, the range of possible values of the parameters may be constrained to only include physically plausible values. Model parameters which are not supposed to be varied can be set to fixed values either

[1]Program for the Analysis of Noble GAs data

globally or independently for each sample. In the same manner, initial parameter values for the fitting algorithm may be specified either globally or individually. The user can also exclude selected noble gases from the fit globally.

Alternatively, the software can also perform a so-called "ensemble fit", i.e., it does not treat the samples individually but combines all of them in a single fit. In this mode each fit parameter can either be varied independently for each sample or it can be fitted to the ensemble of samples, i.e., it has the same value for every sample.

In both modes the user can choose to additionally carry out a specific number of Monte Carlo calculations which can be used to determine more accurate estimates for the fit parameters and their errors and to check whether or not the fit shows any kind of abnormal behavior. A few 100 000 up to several millions of Monte Carlo realizations can be fitted in a few minutes, depending on computation power and on the properties of the samples. All the values calculated for normal fits are also saved for every Monte Carlo realization and one-dimensional histograms of their frequencies of occurrence can be created for all of them. Additionally, any two parameters may be combined into a two-dimensional histogram. Both kinds of histograms can be used to select a subset of the Monte Carlo realizations which can then be applied as a mask to the Monte Carlo results of the sample's other parameters. This can be used, for example, to restrict the analysis of the sample to one of multiple clusters as suggested in chapters 3 and 4.3 for certain problematic cases. The program performs a basic statistical analysis, consisting of mean value and standard deviation, for the remaining Monte Carlo realizations and summarizes these statistical characteristics in a table.

The output of the fitting procedure consists of the following quantities for every single fit: degrees of freedom of the fit; the final χ^2 value together with the probability of obtaining this or a higher value based on the χ^2 distribution (cf. equation 2.43); the best estimates of the fit parameters together with their uncertainties derived from the covariance matrix (cf. equation 2.41); the off-diagonal elements of the covariance matrix in the form of correlation coefficients, which need to be used for error propagation if further quantities are derived from the fit parameters; the residuals of the fit; the equilibrium components of the modeled concentrations; the final modeled concentrations.

Because χ^2 may have multiple local minima it can sometimes be helpful to analyze the structure of the χ^2 surface in the parameter space in order to find irregularities or to verify that the obtained minimum also is the global one. For these cases *PANGA* provides the χ^2 explorer mode, which shows contour plots of two-dimensional cross sections of the χ^2 surface. The user can freely choose the parameters on the plot's axes and can either fix the remaining parameters to a certain value or leave them to be fitted. This feature allows for an interactive exploration of the χ^2 surface, as the results are calculated on-the-fly while the user pans or zooms the plot.

6.2. Implementation

PANGA is written completely in C++ and runs on Microsoft Windows, Linux and OS X. For its calculations and user interface it makes use of several libraries: the *Eigen*

template library (`http://eigen.tuxfamily.org`) is used for all algebraic computations. *Eigen* comes with a port of *MINPACK*, which is a reliable and robust implementation of the Levenberg-Marquardt algorithm. The user interface was implemented using the *Qt* framework (`http://qt-project.org`); plots are created with the *Qwt* library (`http://qwt.sourceforge.net`). The *boost* libraries (`http://www.boost.org`) are used for many other aspects like multi-threading, serialization of data and compression.

6.3. PANGA output

The output of *PANGA*, which is calculated for normal fits as well as for each Monte Carlo realization, consists of the following quantities:

Degrees of freedom The degrees of freedom of a fit, i.e., the number of measurements used for the fit minus the number of fit parameters.

Chi Square The χ^2 of the fit, as defined in (2.23).

Probability The probability of measuring this or a higher χ^2 in repeated experiments, as defined in (2.43).

Fit parameter estimates The best estimates for the fitted model parameters, found by the Levenberg-Marquardt algorithm as described in chapter 2.2.3.

Fit parameter errors Approximations of the errors of the model parameter estimates, as found by equation (2.41).

Convergence Information about the convergence of the fit. This may be one out of:

- **Converged** The fit converged according to one of the convergence criteria.
- **Too many function calls** Fitting was stopped, because 10 000 iterations were reached without the fit being converged.
- **Improper Input Parameters** No fit was be carried out because of a problem with the input. This normally means that there were more fit parameters than data points to fit to.
- **Error** An unknown error occured.

Correlation coefficients The correlation coefficients of the fit, i.e., the off-diagonal elements of the correlation matrix (2.18). Like the fit parameter errors they are calculated from (2.41).

Residuals Residuals of the fit. They are given as relative deviations of the observed noble gas concentrations from the modeled ones: $(c_i - c_i^{\mathrm{mod}})/c_i^{\mathrm{mod}}$.

Equilibrium concentrations The equilibrium components of the modeled concentrations. Their errors are calculated from those of the fit parameters using the general law of error propagation (2.46).

Modeled concentrations Sum of the equilibrium and excess air components, which is equal to the total modeled concentrations. Their errors are calculated from those of the fit parameters using the general law of error propagation (2.46).

Measured concentrations These are the input concentrations of the samples. It also includes the concentrations which were not fitted. In the case of Monte Carlo simulations, these are the randomly altered input concentrations.

6.4. Errors of derived quantities

In order to estimate the errors of quantities derived from the *PANGA* output, equation (2.46) needs to be applied. If expressed in terms of the correlation matrix $\boldsymbol{P}$, it reads

$$\boldsymbol{V}_y = \boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{P}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}} \tag{6.1}$$

with $\boldsymbol{\Sigma}$ being the matrix with the errors of the fit parameters on its diagonal and 0 otherwise. The correlation matrix can be put together from the correlation coefficients *PANGA* gives out (its diagonal elements are 1).

6.5. Comparison with existing results

In order to verify the accuracy of the fits produced by *PANGA*, its output was compared to the results generated by *NOBLE* by Peeters et al. (2003) using two published noble gas data sets. The first one is the Belgium data set by Blaser et al. (2010) which was reviewed in chapter 5.10. The second dataset was taken from the study of Castro et al. (2007) and consists of 41 samples from the Carrizo Aquifer in Texas, USA. For the comparison, both *NOBLE* and *PANGA* were run without constraints on the range of parameter values.

6.5.1. Ledo-Paniselian Aquifer, Belgium

All 44 samples from 39 wells reported in the study of Blaser et al. (2010) were fitted with the CE model; 34 had excess air whereas degassing occurred in 10 of them. Here, the samples are labeled with BE followed by the well number or name as used in the original publication. The fitting was carried out with the following initial values for the parameters: T was set to 2 ℃, A to 0.01 cm^3 STP/g. F has to be chosen according to the type of sample to prevent the fitter from falling into a wrong local minimum. For the samples with excess air and with degassing the initial values 0.5 and 3 were used, respectively.

Except for five special cases, which will be discussed later on, most results were very close to those provided by *NOBLE*, i.e., the best parameter estimates showed typical deviations of 0.002 to 0.017 ℃ for T, $0.03 \cdot 10^{-3}$ to $1.50 \cdot 10^{-3}$ cm^3 STP/g for A and $0.02 \cdot 10^{-3}$ to $2.50 \cdot 10^{-3}$ for F. The samples BE_7 and BE_9, however, showed significant deviations of the F parameter of almost 0.1. For sample BE_29, the F deviation is 0.02. These three samples have in common that their F and A parameters are ill-determined: their

errors are between 5 and almost 150 times as high as their values. The deviations of estimated parameter errors are typically of the same order of magnitude as the deviations of their parameter values. If the parameters are, however, not very well determined, i.e., their uncertainties become about the same size as their values, the deviations of the estimated errors can increase drastically. An overview of the fit results and their deviations is given in Table D.1.

A detailed comparison of the results led to the conclusion that discrepancies in the modeled Xe equilibrium concentrations of about 0.01% caused these deviations. Further investigation showed that these discrepancies were due to *NOBLE* using an additional factor for the conversion of molar units derived from Clever (1979b). Instead of the simple factor X_{Xe}, as recommended by Kipfer et al. (2002), *NOBLE* uses $X_{Xe}/(1 - X_{Xe})$, with X_{Xe} being the mole fraction solubility of Xe, which is small compared to 1. For a better comparison, the additional $1/(1 - X_{Xe})$ factor was temporarily incorporated into *PANGA* and the analysis was repeated. Now all the samples, except for the five aforementioned special cases, were in very good agreement with *NOBLE*, i.e., the deviations of the parameters were, in the majority of the cases, smaller than $5 \cdot 10^{-4}$ °C, 10^{-5} cm^3 STP/g and 10^{-5} for the parameters T, A and F respectively. Even in the more problematic cases with high parameter uncertainties, the deviations of the parameter estimates did not exceed 0.005 °C, $4 \cdot 10^{-4}$ cm^3 STP/g and $7 \cdot 10^{-4}$. In the extreme cases with ill-determined parameters like the ones mentioned above, the absolute deviations of the parameter errors were still higher, but they were always lower than 0.7 % of the estimated errors. The fit results of the modified *PANGA* version and their deviations from the *NOBLE* results are listed in Table D.2.

The five remaining samples are in the UA limit case described in section 4.4.3. Their fit results are listed in Table D.3. The samples exhibit special values of the F parameter, i.e., they are highly negative (ranging from −1200 to −2200) or, in one case, highly positive (about 4200). The corresponding uncertainties of F are very large and range from 10^7 to 10^8. The A values are quite low and range from $3 \cdot 10^{-7}$ to 10^{-6} cm^3 STP/g with uncertainties of $6 \cdot 10^{-3}$ up to $3 \cdot 10^{-2}$ cm^3 STP/g. A comparison of these samples with *NOBLE* shows good agreement for the T parameter, with differences of $1.6 \cdot 10^{-3}$ °C at most, but significant discrepancies otherwise: The results of *NOBLE* for the absolute values of F are much lower, roughly by an order of magnitude, whereas the values for A are higher, again by about one order of magnitude. The temperature uncertainties also deviate from each other by up to 47 %. The product $A \cdot F$, however, is the same for both fitters, with differences of less than 0.8 %. These results can be understood if one considers the limit case of the CE model given in equation (4.1). It shows that in the case of $|F| \gg 1$ the modeled concentrations only depend on the product $A \cdot F$. Therefore, one can expect that A and F individually can only be determined with a high degree of uncertainty and that small numeric differences in the calculations can have a large impact on the estimates of these two parameters. The T parameter, however, is not influenced by this effect and, for this reason, the temperature estimates do not show any significant deviations. The differences in T uncertainties may be understood if we consider equation (2.41): the covariance matrix $\boldsymbol{V}_p$, which is used to determine the temperature error depends on the Jacobian $\boldsymbol{J}$. $\boldsymbol{J}$ is evaluated at the estimates $\boldsymbol{p}^*$, i.e., at the

fit parameters T, A and F. As the two fitters find different values of A and F, it is thus to be expected that the temperature error is also different.

It is also interesting to consider the UA limit case given in equation 4.2, which holds if, besides $|F| \gg 1$, the product $|FA|$ is small compared to 0.01 cm^3 STP/g. For the samples just discussed, $|FA|$ lies between $4 \cdot 10^{-4}$ cm^3 STP/g and 0.002 cm^3 STP/g, i.e., the latter approximation is not fully justified in all cases. Nevertheless, equation 4.2 shows that large negative values of F in combination with small positive values of A can lead to a reasonable description of the observed noble gas concentrations which approximates the simple case of unfractionated excess air. The one case with a large positive F and small positive A actually corresponds to a degassed sample that approximates a UA model with a hypothetical negative excess air parameter.

6.5.2. Carrizo Aquifer in Texas, USA

The study of Castro et al. (2007) consists of 49 samples from 20 different wells, labeled by TX followed by a well number and a number for the replicate analyses. Following the treatment in the original paper, eight of the 49 samples were removed from the analysis because of their χ^2_{UA} values being too high. The remaining 41 samples were fitted with the OD model in an ensemble fit, i.e., the parameters A and T were fitted independently for each sample but only a single P_{OD} value was determined for the whole dataset. The value obtained in this way is $P_{OD} = 1.114 \pm 0.020$ at $\chi^2_{OD} = 73.9$. Castro et al. (2007), however, calculated a minimal χ^2_{OD} value of 113.0 at $P_{OD} = 1.14$. A comparison of UA model results also showed significant deviations: here, a total χ^2_{UA} of 109.2 was determined whereas Castro et al. (2007) found $\chi^2_{UA} = 201.6$. Furthermore, the temperature estimates calculated with *PANGA* were systematically lower, roughly by 0.1 ℃. The reasons for these differences remained somewhat unclear but are likely due to the use of different evaluation algorithms incorporating for example different solubility equations on the one hand and the use of slightly different input data on the other hand. Differences in the input may include the assumed fixed values for pressure and salinity, which we were unable to reproduce from the information given in the article, and the uncertainties of the measured noble gas concentrations, which were only specified as general percentages in the paper but might have been available in more detail for the original analysis.

In order to test the reliability of *PANGA's* fitting algorithm, a comparison was carried out with *NOBLE* also for the Carizzo data set, again using the aforementioned additional factor $1/(1 - X_{Xe})$ so as to match *NOBLE's* solubility equilibrium concentrations in the best possible way. The UA model results are in very good agreement with each other: none of the samples (including the eight samples with high χ^2_{UA}) showed relative deviations in the model parameters or their uncertainties which exceeded 10 ppm. Because *NOBLE* cannot natively fit the OD model, the best estimate of P_{OD} was determined by running multiple UA fits with pressure $P = P^* \cdot P_{OD}$ for different values of P_{OD} (P^* is the pressure estimated from the recharge altitude). In this way a minimal χ^2_{OD} of 73.8 was found at $P_{OD} = 1.112$. This compares well with the value obtained by the ensemble fit of *PANGA*, which—with the Xe modifications in place—slightly increased to

$P_{OD} = 1.115 \pm 0.020$ at $\chi^2_{OD} = 73.9$. The deviation between the two values corresponds to 13% of the estimated error.

7. Summary

The analyses in this thesis confirm that the CE model sometimes gives unrealistic solutions when fitting noble gas data from groundwater. These solutions are characterized by unrealistically high values of A (above 0.1 cm^3 STP/g) and are accompanied by comparatively high values of T, with often very high errors obtained from the fitting algorithm. These large uncertainties are related to a high parameter correlation in this range of A values, but not to the occurrence of multiple minima. Two local minima in the χ^2 surface do occur for normal samples, but they are well separated and usually do not cause problems. Monte Carlo analyses are very useful to study the behavior of CE model fits and can provide more reliable estimates of the parameter uncertainties. They show that the problematic behavior is usually restricted to a subset of the Monte Carlo realizations, both in the cases of physical and synthetic samples.

It was shown that the problematic fitting behavior is related to and can be caused by deviating noble gas concentrations, particularly by increased Ar in combination with decreased Xe. The synthetic samples show that deviations less than typical measurement uncertainties are already sufficient to create these problems. Samples from warmer regions or from regions where infiltration occurred with a high amount of entrapped air seem to be especially sensitive to these kinds of deviations. Both synthetic and real samples show two clusters in Monte Carlo analyses, one of them corresponding to unrealistically high values of A and overestimated T, the other one apparently closely reflecting the true parameter values.

A new method of data evaluation for groundwater noble gas samples was proposed, which extends the usual CE model evaluation by a Monte Carlo analysis and then restricts the statistical evaluation of the Monte Carlo results to the cluster with realistic, low values of A and T. This method proved to be able to recover the original parameter values from modified synthetic samples. Its application to poorly-fitting physical samples yielded realistic temperatures. It thus seems that the proposed restriction of the Monte Carlo ensemble is permissible and that many problematic samples can be evaluated using this approach.

A step-by-step procedure for the evaluation of noble gas data was given, which starts with simple evaluation methods and uses advanced features like Monte Carlo simulations as needed. An initial UA model fit is followed by a CE model fit. If the CE model fit results exhibit unreasonable values like high values of the excess air parameter A, small values of the fractionation parameter F or large temperature uncertainties, Monte Carlo simulations should be carried out. The Monte Carlo results can be divided in different classes which need specific treatment.

A new, easy-to-use tool—*PANGA*—was developed for the analysis of noble gas samples using the before-mentioned step-by-step approach. It allows for a fast evaluation

workflow for well-behaved samples as well as problematic cases which need special treatment like statistical analysis over a subset of the Monte Carlo results. With its support for ensemble fits and more recent excess air models like the OD model, it can also be used for most special cases that might occur. The correct functioning of *PANGA's* fitting routines was demonstrated by a comparison of its results with the software *NOBLE*, which is traditionally used by many groups. The comparison only showed some deviations in special, unphysical cases of the CE model that are of little practical relevance.

The refined evaluation methods were applied in a review of important groundwater noble gas data sets from the literature. The temperature differences between Holocene and late Pleistocene were re-evaluated with a consistent approach. For the majority of studies, the original results could be verified with minor deviations. The results of some of the studies, however, could not be confirmed. A comparative analysis of the CE, OD and PR models showed that the CE model, though not without problems, can describe most data sets well, whereas the OD and PR models can be used successfully only in a minority of cases.

A. Properties of noble gases in water

The following chapter summarizes the recommended equations for the calculation of equilibrium concentrations of noble gases in water. The recommendations follow Kipfer et al. (2002). Temperatures in K will be denoted by T_K, whereas temperatures in ℃ will be denoted by $T_{℃}$. These notations will only be used where the temperatures have to be plugged in directly as numbers. In all other cases they will just be denoted by T. The equilibrium concentrations are in units of cm^3 STP/g. The following equations are also used by *PANGA*.

A.1. Solubility of He, Ne, Ar and Kr

The equilibrium concentrations of He, Ne, Ar and Kr can be calculated using the Weiss solubilities (Weiss 1970, 1971; Weiss and Kyser 1978). According to Kipfer et al. (2002) they are given by

$$c_i^{\mathrm{eq}} = \exp\begin{pmatrix} t_1^i + t_2^i \cdot \frac{100}{T_\mathrm{K}} + t_3^i \cdot \ln\left(\frac{T_\mathrm{K}}{100}\right) + t_4^i \cdot \frac{T_\mathrm{K}}{100} + \\ S \cdot \left[s_1^i + s_2^i \cdot \frac{T_\mathrm{K}}{100} + s_3^i \cdot \left(\frac{T_\mathrm{K}}{100}\right)^2\right] \end{pmatrix} \cdot \frac{P - e_w(T)}{(1 - e_w(T)) \cdot 1000}. \tag{A.1}$$

with the pressure P in atm. The coefficients t_j^i and s_j^i depend on the noble gas and are listed in table A.1. e_w is the water vapor pressure in atm, which is given in equation (A.8).

Table A.1.: Coefficients for the Weiss solubilities for the different noble gases. From Kipfer et al. (2002).

	He	Ne	Ar	Kr
t_1	−167.2178	−170.6018	−178.1725	−112.684
t_2	216.3442	225.1946	251.8139	153.5817
t_3	139.2032	140.8863	145.2337	74.469
t_4	−22.6202	−22.629	−22.2046	−10.0189
s_1	−0.044 781	−0.127 113	−0.038 729	−0.011 213
s_2	0.023 541	0.079 277	0.017 171	−0.001 844
s_3	−0.003 426 6	−0.012 909 5	−0.002 128 1	0.001 120 1

The ^{3}He component of the total He concentration can be determined using an empirical equation for the ^{3}He/^{4}He ratio R_{eq} (Benson and Krause 1980; Kipfer et al. 2002):

$$R_{eq} = R_a / \exp\left[\left(r_1 + \frac{r_2}{T_\mathrm{K}} + \frac{r_3}{T_\mathrm{K}^2}\right) \cdot (1 + r_4 \cdot S)\right] \tag{A.2}$$

$R_a = 1.384 \cdot 10^{-6}$ is the atmospheric ^{3}He/^{4}He ratio (Clarke et al. 1976).

Table A.2.: Coefficients for the calculation of R_{eq}. From Kipfer et al. (2002).

	r_1	r_2	r_3	r_4
R_{eq}	−0.029 964 5	19.8715	−1833.92	0.000 464

A.2. Solubility of Xe

The solubility of Xe was determined by Clever (1979b) in terms of the dimensionless mole fraction solubility:

$$X_{\mathrm{Xe}} = \exp\left[x_1 + x_2 \cdot \frac{100}{T_{\mathrm{K}}} + x_3 \cdot \ln\left(\frac{T_{\mathrm{K}}}{100}\right)\right] \tag{A.3}$$

Smith and Kennedy (1983) determined the salting coefficient of Xe

$$K_{\mathrm{Xe}} = x_4 + x_5 \cdot \frac{100}{T_{\mathrm{K}}} + x_6 \cdot \ln\left(\frac{T_{\mathrm{K}}}{100}\right). \tag{A.4}$$

in units of l/mol. The coefficients x_i are given in table A.3.

Combining these two properties, Aeschbach-Hertig et al. (1999) proposed an equation for the calculation of equilibrium concentrations (here in the representation of Kipfer et al. 2002):

$$c_{\mathrm{Xe}}^{\mathrm{eq}} = \frac{X_{\mathrm{Xe}}(T)}{M_{\mathrm{H_2O}}} \cdot \frac{(P - e_w(T)) \cdot z_{\mathrm{Xe}}}{P_0} \cdot \frac{\rho(T, S=0)}{\rho(T, S)} \cdot V_{\mathrm{Xe}} \cdot \exp\left(-K_{\mathrm{Xe}}(T) \cdot c_{\mathrm{NaCl}}\right) \tag{A.5}$$

with the pressure P, the reference pressure $P_0 = 1$ atm, the molar mass of water $M_{\mathrm{H_2O}} \approx 18.016$ g/mol, the molar volume of Xe $V_{\mathrm{Xe}} = 22\,280.4\,\mathrm{cm}^3$ STP/g, the density of water $\rho(T, S)$, as given in (A.10), and the molar concentration of NaCl c_{NaCl}. As, in general, we are not dealing with pure NaCl solutions, the salinity S needs to be converted into an equivalent NaCl concentration. According to Aeschbach-Hertig et al. (1999) this may be achieved using

$$c_{\mathrm{NaCl}} = S \cdot \frac{\rho(T, S, P)}{M_{\mathrm{NaCl}}} \tag{A.6}$$

with the molar mass of NaCl M_{NaCl}.

Table A.3.: Coefficients for the Clever solubility of Xe. From Kipfer et al. (2002).

	x_1	x_2	x_3	x_4	x_5	x_6
Xe	−74.7398	105.21	27.4664	−14.1338	21.8772	6.5527

A.3. Diffusion coefficient of noble gases in water

The diffusion coefficients of the noble gases He, Ne, Kr and Xe in water were measured by Jähne et al. (1987) for different temperatures. They described the temperature depen-

dency with the equation

$$D_i = A_i \cdot \exp\left[-\frac{E_i}{RT}\right]. \tag{A.7}$$

with gas-dependent fit parameters E_i and A_i, where E_i is the activation energy of gas i for diffusion in water. R is the universal gas constant. The fit parameters are given in Table A.4. As no Ar diffusion coefficients were determined by Jähne et al. (1987), these coefficients were interpolated by Peeters et al. (2003) for a number of temperatures assuming that D is inversely proportional to the square root of the atomic mass. Fitting (A.7) to these diffusion coefficients yields the values for Ar in Table A.4.

Table A.4.: Fit parameters for the calculation of the diffusion coefficients. Jähne et al. (1987) determined the values for He, Ne, Kr and Xe by fitting to measured diffusion coefficients. The Ar values were calculated from the interpolated diffusion coefficients in Peeters et al. (2003).

Gas	A [10^{-5} cm^2/s]	E [kJ/mol]
He	818	11.70
Ne	1608	14.84
Kr	6393	20.20
Xe	9007	21.61
Ar	2233	16.68

A.4. Properties of water

According to Gill (1982), the saturation vapor pressure $e_w(T)$ of pure water over a plane water surface is given by

$$log_{10}(e_w(T)) = \frac{0.7859 + 0.03477 \cdot T_{°C}}{1 + 0.00412 \cdot T_{°C}} \tag{A.8}$$

Gill (1982) also gives equations for the calculation of the density of water. Note, that the pressure in the following equations is expected to be given in bar above the reference pressure, i.e., $P = 0$ is equivalent to a pressure of 1 atm.

For pure water, the density can be described by

$$\begin{aligned}\rho(T, 0, 0) = 999.842594 + 6.793952 \cdot 10^{-2} T_{°C} - 9.095290 \cdot 10^{-3} T_{°C}^2 + \\ 1.001685 \cdot 10^{-4} T_{°C}^3 - 1.120083 \cdot 10^{-6} T_{°C}^4 + 6.536332 \cdot 10^{-9} T_{°C}^5.\end{aligned} \tag{A.9}$$

The density of water with non-zero salinity at one standard atmosphere is given by

$$\begin{aligned}\rho(T,S,0) = \rho(T,0,0) + \\ S\cdot(0.824493 - 4.0899\cdot10^{-3}T_{°C} + 7.6438\cdot10^{-5}T_{°C}^2 \\ \quad - 8.2467\cdot10^{-7}T_{°C}^3 + 5.3875\cdot10^{-8}T_{°C}^4) + \\ S^{3/2}\cdot(-5.72466\cdot10^{-3} + 1.0227\cdot10^{-4}T_{°C} - 1.6546\cdot10^{-6}T_{°C}^2) + \\ S^2\cdot4.8314\cdot10^{-4}.\end{aligned} \tag{A.10}$$

In the general case at pressure P, the density can be calculated according to

$$\rho(T,S,P) = \frac{\rho(T,S,0)}{1-\frac{P}{K(T,S,P)}}. \tag{A.11}$$

with the secant bulk modulus K. For pure water, K is given by

$$\begin{aligned}K(T,0,0) = 19652.21 + 148.4206T_{°C} - 2.327105T_{°C}^2 + \\ 1.360477\cdot10^{-2}T_{°C}^3 - 5.155288\cdot10^{-5}T_{°C}^4.\end{aligned} \tag{A.12}$$

At one standard atmosphere with non-zero salinity, it becomes

$$\begin{aligned}K(T,S,0) = K(T,0,0) + \\ S\cdot(54.6746 - 0.603459T_{°C} + 1.09987\cdot10^{-2}T_{°C}^2 - 6.1670\cdot10^{-5}T_{°C}^3) + \\ S^{3/2}\cdot(7.944\cdot10^{-2} + 1.6483\cdot10^{-2}T_{°C} - 5.3009\cdot10^{-4}T_{°C}^2).\end{aligned} \tag{A.13}$$

At pressure P with non-zero salinity, K is given by

$$\begin{aligned}K(T,S,P) = K(T,S,0) + \\ P\cdot(3.239908 + 1.43713\cdot10^{-3}T_{°C} + 1.16092\cdot10^{-5}T_{°C}^2 - 5.77905\cdot10^{-7}T_{°C}^3) + \\ PS\cdot(2.2838\cdot10^{-3} - 1.0981\cdot10^{-5}T_{°C} - 1.6078\cdot10^{-6}T_{°C}^2) + \\ PS^{3/2}\cdot1.91075\cdot10^{-4} + \\ P^2\cdot(8.50935\cdot10^{-5} - 6.12293\cdot10^{-6}T_{°C} + 5.2787\cdot10^{-8}T_{°C}^2) + \\ P^2S\cdot(-9.9348\cdot10^{-7} + 2.0816\cdot10^{-8}T_{°C} + 9.1697\cdot10^{-10}T_{°C}^2).\end{aligned} \tag{A.14}$$

B. PANGA manual

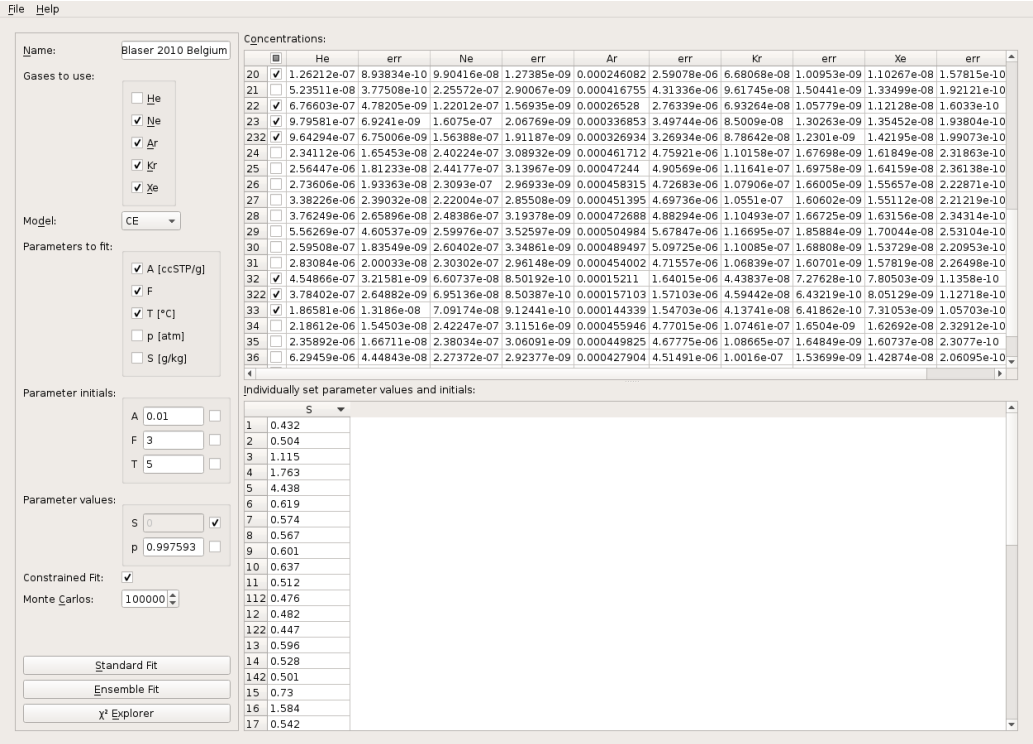


Figure B.1.: *PANGA* main window

This chapter contains the user manual for *PANGA*. A general overview of the software may be found in chapter 6.

B.1. Loading data

Noble gas concentrations can be loaded from a file or imported from the clipboard using the respective entries in the file menu or the shortcuts `Ctrl+L` / `Ctrl+V`. The file or clipboard contents must meet the following specifications:

- The data need to be arranged in eleven columns: sample name, He, He error, Ne, Ne error, Ar, Ar error, Kr, Kr error, Xe and Xe error.
- Noble gas concentrations and there errors need to be given in cm^3 STP/g.

- The columns need to be separated by either comma or tab.
- The decimal mark must be a dot. Other decimal mark system settings will be ignored.
- If no concentrations are available for some gases, their fields may also be left empty.

Lines not fulfilling the above conditions will be ignored. Just using copy and paste from a spreadsheet program should generally work.

B.2. Setting up the fit

On the left side of the main window, the fit can be configured:

Gases to use Here, the individual noble gases can be added to or removed from the fit. If a noble gas is selected here, but no concentrations are given, the gas will not be fitted for that sample.

Model The excess air model to fit to the data.

Parameters to fit The model parameters to be fitted need to be selected here.

Parameter initials Here, the initial parameter guess for the fitter can be configured. To set different initial values for each sample, check the checkbox on the right side of the parameter.

Parameter values Fixed values for the model parameters not to be fitted. To set different parameter values for each sample, check the checkbox on the right side of the parameter.

Constrained Fit If this checkbox is checked, the range of possible values for the fit parameters will be constrained. The ranges can be configured by clicking the *Setup* button.

Monte Carlos The number of Monte Carlo fits to be carried out for each sample.

The table on the bottom can be used to assign each sample individual values for certain model parameters or individual initial guesses for certain fit parameters. First, the parameters need to be selected under *parameter initials* or *parameter values*. A column will be added to the table for each selected parameter. By clicking on the head line cell of a column, it is possible to set the parameter this column is used for. The values may be filled in using copy and paste from, e.g., a spreadsheet software.

After the fit has been set up, one can choose between two different fitting modes and the χ^2 explorer:

Standard Fit Each sample is fitted separately.

Ensemble Fit All samples are combined in a single fit. In this mode each fit parameter can either be varied independently for each sample or it can be fitted to the ensemble of samples, i.e., it has the same value for every sample.

χ^2 Explorer In this mode the χ^2 surface of the fit can be analyzed interactively.

B.3. Evaluation of the fit results

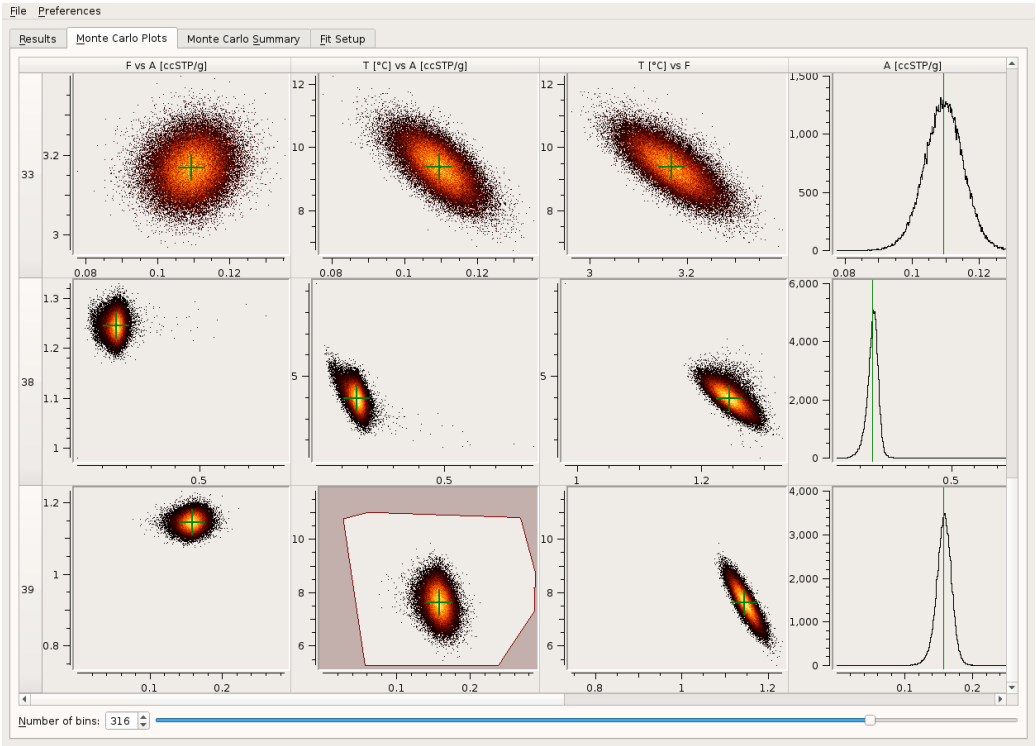


Figure B.2.: Overview of the Monte Carlo results

When the fit has finished, a new window will pop up containing the fit results. The *Results* tab contains the complete fit output for all samples, as listed in section 6.3.

For the CE model, *PANGA* gives hints as to when an additional Monte Carlo analysis is recommended: If any result is outside of the normal range (as described in 4.2), the respective value is shown in red and the whole sample is marked in yellow.

B.3.1. Monte Carlo analysis

The *Monte Carlo Plots* tab contains an overview of the Monte Carlo histograms of the different samples. By default, the following plots are shown: two-dimensional histograms of all possible pairs of fit parameters and one-dimensional plots of all fit parameters. The

green crosses and vertical lines mark the position of the original fit results. The number of histogram bins can be set globally using the slider on the bottom of the window.

The plot window

A double click on a plot will open the plot window. Here, the plot may be zoomed, selections are made and masks are set.

The plot can be zoomed by drawing a rectangle around the desired area using the left mouse button. With a right click you can zoom out again. When the shift button is pressed, a selection is made instead. In a one-dimensional plot an interval of values is selected. In a two-dimensional plot a polygon can be selected using several clicks of the left mouse button. The polygon will be completed if the right mouse button is used. The selected area will then be highlighted. Note, that the red-shaded region indicates the areas that are not selected. With the buttons on the right side, the selection can be modified as well as applied to the sample's mask.

With the slider on the bottom, the bin number of the histogram can be changed individually for this plot.

Using masks

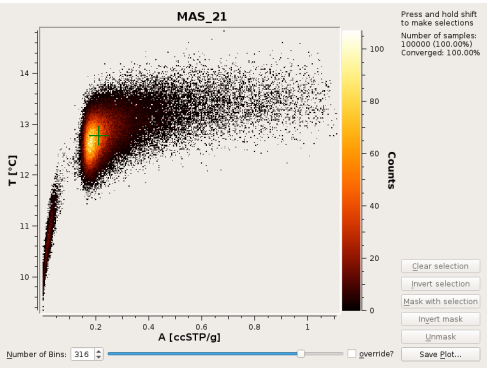
A mask is used to switch on or off specific Monte Carlo realizations for the statistical evaluation. Its use is illustrated in figure B.3. Masks are modified using selections in the plot window: the selection needs to be done so that the Monte Carlo realizations to be removed are in the red-shaded area. After a click on *Mask with selection*, only the Monte Carlo realizations in the selected area (not shaded in red) will remain activated. This process may be repeated to combine several selections in one mask. Note that a mask always affects the whole sample, not only the plot in which it is created.

Monte Carlo results

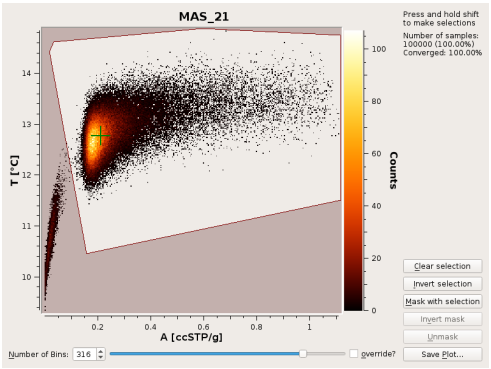
On the right side of the plot window, the results of the statistical analysis are displayed. For 1D plots this analysis consists of mean value and standard deviation of the respective parameter. 2D plots show mean values and standard deviations of both parameters as well as their correlation coefficient. The statistical analysis is restricted to the Monte Carlo realizations not disabled via the mask. The Monte Carlo results of all samples are summarized in a table under the tab *Monte Carlo Summary*.

Add other plots or parameters

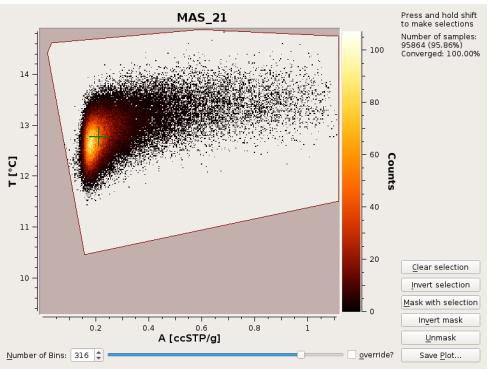
By default, *PANGA* shows plots and carries out a statistical analysis for all fit parameters. Two-dimensional plots are created for all possible pairs of fit parameters. If needed, additional plots can be created via *Preferences* → *Choose Monte Carlo Plots*. A window will show up, which contains a list of all available parameters and the list of the currently shown ones. Parameters can be added to right light list by dragging them from the left-side and dropping them at the end of the list or between two existing entries. If the



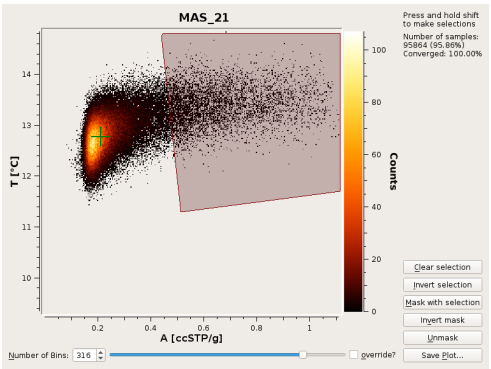
(a) No selection



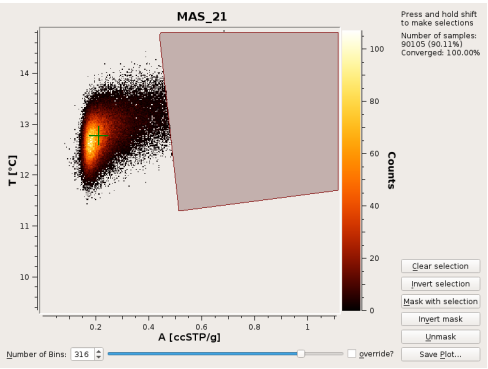
(b) First selection, not yet applied to the mask



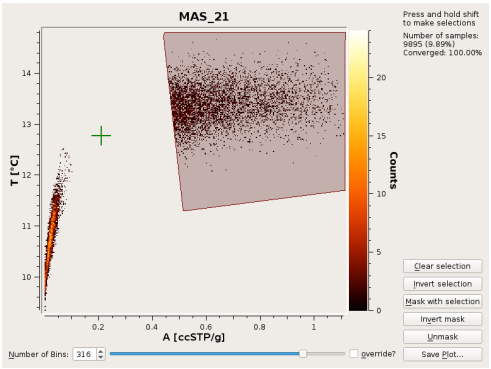
(c) First selection, applied to the mask



(d) Second (inverted) selection, not yet applied to the mask



(e) Second selection applied to the mask



(f) Mask inverted

Figure B.3.: Illustration of masks in *PANGA*. Note that this illustration does not represent a real evaluation. It is only meant to demonstrate the masking feature.

parameter is dropped on an existing list entry, both entries will be combined for a two-dimensional plot. Entries can be removed by dragging them from the right list and dropping them on the left one.

B.3.2. Saving and loading results

The results of a fit, including Monte Carlo results, selections, masks, etc., can be saved to a file using *File* → *Save*. There are two different file types available: binary files and portable files. The binary files can be saved and loaded much faster than the portable files. They are, however, not portable, which means that, e.g., a file saved in the 32-bit Windows version cannot be read with the 64-bit Windows version. Also, the files saved under another operating system cannot be opened. So if cross-architecture or cross-platform compatibility is required, the slower portable format should be used. Saved files can be opened with *File* → *Open results*.

The settings of fits from a saved file (and also of a new fit) can be seen under the *Fit Setup* tab in the results window. If needed, they can be loaded back into the main window by clicking *Load into Main Window*.

If the simple statistical evaluation carried out by *PANGA* is not sufficient, the Monte Carlo results can be exported to a text file, from where they may be processed with other software. The export can be done by choosing *Export Monte Carlo Data* from the *File* window. The file will be in the *csv* format and will contain, for every Monte Carlo realization, all of the columns shown under the *Results* tab. Only the Monte Carlo realizations which are enabled by the sample's mask will be exported.

B.4. The χ^2 explorer

The plot in the χ^2 explorer shows χ^2 as a function of two model parameters. The plot is configured in the table on the right side of the window. It contains options for all the parameters chosen to be fit parameters in the main window. The columns *X* and *Y* are used to set the parameters which are currently displayed on the plot's axes. Their *min* and *max* values determine the range of the axes. The remaining parameters may either be set to fixed values or included in the fit. This is determined by the *fit?* column. If a model parameter is not fitted, a slider will appear, which can be used to set its value. Alternatively, it can also be set in the table. The *min* and *max* values of these parameters only determine the range of the their sliders.

If *auto-update* is selected, the plot will be updated whenever any setting is changed. Otherwise, updates need to be triggered manually with the *Update plot* button.

By default, the scale of the χ^2 colorbar is set to range from the minimal to the maximal value visible in the plot. It may be readjusted with the *Minimum* and *Maximum* fields.

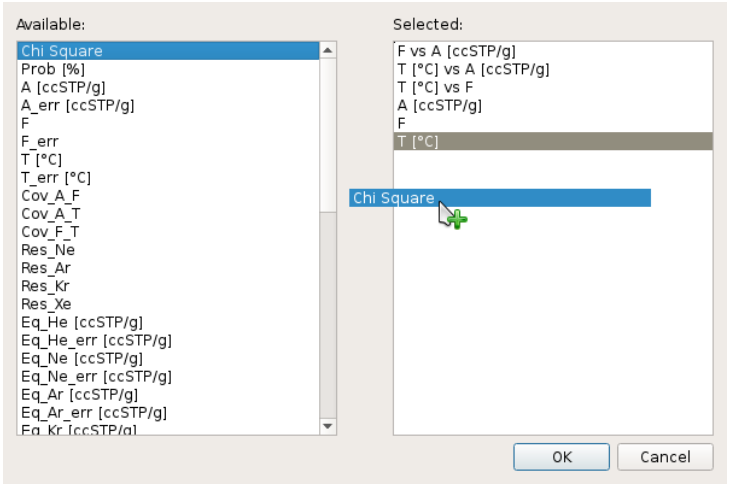


Figure B.4.: Monte Carlo plots setup

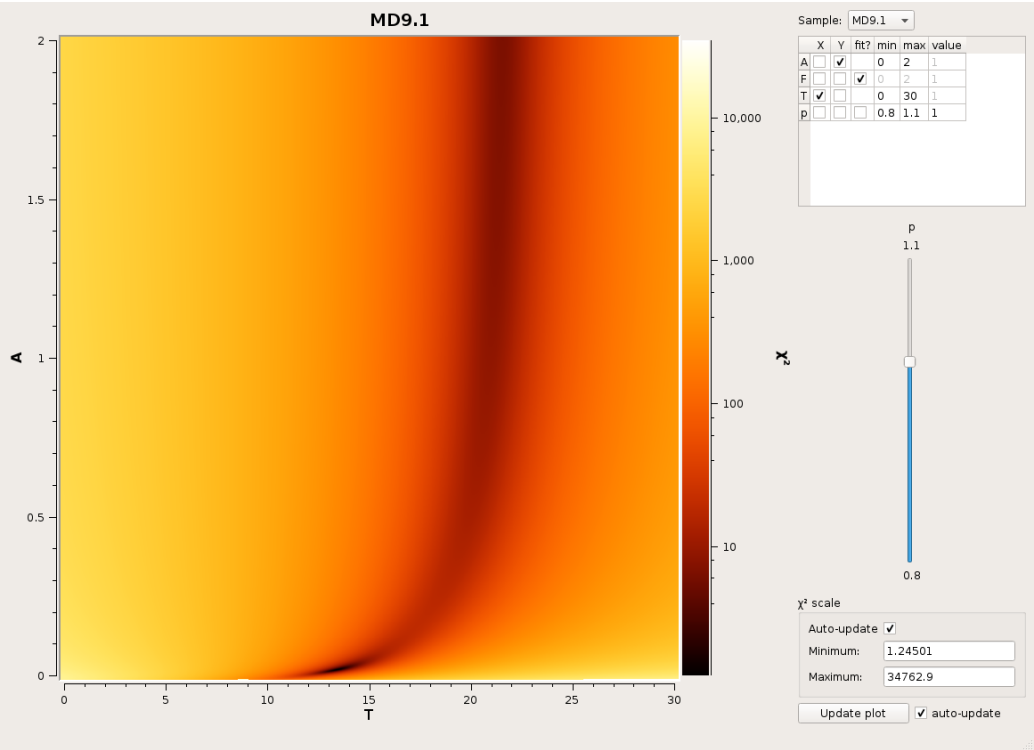


Figure B.5.: The χ^2 explorer

C. Review data

The following tables summarize the fit results for the data sets in chapter 5. The re-evaluation of all data sets was carried out using the CE model. The methods "MC u" and "MC c" stand for unconstrained and constrained Monte Carlo fitting respectively.

Table C.1.: Stute 1995 Brazil

		Original results		Results of re-evaluation					Input parameters	
Sample	Δ NGT [℃]	NGT [℃]	Model	NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	P [atm]	S [g/kg]
7[a]	0.53	29.7 ± 0.8	PR	29.17 ± 0.61	10.47	0.55	2.87	Fit	0.9472	0.0
9[a]	0.69	28.4 ± 0.8	PR	27.71 ± 0.61	12.37	0.56	2.18	Fit	0.9472	0.0
15[a]	0.33	29.8 ± 0.8	PR	28.47 ± 0.67	17.11	0.37	0.17	Fit	0.9472	0.0
18[a]	0.93	30.6 ± 0.8	PR	29.67 ± 0.89	25.95	0.64	0.04	Fit	0.9472	0.0
3[a]	0.47	29.2 ± 0.8	PR	28.73 ± 0.64	14.33	0.35	2.51	Fit	0.9472	0.0
13[a]	2.30	31.1 ± 0.8	PR	28.80 ± 0.71	17.93	0.57	0.13	Fit	0.9472	0.0
8[a]	0.78	30.8 ± 0.8	PR	30.02 ± 0.71	17.17	0.47	0.59	Fit	0.9472	0.0
4[c]	0.15	27.3 ± 0.8	PR	27.15 ± 1.06	13.91	0.10	0.46	MC c	0.9472	0.0
6[a]	1.77	28.4 ± 0.8	PR	26.63 ± 0.60	14.68	0.25	0.61	Fit	0.9472	0.0
19[a]	0.36	29.3 ± 0.8	PR	28.94 ± 0.65	13.12	0.62	2.03	Fit	0.9472	0.0
H10	1.39	28.8 ± 0.8	PR	27.41 ± 0.58	11.22	0.38	7.56	Fit	0.9472	0.0
2[a]	0.52	28.9 ± 0.8	PR	28.38 ± 0.64	13.79	0.53	2.41	Fit	0.9472	0.0
H6[b]	0.66	24.1 ± 0.8	PR	23.44 ± 0.52	12.13	0.13	1.48	Fit	0.9472	0.0
16[a]	1.39	29.7 ± 0.8	PR	28.31 ± 0.68	18.33	0.37	0.37	Fit	0.9472	0.0
H7[b]	1.33	24.4 ± 0.8	PR	23.07 ± 0.57	17.23	0.25	0.00	Fit	0.9472	0.0
1[b]	1.73	25.2 ± 0.8	PR	23.47 ± 0.63	23.67	0.27	0.57	Fit	0.9472	0.0
14[b]	1.69	23.3 ± 0.8	PR	21.61 ± 0.56	19.54	0.29	1.93	Fit	0.9472	0.0
23[b]	1.22	24.8 ± 0.8	PR	23.58 ± 0.60	24.18	0.06	1.10	Fit	0.9472	0.0
11[b]	1.02	23.3 ± 0.8	PR	22.28 ± 0.58	21.55	0.17	1.35	Fit	0.9472	0.0
17	4.48	30.3 ± 0.8	PR	25.82 ± 0.66	22.56	0.24	3.87	Fit	0.9472	0.0
12[b]	0.74	24.0 ± 0.8	PR	23.26 ± 0.59	20.74	0.19	3.86	Fit	0.9472	0.0

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean LGM temperature.
[c] No Ne concentrations were available for this sample.

Table C.2.: Beyerle 1998 Switzerland

Sample	Δ NGT [℃]	Original results NGT [℃]	Model	Results of re-evaluation NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
0[a]	−0.83	7.8 ± 0.6	UA	8.63 ± 0.61	16.64	0.67	0.01	Fit	0.937	0.0
1[a]	−0.23	7.3 ± 0.4	UA	7.53 ± 0.27	2.26	0.11	4.27	MC c	0.937	0.0
2[c]	−0.19	5.6 ± 0.7	UA	5.79 ± 0.30	3.04	0.08	7.39	MC c	0.937	0.0
2c[bc]	−0.32	4.6 ± 1.1	UA	4.92 ± 0.43	2.85	0.05	6.59	MC c	0.937	0.0
3	−0.12	3.8 ± 1.1	UA	3.92 ± 0.24	2.22	0.01	20.32	MC c	0.937	0.0
4[c]	−1.19	6.8 ± 0.6	UA	7.99 ± 0.62	18.58	0.68	0.06	Fit	0.937	0.0
4c[ac]	−1.29	6.1 ± 0.7	UA	7.39 ± 1.38	20.68	0.68	0.07	Fit	0.937	0.0
5	−0.34	2.4 ± 1.4	UA	2.74 ± 0.23	0.61	0.00	30.16	MC c	0.937	0.0
6[c]	−0.07	4.9 ± 0.6	UA	4.97 ± 0.28	3.11	0.12	5.70	MC c	0.937	0.0
6c[bc]	−0.08	4.7 ± 0.6	UA	4.78 ± 0.32	3.24	0.12	5.09	MC c	0.937	0.0
7	−0.96	3.2 ± 1.7	UA	4.16 ± 0.26	6.01	0.00	33.39	MC c	0.937	0.0

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean LGM temperature.
[c] Samples with mixture. The samples with c in their names were corrected for the older component.

Table C.3.: Weyhenmeyer 2000 Oman

Sample	Δ NGT [℃]	Original results NGT [℃]	Model	Results of re-evaluation NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
SLU-2B[a]	−0.04	34.4 ± 1.3	CE	34.44 ± 1.35	18.48	0.37	2.10	Fit	1.0	0.0
RGS-5F[a]	−0.18	33.7 ± 2.3	CE	33.88 ± 1.70	41.45	0.58	4.78	MC u	1.0	0.0
RGS-2U[a]	0.12	32.5 ± 4.4	CE	32.38 ± 1.79	73.80	0.54	6.25	MC u	1.0	0.0
BZ-4	−0.03	29.7 ± 1.4	CE	29.73 ± 1.37	22.51	0.58	3.99	Fit	1.0	0.0
JT-31	−0.03	29.6 ± 1.3	CE	29.63 ± 1.34	20.40	0.70	3.15	Fit	1.0	0.0
21/7	0.03	30.6 ± 1.0	CE	30.57 ± 1.01	9.85	0.45	0.57	Fit	1.0	0.0
RGS-2L[b]	0.05	26.8 ± 1.1	CE	26.75 ± 1.10	15.20	0.71	0.10	Fit	1.0	0.0
KWD-3L[b]	0.01	26.8 ± 1.0	CE	26.79 ± 0.99	10.67	0.79	0.28	Fit	1.0	0.0
KWD-1[b]	0.03	26.2 ± 1.0	CE	26.17 ± 1.02	13.00	0.74	0.21	Fit	1.0	0.0
DP-2[c]	—	—	CE	33.08 ± 2.50	46.79	0.45	8.85	Fit	1.0	0.0
21/6[c]	—	—	CE	34.67 ± 1.10	9.15	0.15	47.56	Fit	1.0	0.0

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean Pleistocene temperature.
[c] No errors of the noble gas concentrations were available for these samples. They were reconstructed for the re-evaluation from the mean values of the relative errors of the other samples in this study.

Table C.4.: Aeschbach-Hertig 2002 Maryland

		Original results		Results of re-evaluation					Input parameters	
Sample	Δ NGT [℃]	NGT [℃]	Model	NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	P [atm]	S [g/kg]
MD1.1	0.03	7.3 ± 1.0	CE	7.27 ± 0.20	15.11	0.00	29.07	MC c	0.994	0.0
MD1.3	−0.10	7.3 ± 1.0	CE	7.40 ± 0.22	13.79	0.01	18.38	MC c	0.994	0.0
MD2.1	0.30	11.0 ± 0.4	CE	10.70 ± 0.55	42.11	0.72	2.57	Fit	0.994	0.0
MD2.2	−0.31	11.0 ± 0.4	CE	11.31 ± 0.60	46.23	0.74	0.05	Fit	0.994	0.0
MD3.3	−0.09	6.1 ± 0.3	CE	6.19 ± 0.49	46.00	0.80	4.27	Fit	0.994	0.0
MD3.4	0.03	6.1 ± 0.3	CE	6.07 ± 0.48	44.79	0.80	1.21	Fit	0.994	0.0
MD4.1	−0.96	8.1 ± 0.3	CE	9.06 ± 0.54	44.75	0.76	0.01	Fit	0.994	0.0
MD4.3	0.21	8.1 ± 0.3	CE	7.89 ± 0.47	38.22	0.77	2.67	Fit	0.994	0.0
MD4.4	0.35	8.1 ± 0.3	CE	7.75 ± 0.37	21.00	0.71	6.29	Fit	0.994	0.0
MD5.2[a]	0.02	14.7 ± 0.7	CE	14.68 ± 0.70	45.10	0.80	0.04	Fit	0.994	0.0
MD6.2[a]	0.64	13.5 ± 1.1	CE	12.88 ± 0.42	46.57	0.88	0.64	MC c	0.994	0.0
MD7.1	−0.13	5.5 ± 1.5	CE	5.63 ± 0.17	3.29	0.00	67.43	MC c	0.994	0.0
MD7.2	0.09	5.5 ± 1.5	CE	5.41 ± 0.17	3.28	0.00	52.15	MC c	0.994	0.0
MD8.2[a]	0.03	13.8 ± 0.5	CE	13.77 ± 0.53	31.26	0.79	0.55	Fit	0.994	0.0
MD9.1[a]	−0.04	13.2 ± 0.4	CE	13.24 ± 0.42	19.80	0.66	1.28	Fit	0.994	0.0
MD10.1	0.05	13.4 ± 0.5	CE	13.35 ± 0.53	32.87	0.73	0.22	Fit	0.994	0.0
MD11.1[c]	0.03	7.2 ± 0.7	CE	7.17 ± 0.70	67.98	0.81	0.18	Fit	0.994	0.0
MD12.1	0.05	5.3 ± 0.3	CE	5.25 ± 0.35	36.04	0.09	0.27	Fit	0.994	0.0
MD12.2	−0.08	5.3 ± 0.3	CE	5.38 ± 0.36	37.79	0.09	0.13	Fit	0.994	0.0
MD13.1	0.02	6.0 ± 0.4	CE	5.98 ± 0.46	41.78	0.76	0.01	Fit	0.994	0.0
MD13.2	0.08	6.0 ± 0.4	CE	5.92 ± 0.52	56.40	0.78	0.90	MC c	0.994	0.0
MD14.1	−0.03	12.6 ± 0.6	CE	12.63 ± 0.58	39.88	0.76	1.17	Fit	0.994	0.0
MD15.2[b]	0.02	5.3 ± 0.4	CE	5.28 ± 0.36	23.86	0.75	2.90	Fit	0.994	0.0
MD16.1	−0.01	6.5 ± 0.5	CE	6.51 ± 0.52	48.91	0.80	0.01	Fit	0.994	0.0
MD17.2[b]	0.02	4.4 ± 0.3	CE	4.38 ± 0.32	17.00	0.78	4.72	Fit	0.994	0.0
MD18.1[b]	0.10	3.3 ± 0.3	CE	3.20 ± 0.28	9.78	0.65	4.33	Fit	0.994	0.0
MD18.2[b]	−0.07	3.3 ± 0.3	CE	3.37 ± 0.29	9.98	0.65	5.02	Fit	0.994	0.0
MD19.1[c]	0.05	5.9 ± 0.7	CE	5.85 ± 0.68	72.11	0.78	0.01	Fit	0.994	0.0
MD20.2	0.05	10.6 ± 0.5	CE	10.55 ± 0.47	30.36	0.81	2.66	Fit	0.994	0.0
MD21.1	0.04	12.1 ± 0.3	CE	12.06 ± 0.42	19.64	0.77	1.09	Fit	0.994	0.0
MD21.2	0.05	12.1 ± 0.3	CE	12.05 ± 0.40	17.60	0.76	1.44	Fit	0.994	0.0
MD22.1[b]	0.11	5.5 ± 0.2	CE	5.39 ± 0.33	16.76	0.75	1.15	Fit	0.994	0.0
MD22.2[b]	−0.06	5.5 ± 0.2	CE	5.56 ± 0.32	13.44	0.75	4.31	Fit	0.994	0.0
MD23.1	0.01	5.2 ± 1.3	CE	5.19 ± 0.17	2.51	0.01	22.19	MC c	0.994	0.0

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean LGM temperature.
[c] The original fit results were used, because the two clusters in the Monte Carlo simulations were not clearly separable.

Table C.5.: Beyerle 2003 Niger

Sample	Δ NGT [°C]	Original results NGT [°C]	Model	Results of re-evaluation NGT [°C]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
2	−0.23	31.4 ± 1.0	CE	31.63 ± 1.02	6.58	0.31	1.15	Fit	0.9703	0.0
5[a]	−0.03	32.7 ± 1.3	CE	32.73 ± 1.35	19.33	0.46	0.04	Fit	0.9703	0.0
18[a]	−0.09	31.9 ± 1.1	CE	31.99 ± 1.13	12.88	0.40	0.21	Fit	0.9703	0.0
19[a]	0.04	30.2 ± 1.0	CE	30.16 ± 0.95	6.50	0.47	0.36	Fit	0.9703	0.0
20[a]	0.08	30.2 ± 0.8	CE	30.12 ± 0.75	10.58	0.55	0.00	Fit	0.9703	0.0
21[a]	−0.02	32.4 ± 0.8	CE	32.42 ± 0.80	10.03	0.60	1.00	Fit	0.9703	0.0
22[a]	−0.04	33.2 ± 1.0	CE	33.24 ± 1.03	20.22	0.73	2.29	Fit	0.9703	0.0
24[a]	−0.09	32.9 ± 1.1	CE	32.99 ± 1.13	10.79	0.54	0.40	Fit	0.9703	0.0
28[a]	−0.03	31.6 ± 1.3	CE	31.63 ± 1.34	16.81	0.48	0.04	Fit	0.9703	0.0
30	−0.12	31.7 ± 1.6	CE	31.82 ± 1.61	26.81	0.57	0.89	Fit	0.9703	0.0
31[a]	0.06	30.2 ± 0.7	CE	30.14 ± 0.72	8.56	0.46	0.04	Fit	0.9703	0.0
33[a]	−0.06	30.5 ± 0.8	CE	30.56 ± 0.84	16.53	0.44	0.10	Fit	0.9703	0.0
34[a]	0.03	32.5 ± 0.9	CE	32.47 ± 0.94	18.66	0.58	0.85	Fit	0.9703	0.0
37[a]	−0.03	31.5 ± 0.7	CE	31.53 ± 0.72	7.13	0.57	0.13	Fit	0.9703	0.0
40[a]	0.02	31.2 ± 0.7	CE	31.18 ± 0.69	4.31	0.43	0.42	Fit	0.9703	0.0
41[a]	0.05	31.5 ± 0.9	CE	31.45 ± 0.93	3.82	0.46	0.23	Fit	0.9703	0.0
85	−0.25	32.4 ± 1.0	CE	32.65 ± 1.04	24.46	0.46	7.70	Fit	0.9703	0.0
80[b]	−0.26	29.7 ± 0.8	CE	29.96 ± 0.85	23.63	0.44	5.38	Fit	0.9703	0.0
84[b]	−0.23	31.2 ± 0.7	CE	31.43 ± 0.73	18.78	0.37	3.56	Fit	0.9703	0.0
82[b]	−0.20	28.7 ± 1.0	CE	28.90 ± 1.03	16.34	0.45	0.03	Fit	0.9703	0.0
83[b]	−0.20	29.3 ± 1.1	CE	29.50 ± 1.08	20.37	0.30	0.05	Fit	0.9703	0.0
1[d]	−0.33	29.6 ± 1.3	CE	29.93 ± 1.34	19.28	0.49	0.27	Fit	0.9703	0.0
12[d]	−0.19	28.9 ± 0.8	CE	29.09 ± 0.82	19.71	0.32	1.24	Fit	0.9703	0.0
4[d]	−0.16	26.8 ± 1.1	CE	26.96 ± 1.14	21.47	0.29	0.47	Fit	0.9703	0.0
7[cd]	−0.16	25.7 ± 0.8	CE	25.86 ± 0.79	15.72	0.30	0.07	Fit	0.9703	0.0
9[cd]	−0.05	26.3 ± 1.1	CE	26.35 ± 1.12	20.96	0.34	2.42	Fit	0.9703	0.0
10[cd]	−0.21	26.1 ± 1.1	CE	26.31 ± 1.14	20.26	0.47	1.27	Fit	0.9703	0.0
17[d]	−0.19	27.6 ± 1.3	CE	27.79 ± 1.29	24.03	0.28	0.03	Fit	0.9703	0.0
27[d]	−0.24	26.7 ± 0.8	CE	26.94 ± 0.78	12.95	0.31	0.00	Fit	0.9703	0.0
29[d]	−0.20	27.6 ± 0.8	CE	27.80 ± 0.82	15.38	0.28	0.62	Fit	0.9703	0.0
35[d]	−0.27	27.2 ± 1.7	CE	27.47 ± 1.67	41.29	0.34	1.01	Fit	0.9703	0.0
3[d]	−0.28	26.0 ± 1.2	CE	26.28 ± 1.18	25.19	0.27	3.47	Fit	0.9703	0.0
8[d]	−0.33	27.3 ± 1.3	CE	27.63 ± 1.29	28.02	0.26	2.93	Fit	0.9703	0.0
25[d]	−0.31	26.9 ± 1.3	CE	27.21 ± 1.31	29.84	0.25	0.81	Fit	0.9703	0.0

[a] Belongs to the group of modern samples from the upper (CT3) aquifer.
[b] Belongs to the group of samples from the recharge area of the middle aquifer (CT2).
[c] Belongs to the group of CT2 samples between 10 and 14 kyr BP.
[d] Belongs to the group of all confined samples from the middle and lower aquifers.

Table C.6.: Kulongoski 2004 Kalahari

		Original results		Results of re-evaluation					Input parameters	
Sample	Δ NGT [℃]	NGT [℃]	Model	NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	P [atm]	S [g/kg]
4507[a]	0.04	23.3 ± 0.8	CE	23.26 ± 1.46	32.43	0.25	0.30	Fit	0.8759	0.0
769[b]	0.02	16.2 ± 2.0	PD	16.18 ± 1.11	33.22	0.03	7.08	MC c	0.8759	0.0
6742	−0.20	24.5 ± 1.6	PD	24.70 ± 1.12	16.70	0.06	19.91	Fit	0.8759	0.0
736[a]	0.99	23.0 ± 0.5	PR	22.01 ± 1.02	10.07	0.60	0.48	Fit	0.8759	0.0
6767	−0.50	22.1 ± 1.3	PD	22.60 ± 0.99	17.14	0.04	9.73	MC c	0.8759	0.0
8866[a]	1.55	24.1 ± 1.3	PR	22.55 ± 1.05	13.37	0.23	0.84	Fit	0.8759	0.0
953[b]	−0.14	17.4 ± 0.6	CE	17.54 ± 0.75	13.70	0.03	3.20	MC c	0.8759	0.0
6586[b]	0.09	17.8 ± 0.4	CE	17.71 ± 0.60	7.16	0.05	1.65	MC c	0.8759	0.0
6586_2[b]	0.34	18.9 ± 0.5	CE	18.56 ± 0.57	7.91	0.02	1.98	MC c	0.8759	0.0
8863	−4.88	17.3 ± 1.2	PD	22.18 ± 0.54	6.97	0.00	24.13	MC c	0.8759	0.0
6462	−3.43	18.6 ± 1.1	PD	22.03 ± 0.50	6.02	0.00	8.21	MC c	0.8759	0.0
2869[b]	−0.03	20.1 ± 0.5	CE	20.13 ± 0.73	8.20	0.08	1.19	MC c	0.8759	0.0

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean temperature of samples older than 24 kyr BP.

Table C.7.: Ma 2004 Michigan

		Original results		Results of re-evaluation					Input parameters	
Sample	Δ NGT [℃]	NGT [℃]	Model	NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	P [atm]	S [g/kg]
MI-1	1.09	6.8 ± 0.9	UA	6.69 ± 0.39	8.96	0.00	13.10	MC c	0.9645	0.0
MI-2a	0.07	3.5 ± 0.8	UA	3.38 ± 0.37	9.08	0.00	10.94	MC c	0.9645	0.0
MI-2b	0.09	3.1 ± 0.6	UA	3.05 ± 0.33	2.13	0.00	7.59	MC c	0.9645	0.0
MI-3	1.08	5.8 ± 0.9	UA	5.70 ± 0.34	0.62	0.00	17.05	MC c	0.9645	0.0
MI-4a	−0.04	3.0 ± 0.3	UA	3.02 ± 0.47	5.73	0.19	2.50	MC c	0.9645	0.0
MI-4b	−0.63	2.5 ± 0.5	UA	3.13 ± 0.74	15.40	0.69	2.46	Fit	0.9645	0.0
MI-6	0.02	6.5 ± 0.9	UA	6.44 ± 0.35	1.23	0.03	14.28	MC c	0.9645	0.0
MI-7	0.16	7.1 ± 0.8	UA	6.89 ± 0.35	1.15	0.00	12.81	MC c	0.9645	0.0
MI-8[a]	−0.16	5.0 ± 0.4	UA	5.14 ± 0.55	13.78	0.05	1.06	MC c	0.9645	0.0
MI-9	−0.07	4.5 ± 0.3	UA	4.55 ± 0.47	5.72	0.23	1.31	MC c	0.9645	0.0
MI-10a[c]	−2.03	7.0 ± 2.0	UA	8.99 ± 0.79	30.86	0.64	34.25	MC c	0.9645	0.0
MI-10b	−7.54	4.2 ± 2.8	UA	11.70 ± 0.63	247.89	0.69	57.43	MC c	0.9645	0.0
MI-11	−0.30	4.0 ± 0.3	UA	4.32 ± 0.68	8.60	0.51	1.06	Fit	0.9645	0.0
MI-12a	−2.25	3.9 ± 0.9	UA	6.17 ± 1.18	43.37	0.70	3.04	Fit	0.9645	0.0
MI-12b	−1.08	4.9 ± 0.5	UA	5.99 ± 0.83	19.09	0.62	0.02	Fit	0.9645	0.0
MI-13	0.02	2.8 ± 0.9	UA	2.76 ± 0.37	3.08	0.08	14.98	MC c	0.9645	0.0
MI-14a	−2.06	1.1 ± 1.1	UA	3.17 ± 0.93	37.93	0.64	13.98	Fit	0.9645	0.0
MI-14b[b]	−0.65	1.3 ± 0.5	UA	1.96 ± 0.68	13.40	0.45	1.67	Fit	0.9645	0.0
MI-15	0.02	2.3 ± 0.3	UA	2.28 ± 0.44	5.67	0.15	2.02	MC c	0.9645	0.0

[a] Used for the calculation of the Holocene temperature.
[b] Used for the calculation of the LGM temperature.
[c] Only 6.5 % of the Monte Carlo realizations were in the realistic cluster.

Table C.8.: Kreuzer 2009 China

		Original results		Results of re-evaluation					Input parameters	
Sample	Δ NGT [℃]	NGT [℃]	Model	NGT [℃]	$A \cdot 10^3$ [cm^3 STP/g]	F	χ^2	Method	P [atm]	S [g/kg]
1[ab]	−0.25	12.4 ± 0.4	CE	12.68 ± 0.27	3.37	0.22	1.22	MC c	0.9873	0.0
2[ab]	−0.33	13.8 ± 0.5	CE	14.14 ± 0.29	2.34	0.05	3.24	MC c	0.9900	0.0
3[ab]	−0.14	12.0 ± 0.6	CE	12.16 ± 0.58	15.47	0.87	0.00	Fit	0.9908	0.0
4[b]	−0.15	11.7 ± 0.5	CE	11.89 ± 0.30	8.29	0.13	3.04	MC c	0.9939	0.0
5[ab]	−0.13	11.8 ± 0.5	CE	11.90 ± 0.50	5.25	0.30	0.56	Fit	0.9912	0.0
7[ab]	−0.20	14.8 ± 0.6	CE	15.04 ± 0.55	7.64	0.58	0.00	Fit	0.9928	0.0
8[ab]	−0.16	13.9 ± 0.5	CE	14.08 ± 0.53	7.61	0.76	0.01	Fit	0.9883	0.0
9[cd]	−0.19	8.2 ± 0.3	CE	8.42 ± 0.32	1.90	0.16	1.00	Fit	0.9940	0.0
10[b]	−0.12	12.1 ± 0.5	CE	12.19 ± 0.28	3.90	0.21	4.32	MC c	0.9940	0.0
11[c]	−0.16	11.2 ± 0.5	CE	11.37 ± 0.49	7.92	0.12	0.46	Fit	0.9940	0.0
12[cd]	−0.11	9.2 ± 0.5	CE	9.35 ± 0.46	3.95	0.34	0.01	Fit	0.9940	0.0
13	—	—	CE	14.55 ± 0.22	0.33	0.03	10.40	MC c	0.9940	0.0
14[cd]	−0.25	8.9 ± 0.5	CE	9.10 ± 0.33	4.06	0.22	1.60	MC c	0.9940	0.0
15	−0.15	14.4 ± 0.7	CE	14.58 ± 0.66	18.89	0.77	0.88	Fit	0.9940	0.0
16[cd]	−0.15	9.0 ± 0.5	CE	9.12 ± 0.47	1.63	0.07	0.58	Fit	0.9940	0.0
17[cd]	−0.28	8.3 ± 0.5	CE	8.61 ± 0.28	2.13	0.06	3.03	MC c	0.9940	0.0
18[c]	−0.17	10.9 ± 0.4	CE	11.09 ± 0.43	11.70	0.67	2.00	Fit	0.9940	0.0
19[cd]	−0.31	8.8 ± 0.5	CE	9.12 ± 0.37	5.04	0.27	1.23	MC c	0.9940	0.0
20[c]	−0.27	9.9 ± 0.5	CE	10.18 ± 0.30	3.24	0.08	2.44	MC c	0.9940	0.0
31[ab]	−0.12	15.5 ± 1.2	CE	15.60 ± 1.19	26.04	0.52	0.12	Fit	0.9926	0.0
32[ab]	−0.14	13.8 ± 0.6	CE	13.94 ± 0.62	21.71	0.51	0.88	Fit	0.9892	0.0
33[ab]	−0.14	13.7 ± 1.1	CE	13.84 ± 1.06	21.17	0.57	0.02	Fit	0.9885	0.0
36[b]	−0.09	13.2 ± 0.7	CE	13.26 ± 0.69	4.89	0.50	0.82	Fit	0.9940	0.0
37[c]	−0.11	10.6 ± 0.9	CE	10.66 ± 0.87	12.50	0.73	0.36	Fit	0.9940	0.0
38[c]	−0.16	10.4 ± 0.9	CE	10.53 ± 0.88	12.58	0.68	0.21	Fit	0.9924	0.0
39	−0.12	17.7 ± 0.8	CE	17.77 ± 0.80	29.47	0.51	0.72	Fit	0.9916	0.0
40[ab]	−0.12	13.0 ± 0.6	CE	13.15 ± 0.57	14.47	0.60	0.58	Fit	0.9909	0.0
41[b]	−0.08	12.7 ± 0.8	CE	12.81 ± 0.80	11.43	0.88	0.87	Fit	0.9928	0.0
43[b]	−0.09	11.6 ± 0.5	CE	11.71 ± 0.50	10.14	0.86	0.23	Fit	0.9939	0.0
44[ab]	−0.16	13.8 ± 1.1	CE	13.97 ± 1.09	27.59	0.84	0.64	Fit	0.9940	0.0
45	−0.18	16.5 ± 0.8	CE	16.70 ± 0.82	9.36	0.65	0.00	Fit	0.9940	0.0
46	−0.10	16.7 ± 1.0	CE	16.83 ± 1.02	19.00	0.61	0.98	Fit	0.9940	0.0
47[b]	−0.08	13.0 ± 0.8	CE	13.08 ± 0.79	10.46	0.74	0.33	Fit	0.9940	0.0
48[ab]	−0.11	14.4 ± 0.6	CE	14.48 ± 0.60	15.15	0.50	6.51	Fit	0.9940	0.0
49[ab]	−0.12	14.5 ± 0.6	CE	14.62 ± 0.56	12.93	0.68	1.26	Fit	0.9940	0.0
50[b]	−0.08	13.9 ± 0.7	CE	13.94 ± 0.65	17.44	0.74	0.59	Fit	0.9926	0.0
51[e]	—	—	CE	14.89 ± 0.30	0.00	−6165	13.03	MC u	0.9926	0.0
52[ab]	−0.14	13.8 ± 1.1	CE	13.91 ± 0.70	57.11	0.81	5.92	MC u	0.9916	0.0
53[ab]	−0.13	14.6 ± 1.1	CE	14.73 ± 1.05	25.35	0.68	0.30	Fit	0.9896	0.0
56[ab]	−0.11	14.9 ± 1.1	CE	15.03 ± 1.10	29.63	0.49	0.05	Fit	0.9871	0.0
57[b]	−0.08	12.1 ± 1.1	CE	12.18 ± 1.07	20.66	0.96	0.28	Fit	0.9932	0.0

[a] Belongs to the group of modern samples.
[b] Belongs to the group of all Holocene samples.
[c] Belongs to the group of Pleistocene samples.
[d] Belongs to the group of the six coldest samples.
[e] Constrained fits could not describe this sample. 99.96 % of the MC realizations lie in the region of $F = -6000$.

Table C.9.: Kulongoski 2009 California

Sample	Δ NGT [℃]	Original results NGT [℃]	Model	Results of re-evaluation NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
1[b]	0.00	12.9 ± 1.1	CE	12.90 ± 1.14	23.84	0.38	0.01	Fit	0.8759	0.030
2a[b]	0.02	11.9 ± 1.1	CE	11.88 ± 1.07	22.63	0.31	0.08	Fit	0.8759	0.030
2b[b]	0.58	11.7 ± 1.3	PR	11.12 ± 0.88	10.92	0.32	0.67	Fit	0.8759	0.030
2c[b]	0.74	12.6 ± 1.6	PR	11.86 ± 1.03	17.23	0.56	0.10	Fit	0.8759	0.030
3a[b]	−0.31	10.6 ± 0.4	CE	10.91 ± 0.64	6.43	0.23	1.12	MC c	0.8759	0.030
3b[b]	−0.16	10.7 ± 0.4	CE	10.86 ± 0.61	5.47	0.17	1.52	MC c	0.8759	0.030
3c[b]	−0.42	12.2 ± 0.5	CE	12.62 ± 0.91	8.00	0.62	0.68	Fit	0.8759	0.030
4a[b]	0.01	12.7 ± 1.0	CE	12.69 ± 1.05	19.49	0.32	0.00	Fit	0.8759	0.030
4b[b]	0.05	11.7 ± 1.1	CE	11.65 ± 1.13	24.09	0.50	0.05	Fit	0.8759	0.030
5[b]	−0.13	12.7 ± 0.5	CE	12.83 ± 0.59	8.91	0.03	1.66	MC c	0.8759	0.030
6a[b]	−0.26	12.1 ± 0.5	CE	12.36 ± 0.68	6.59	0.24	1.05	MC c	0.8759	0.030
6b	−2.44	12.9 ± 1.6	PD	15.34 ± 0.55	5.56	0.00	7.33	MC c	0.8759	0.030
7a[b]	−0.07	12.2 ± 0.5	CE	12.27 ± 0.53	5.18	0.05	2.25	MC c	0.8759	0.030
7b[b]	−0.13	12.1 ± 0.5	CE	12.23 ± 0.82	5.01	0.10	1.56	Fit	0.8759	0.030
8a[a]	−0.11	15.6 ± 0.5	CE	15.71 ± 0.61	4.67	0.09	2.85	MC c	0.8759	0.030
8b[a]	0.19	16.2 ± 1.3	PR	16.01 ± 0.93	6.57	0.23	0.00	Fit	0.8759	0.030
9[a]	0.00	17.2 ± 1.4	CE	17.20 ± 1.43	26.35	0.54	0.03	Fit	0.8759	0.030
10[a]	0.12	16.4 ± 0.5	CE	16.28 ± 0.68	5.36	0.18	1.46	MC c	0.8759	0.030

[a] Used for the calculation of the mean Holocene temperature.
[b] Used for the calculation of the mean temperature of the late Pleistocene.

Table C.10.: Blaser 2010 Belgium (continued on the following page)

Sample	Δ NGT [℃]	Original results NGT [℃]	Model	Results of re-evaluation NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
TB532[a]	0.57	10.1 ± 1.4	CE	9.53 ± 0.84	67.92	0.77	1.01	MC u	0.9976	0.432
MW901	0.00	5.6 ± 0.6	CE	5.55 ± 0.63	13.19	0.55	1.77	Fit	0.9976	0.504
VZELE	0.49	6.1 ± 1.1	CE	5.56 ± 0.57	32.99	0.77	0.66	MC c	0.9976	1.115
GD034	0.21	6.2 ± 0.7	CE	6.03 ± 0.58	6.49	0.79	0.05	Fit	0.9976	1.763
GROEDE[c]	1.70	10.5 ± 0.6	CE	8.83 ± 0.52	57.70	0.87	0.77	MC c	0.9976	4.438
1[d]	0.01	7.7 ± 1.2	CE	7.72 ± 1.18	68.26	0.72	0.14	Fit	0.9976	0.619
2	−0.11	7.4 ± 1.0	CE	7.53 ± 1.01	77.03	0.74	1.05	MC u	0.9976	0.574
3	0.01	8.2 ± 0.9	CE	8.14 ± 0.77	36.86	0.66	0.24	Fit	0.9976	0.567
4[d]	0.01	8.2 ± 1.7	CE	8.23 ± 1.19	68.24	0.70	0.08	Fit	0.9976	0.601
5[c]	1.00	9.0 ± 0.6	CE	7.99 ± 0.48	54.75	0.88	0.77	MC c	0.9976	0.637
6[b]	1.89	3.2 ± 1.0	CE	1.35 ± 0.56	75.63	0.82	0.73	MC c	0.9976	0.512
6new[b]	0.01	0.6 ± 2.0	CE	0.59 ± 0.58	35.27	0.80	4.91	Fit	0.9976	0.476
7[b]	0.00	1.1 ± 0.4	CE	1.10 ± 0.42	0.95	0.14	0.00	Fit	0.9976	0.482
7new[b]	−0.01	0.7 ± 0.3	CE	0.67 ± 0.24	1.33	0.07	4.13	MC c	0.9976	0.447
8	−0.05	4.6 ± 0.6	CE	4.69 ± 0.64	29.25	0.66	0.81	MC c	0.9976	0.596
9[b]	−0.14	1.8 ± 0.4	CE	1.89 ± 0.35	11.45	0.39	1.07	MC c	0.9976	0.528
9new[b]	−0.01	1.3 ± 0.3	CE	1.27 ± 0.24	1.05	0.02	6.50	MC c	0.9976	0.501

Table C.10 (cont.): Blaser 2010 Belgium

Sample	Δ NGT [℃]	Original results NGT [℃]	Model	Results of re-evaluation NGT [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	χ^2	Method	Input parameters P [atm]	S [g/kg]
10	0.49	5.0 ± 1.4	CE	4.48 ± 0.60	45.71	0.81	0.82	MC c	0.9976	0.730
11	0.72	6.3 ± 0.9	CE	5.58 ± 0.50	40.27	0.82	0.68	MC c	0.9976	1.584
12[e]	−0.15	10.4 ± 0.5	CE	10.53 ± 0.33	0.85	3394.30	2.00	MC u	0.9976	0.542
13	−0.03	8.6 ± 1.1	CE	8.64 ± 0.92	56.97	0.70	0.71	MC u	0.9976	0.615
14[ad]	0.01	9.7 ± 1.2	CE	9.72 ± 1.16	58.97	0.76	0.08	Fit	0.9976	0.518
15[f]	0.09	8.9 ± 1.1	CE	8.80 ± 1.10	32.88	2.43	0.72	MC c	0.9976	0.712
16[a]	1.51	11.6 ± 1.0	CE	10.13 ± 0.61	51.60	0.84	0.67	MC c	0.9976	0.470
17[f]	−0.07	6.6 ± 1.7	CE	6.66 ± 1.89	136.11	1.78	1.12	MC c	0.9976	0.795
18[f]	−0.04	4.9 ± 0.7	CE	4.97 ± 0.82	191.96	1.35	3.09	MC c	0.9976	0.780
18new[f]	−0.01	3.4 ± 2.0	CE	3.41 ± 0.49	158.72	1.44	19.03	MC c	0.9976	0.657
19	0.01	4.6 ± 0.6	CE	4.58 ± 0.56	17.67	0.80	0.10	Fit	0.9976	0.656
20	0.01	4.5 ± 0.8	CE	4.46 ± 0.66	31.28	0.82	0.00	Fit	0.9976	0.819
21	1.14	6.8 ± 0.8	CE	5.61 ± 0.50	58.96	0.86	0.74	MC c	0.9976	1.142
22[g]	0.76	6.1 ± 0.5	CE	5.36 ± 0.44	59.16	0.90	1.81	MC c	0.9976	1.862
23	0.12	4.7 ± 0.7	CE	4.60 ± 0.62	25.99	0.78	0.13	Fit	0.9976	1.254
24[d]	0.00	4.6 ± 1.1	CE	4.64 ± 1.18	74.04	0.79	0.22	Fit	0.9976	0.435
25[g]	1.22	9.1 ± 1.0	CE	7.89 ± 0.68	96.89	0.75	2.08	MC u	0.9976	0.472
26	0.09	5.5 ± 0.8	CE	5.45 ± 0.88	49.23	0.89	0.17	Fit	0.9976	1.089
27[f]	−0.14	6.9 ± 1.4	CE	7.06 ± 1.35	95.68	3.43	6.62	MC c	0.9976	0.703
27new[f]	−0.09	6.8 ± 3.2	CE	6.86 ± 0.99	99.13	3.27	12.64	MC c	0.9976	0.786
28[f]	−0.01	9.4 ± 3.7	CE	9.39 ± 0.62	109.36	3.17	39.67	MC c	0.9976	0.672
29	0.00	4.4 ± 0.4	CE	4.43 ± 0.46	1.93	0.13	0.78	Fit	0.9976	0.637
30	−0.12	4.6 ± 0.3	CE	4.67 ± 0.36	5.68	0.27	1.36	MC c	0.9976	0.636
31	0.65	8.4 ± 1.2	CE	7.76 ± 0.58	36.45	0.80	0.64	MC c	0.9976	1.977
32	−0.01	2.7 ± 0.3	CE	2.66 ± 0.25	1.35	0.04	4.83	MC c	0.9976	0.500
33[f]	−0.04	4.0 ± 1.6	CE	4.00 ± 0.50	155.34	1.25	9.46	MC c	0.9976	0.585
34[f]	0.00	7.6 ± 2.1	CE	7.63 ± 0.50	157.47	1.15	21.84	MC c	0.9976	0.594

[a] Belongs to the group of three modern non-degassed samples.
[b] Belongs to the group of the coldest six samples.
[c] Only few Monte Carlo realizations in the realistic cluster (< 6%).
[d] The clusters in the Monte Carlo results cannot be separated. The fit result were used as they still look realistic.
[e] Degassed UA limit case. Around 78 % of the results lie in the area of $F = 5000$. Most of the remaining MC realizations have $1 < F < 5$. A few results are between both groups.
[f] Degassed samples. An initial parameter value of $F > 1$ was used for fitting.
[g] The clusters in the Monte Carlo results cannot be separated very well. They had to be used, however, because the fit results lie in the physically unrealistic cluster.

D. Comparison of NOBLE and PANGA

Table D.1.: Comparison of *NOBLE* and *PANGA* (continued on the following page)

Sample	*NOBLE*			*PANGA*			Deviations					
	T	$A \cdot 10^3$	F	T	$A \cdot 10^3$	F	T	σ_T	$A \cdot 10^3$	$\sigma_A \cdot 10^3$	F	σ_F
	[℃]	[cm³ STP/g]		[℃]	[cm³ STP/g]		[℃]		[cm³ STP/g]			
TB532	10.100 ± 2.044	85.89 ± 90.79	0.780 ± 0.019	10.091 ± 2.034	85.55 ± 90.05	0.780 ± 0.019	0.0093	0.0104	0.34	0.75	0.000 02	0.000 05
MW901	5.551 ± 0.633	13.22 ± 7.44	0.553 ± 0.116	5.547 ± 0.633	13.19 ± 7.43	0.552 ± 0.116	0.0037	0.0002	0.03	0.01	0.000 43	0.000 43
VZELE	6.047 ± 1.116	55.09 ± 82.38	0.887 ± 0.027	6.039 ± 1.108	54.58 ± 81.28	0.887 ± 0.027	0.0084	0.0076	0.51	1.09	0.000 10	0.000 18
GD034	6.242 ± 0.575	5.55 ± 19.18	0.779 ± 0.509	6.238 ± 0.574	5.46 ± 19.11	0.776 ± 0.525	0.0038	0.0007	0.10	0.07	0.002 50	0.015 97
GROEDE	10.528 ± 7.685	177.20 ± 1475.80	0.889 ± 0.021	10.524 ± 7.687	177.00 ± 1475.11	0.889 ± 0.021	0.0035	0.0024	0.20	0.69	0.000 05	0.000 04
1	7.732 ± 1.188	68.45 ± 37.94	0.724 ± 0.021	7.724 ± 1.185	68.26 ± 37.77	0.724 ± 0.021	0.0082	0.0031	0.19	0.17	0.000 04	0.000 04
2	7.425 ± 1.147	66.15 ± 40.81	0.750 ± 0.022	7.417 ± 1.144	65.96 ± 40.63	0.750 ± 0.022	0.0077	0.0030	0.20	0.19	0.000 04	0.000 04
3	8.148 ± 0.775	36.94 ± 13.73	0.656 ± 0.031	8.142 ± 0.774	36.86 ± 13.70	0.655 ± 0.032	0.0061	0.0007	0.07	0.03	0.000 11	0.000 06
4	8.244 ± 1.197	68.43 ± 34.23	0.700 ± 0.021	8.235 ± 1.194	68.24 ± 34.07	0.700 ± 0.021	0.0090	0.0032	0.19	0.16	0.000 05	0.000 04
5	8.988 ± 7.529	184.08 ± 1606.50	0.893 ± 0.022	8.986 ± 7.535	184.15 ± 1609.18	0.893 ± 0.022	0.0027	0.0058	0.08	2.68	0.000 05	0.000 05
6	3.243 ± 3.983	331.78 ± 1023.69	0.834 ± 0.029	3.243 ± 3.966	332.72 ± 1022.93	0.834 ± 0.029	0.0003	0.0167	0.93	0.76	0.000 05	0.000 04
6new	0.597 ± 0.582	35.51 ± 23.77	0.804 ± 0.035	0.589 ± 0.581	35.27 ± 23.63	0.804 ± 0.036	0.0077	0.0014	0.24	0.14	0.000 25	0.000 23
7	1.103 ± 0.424	1.09 ± 18.70	0.239 ± 11.996	1.099 ± 0.424	0.95 ± 18.62	0.140 ± 15.551	0.0045	0.0005	0.13	0.09	0.098 26	3.554 76
8	4.639 ± 0.615	26.78 ± 11.93	0.680 ± 0.045	4.634 ± 0.615	26.71 ± 11.91	0.680 ± 0.045	0.0052	0.0004	0.07	0.02	0.000 19	0.000 13
9	1.751 ± 0.432	1.09 ± 17.33	0.194 ± 11.750	1.747 ± 0.431	0.97 ± 17.25	0.101 ± 14.838	0.0044	0.0005	0.12	0.08	0.092 90	3.087 79
10	4.971 ± 1.162	73.30 ± 86.17	0.854 ± 0.023	4.958 ± 1.150	72.52 ± 84.59	0.854 ± 0.023	0.0128	0.0123	0.78	1.57	0.000 09	0.000 15
11	6.302 ± 1.705	93.07 ± 239.11	0.907 ± 0.021	6.288 ± 1.666	91.53 ± 230.68	0.907 ± 0.021	0.0136	0.0384	1.54	8.43	0.000 03	0.000 22
13	8.614 ± 1.001	53.16 ± 25.44	0.710 ± 0.025	8.607 ± 0.999	53.04 ± 25.36	0.710 ± 0.025	0.0069	0.0017	0.12	0.08	0.000 06	0.000 04
14	9.731 ± 1.163	59.14 ± 38.08	0.757 ± 0.024	9.723 ± 1.160	58.97 ± 37.91	0.757 ± 0.024	0.0074	0.0029	0.17	0.17	0.000 05	0.000 04

Table D.1 (cont.): Comparison of *NOBLE* and *PANGA*

Sample	*NOBLE*			*PANGA*			Deviations					
	T	$A \cdot 10^3$	F	T	$A \cdot 10^3$	F	T	σ_T	$A \cdot 10^3$	$\sigma_A \cdot 10^3$	F	σ_F
	[℃]	[cm^3 STP/g]		[℃]	[cm^3 STP/g]		[℃]		[cm^3 STP/g]			
15	8.893 ± 0.982	30.91 ± 7.84	2.420 ± 0.092	8.885 ± 0.984	30.96 ± 7.86	2.420 ± 0.092	0.0086	0.0011	0.05	0.02	0.000 37	0.000 11
16	11.642 ± 7.560	201.04 ± 1354.31	0.867 ± 0.026	11.639 ± 7.561	201.01 ± 1354.75	0.867 ± 0.026	0.0031	0.0008	0.02	0.44	0.000 05	0.000 01
17	6.586 ± 4.327	120.08 ± 147.32	1.769 ± 0.027	6.569 ± 4.363	120.55 ± 149.08	1.769 ± 0.027	0.0170	0.0358	0.47	1.76	0.000 13	0.000 03
18	4.932 ± 6.233	183.18 ± 586.63	1.344 ± 0.039	4.927 ± 6.236	183.30 ± 587.15	1.344 ± 0.039	0.0046	0.0034	0.12	0.52	0.000 11	0.000 09
18new	3.396 ± 4.627	159.61 ± 318.17	1.439 ± 0.028	3.393 ± 4.627	159.50 ± 317.83	1.439 ± 0.028	0.0023	0.0002	0.11	0.34	0.000 14	0.000 02
19	4.587 ± 0.558	17.79 ± 17.97	0.805 ± 0.080	4.582 ± 0.557	17.67 ± 17.91	0.804 ± 0.080	0.0052	0.0007	0.11	0.07	0.000 43	0.000 69
20	4.469 ± 0.662	31.43 ± 25.44	0.819 ± 0.042	4.463 ± 0.661	31.28 ± 25.34	0.819 ± 0.042	0.0059	0.0012	0.15	0.11	0.000 17	0.000 20
21	6.754 ± 6.666	192.53 ± 1420.37	0.886 ± 0.023	6.749 ± 6.654	192.07 ± 1415.32	0.886 ± 0.023	0.0054	0.0115	0.46	5.05	0.000 07	0.000 13
22	6.120 ± 4.761	163.29 ± 1374.82	0.926 ± 0.016	6.116 ± 4.750	163.06 ± 1370.61	0.926 ± 0.016	0.0037	0.0113	0.23	4.21	0.000 06	0.000 03
23	4.723 ± 0.620	25.93 ± 18.60	0.784 ± 0.050	4.718 ± 0.619	25.83 ± 18.54	0.784 ± 0.050	0.0054	0.0008	0.11	0.06	0.000 21	0.000 23
24	4.644 ± 1.180	74.33 ± 57.93	0.786 ± 0.022	4.636 ± 1.176	74.04 ± 57.57	0.786 ± 0.022	0.0083	0.0043	0.29	0.37	0.000 03	0.000 05
25	9.108 ± 5.487	163.00 ± 402.48	0.758 ± 0.012	9.101 ± 5.468	162.73 ± 400.54	0.758 ± 0.012	0.0066	0.0190	0.27	1.94	0.000 05	0.000 02
26	5.538 ± 0.887	49.78 ± 68.76	0.889 ± 0.030	5.531 ± 0.882	49.42 ± 68.13	0.889 ± 0.030	0.0069	0.0042	0.36	0.63	0.000 07	0.000 15
27	6.917 ± 4.745	96.03 ± 71.00	3.424 ± 0.056	6.906 ± 4.760	96.15 ± 71.30	3.424 ± 0.056	0.0116	0.0146	0.12	0.30	0.000 24	0.000 09
27new	6.765 ± 4.456	99.70 ± 70.79	3.263 ± 0.049	6.759 ± 4.463	99.74 ± 70.93	3.263 ± 0.049	0.0062	0.0072	0.04	0.14	0.000 26	0.000 02
28	9.381 ± 7.224	109.38 ± 124.80	3.167 ± 0.051	9.374 ± 7.236	109.45 ± 125.08	3.168 ± 0.051	0.0069	0.0119	0.07	0.28	0.000 20	0.000 07
29	4.429 ± 0.461	1.98 ± 9.11	0.146 ± 3.346	4.425 ± 0.461	1.93 ± 9.09	0.126 ± 3.524	0.0041	0.0002	0.05	0.02	0.019 69	0.178 04
31	8.406 ± 1.141	58.37 ± 80.37	0.875 ± 0.027	8.396 ± 1.132	57.85 ± 79.24	0.875 ± 0.027	0.0095	0.0089	0.52	1.13	0.000 09	0.000 15
33	3.964 ± 4.308	157.66 ± 466.80	1.245 ± 0.021	3.959 ± 4.320	157.81 ± 468.27	1.245 ± 0.021	0.0047	0.0116	0.15	1.47	0.000 10	0.000 06
34	7.635 ± 5.923	157.55 ± 965.92	1.146 ± 0.022	7.628 ± 5.950	157.95 ± 971.85	1.146 ± 0.022	0.0066	0.0267	0.39	5.94	0.000 13	0.000 16

Table D.2.: Comparison of *NOBLE* and a modified version of *PANGA*, incorporating an additional $1/(1 - X_{\mathrm{Xe}})$ factor for the calculation of Xe, as it is used in *NOBLE* (continued on the following page)

	NOBLE			*PANGA*			Deviations					
Sample	T	$A \cdot 10^3$	F	T	$A \cdot 10^3$	F	T	σ_T	$A \cdot 10^3$	$\sigma_A \cdot 10^3$	F	σ_F
	[℃]	[cm³ STP/g]		[℃]	[cm³ STP/g]		[℃]		[cm³ STP/g]			
TB532	10.100 ± 2.044	85.89 ± 90.79	0.780 ± 0.019	10.100 ± 2.044	85.88 ± 90.77	0.780 ± 0.019	0.2310	0.2981	0.011	0.021	0.002	0.001
MW901	5.551 ± 0.633	13.22 ± 7.44	0.553 ± 0.116	5.551 ± 0.633	13.22 ± 7.44	0.553 ± 0.116	0.0469	0.0098	0.001	0.000	0.017	0.014
VZELE	6.047 ± 1.116	55.09 ± 82.38	0.887 ± 0.027	6.048 ± 1.116	55.10 ± 82.40	0.887 ± 0.027	0.1948	0.1740	0.012	0.021	0.001	0.003
GD034	6.242 ± 0.575	5.55 ± 19.18	0.779 ± 0.509	6.242 ± 0.575	5.56 ± 19.18	0.779 ± 0.509	0.0794	0.0099	0.001	0.001	0.027	0.145
GROEDE	10.528 ± 7.685	177.20 ± 1475.80	0.889 ± 0.021	10.527 ± 7.684	177.09 ± 1474.64	0.889 ± 0.021	0.5484	1.0817	0.114	1.156	0.001	0.030
1	7.732 ± 1.188	68.45 ± 37.94	0.724 ± 0.021	7.732 ± 1.188	68.45 ± 37.94	0.724 ± 0.021	0.1567	0.0845	0.004	0.004	0.001	0.001
2	7.425 ± 1.147	66.15 ± 40.81	0.750 ± 0.022	7.425 ± 1.147	66.15 ± 40.81	0.750 ± 0.022	0.0339	0.0303	0.001	0.002	0.001	0.000
3	8.148 ± 0.775	36.94 ± 13.73	0.656 ± 0.031	8.148 ± 0.775	36.94 ± 13.73	0.656 ± 0.031	0.1144	0.0189	0.001	0.000	0.001	0.001
4	8.244 ± 1.197	68.43 ± 34.23	0.700 ± 0.021	8.244 ± 1.197	68.43 ± 34.23	0.700 ± 0.021	0.2363	0.1092	0.006	0.005	0.001	0.001
5	8.988 ± 7.529	184.08 ± 1606.50	0.893 ± 0.022	8.989 ± 7.534	184.26 ± 1609.02	0.893 ± 0.022	0.9989	4.8893	0.187	2.521	0.006	0.062
6	3.243 ± 3.983	331.78 ± 1023.69	0.834 ± 0.029	3.243 ± 3.984	331.73 ± 1023.74	0.834 ± 0.029	0.2683	0.8976	0.055	0.055	0.002	0.002
6new	0.597 ± 0.582	35.51 ± 23.77	0.804 ± 0.035	0.597 ± 0.583	35.52 ± 23.77	0.804 ± 0.035	0.0997	0.0498	0.005	0.005	0.008	0.007
7	1.103 ± 0.424	1.09 ± 18.70	0.239 ± 11.996	1.103 ± 0.424	1.09 ± 18.70	0.238 ± 12.019	0.0542	0.0046	0.001	0.000	0.660	22.969
8	4.639 ± 0.615	26.78 ± 11.93	0.680 ± 0.045	4.639 ± 0.615	26.78 ± 11.93	0.680 ± 0.045	0.0143	0.0057	0.000	0.000	0.000	0.000
9	1.751 ± 0.432	1.09 ± 17.33	0.194 ± 11.750	1.751 ± 0.432	1.09 ± 17.33	0.195 ± 11.736	0.0787	0.0002	0.001	0.000	0.432	14.242
10	4.971 ± 1.162	73.30 ± 86.17	0.854 ± 0.023	4.971 ± 1.163	73.31 ± 86.18	0.854 ± 0.023	0.0382	0.1396	0.010	0.019	0.005	0.001
11	6.302 ± 1.705	93.07 ± 239.11	0.907 ± 0.021	6.302 ± 1.705	93.06 ± 239.05	0.907 ± 0.021	0.0754	0.2312	0.011	0.061	0.001	0.002
13	8.614 ± 1.001	53.16 ± 25.44	0.710 ± 0.025	8.614 ± 1.001	53.16 ± 25.44	0.710 ± 0.025	0.0554	0.0121	0.001	0.001	0.001	0.000
14	9.731 ± 1.163	59.14 ± 38.08	0.757 ± 0.024	9.731 ± 1.163	59.14 ± 38.09	0.757 ± 0.024	0.1391	0.0983	0.006	0.006	0.002	0.002
15	8.893 ± 0.982	30.91 ± 7.84	2.420 ± 0.092	8.893 ± 0.982	30.91 ± 7.84	2.420 ± 0.092	0.1651	0.0078	0.001	0.000	0.006	0.002
16	11.642 ± 7.560	201.04 ± 1354.31	0.867 ± 0.026	11.642 ± 7.563	200.99 ± 1354.50	0.867 ± 0.026	0.1864	2.9014	0.045	0.189	0.000	0.004
17	6.586 ± 4.327	120.08 ± 147.32	1.769 ± 0.027	6.586 ± 4.326	120.06 ± 147.27	1.769 ± 0.027	0.4499	0.8609	0.014	0.047	0.005	0.001
18	4.932 ± 6.233	183.18 ± 586.63	1.344 ± 0.039	4.931 ± 6.233	183.27 ± 586.92	1.344 ± 0.039	0.8004	0.2308	0.082	0.298	0.000	0.035
18new	3.396 ± 4.627	159.61 ± 318.17	1.439 ± 0.028	3.397 ± 4.620	159.45 ± 317.40	1.439 ± 0.027	1.7017	6.5294	0.154	0.770	0.021	0.062
19	4.587 ± 0.558	17.79 ± 17.97	0.805 ± 0.080	4.587 ± 0.558	17.79 ± 17.97	0.805 ± 0.080	0.0370	0.0119	0.001	0.001	0.005	0.005
20	4.469 ± 0.662	31.43 ± 25.44	0.819 ± 0.042	4.469 ± 0.662	31.43 ± 25.45	0.819 ± 0.042	0.0399	0.0119	0.002	0.001	0.003	0.003
21	6.754 ± 6.666	192.53 ± 1420.37	0.886 ± 0.023	6.753 ± 6.660	192.35 ± 1417.78	0.886 ± 0.023	0.8168	6.0624	0.178	2.588	0.001	0.060

Table D.2 (cont.): Comparison of *NOBLE* and a modified version of *PANGA*

Sample	*NOBLE*			*PANGA*			Deviations					
	T	$A \cdot 10^3$	F	T	$A \cdot 10^3$	F	T	σ_T	$A \cdot 10^3$	$\sigma_A \cdot 10^3$	$F \cdot 10^3$	$\sigma_F \cdot 10^3$
	[℃]	[cm³ STP/g]		[℃]	[cm³ STP/g]		[℃]		[cm³ STP/g]			
22	6.120 ± 4.761	163.29 ± 1374.82	0.926 ± 0.016	6.121 ± 4.764	163.34 ± 1375.97	0.926 ± 0.016	0.5139	3.3383	0.051	1.151	0.011	0.007
23	4.723 ± 0.620	25.93 ± 18.60	0.784 ± 0.050	4.723 ± 0.620	25.93 ± 18.59	0.784 ± 0.050	0.0784	0.0059	0.001	0.001	0.001	0.003
24	4.644 ± 1.180	74.33 ± 57.93	0.786 ± 0.022	4.644 ± 1.180	74.33 ± 57.94	0.786 ± 0.022	0.1296	0.1184	0.006	0.008	0.001	0.001
25	9.108 ± 5.487	163.00 ± 402.48	0.758 ± 0.012	9.108 ± 5.487	163.00 ± 402.47	0.758 ± 0.012	0.0285	0.0760	0.002	0.006	0.002	0.000
26	5.538 ± 0.887	49.78 ± 68.76	0.889 ± 0.030	5.538 ± 0.887	49.77 ± 68.76	0.889 ± 0.030	0.1363	0.0242	0.003	0.001	0.003	0.001
27	6.917 ± 4.745	96.03 ± 71.00	3.424 ± 0.056	6.917 ± 4.746	96.04 ± 71.01	3.424 ± 0.056	0.4045	0.2113	0.007	0.009	0.006	0.004
27new	6.765 ± 4.456	99.70 ± 70.79	3.263 ± 0.049	6.767 ± 4.453	99.67 ± 70.72	3.263 ± 0.049	1.8235	2.9414	0.030	0.078	0.012	0.018
28	9.381 ± 7.224	109.38 ± 124.80	3.167 ± 0.051	9.376 ± 7.229	109.45 ± 124.98	3.167 ± 0.051	4.4510	5.1141	0.077	0.183	0.004	0.045
29	4.429 ± 0.461	1.98 ± 9.11	0.146 ± 3.346	4.429 ± 0.461	1.98 ± 9.11	0.146 ± 3.347	0.0324	0.0045	0.000	0.000	0.119	1.161
31	8.406 ± 1.141	58.37 ± 80.37	0.875 ± 0.027	8.405 ± 1.140	58.35 ± 80.33	0.875 ± 0.027	0.2895	0.3110	0.018	0.041	0.005	0.005
33	3.964 ± 4.308	157.66 ± 466.80	1.245 ± 0.021	3.963 ± 4.310	157.69 ± 467.08	1.245 ± 0.021	0.4152	1.7426	0.035	0.285	0.002	0.009
34	7.635 ± 5.923	157.55 ± 965.92	1.146 ± 0.022	7.631 ± 5.944	157.93 ± 971.30	1.146 ± 0.022	3.8246	20.9385	0.379	5.384	0.065	0.134

Table D.3.: Comparison of *NOBLE* and *PANGA* for the UA limit case samples. The modified version of *PANGA* was used, which incorporates an additional $1/(1 - X_{\mathrm{Xe}})$ factor for the calculation of Xe, as it is used in *NOBLE*.

Sample	*NOBLE*			*PANGA*		
	T [℃]	$A \cdot 10^3$ [cm³ STP/g]	F	T [℃]	$A \cdot 10^3$ [cm³ STP/g]	F
7new	0.662 ± 0.771	$1.59 \cdot 10^{-3}$ ± 1053.7	$-242.5 \pm 1.61 \cdot 10^8$	0.663 ± 0.406	$2.90 \cdot 10^{-4}$ ± 34.43	$-1342 \pm 1.59 \cdot 10^8$
9new	1.247 ± 0.495	$2.61 \cdot 10^{-3}$ ± 375.4	$-301.4 \pm 4.36 \cdot 10^7$	1.249 ± 0.411	$3.71 \cdot 10^{-4}$ ± 16.43	$-2130 \pm 9.44 \cdot 10^7$
12	10.391 ± 0.823	$4.29 \cdot 10^{-3}$ ± 161.2	$454.0 \pm 1.70 \cdot 10^7$	10.391 ± 0.515	$4.57 \cdot 10^{-4}$ ± 5.91	$4249 \pm 5.49 \cdot 10^7$
30	4.493 ± 0.642	$3.96 \cdot 10^{-3}$ ± 74.5	$-319.6 \pm 6.03 \cdot 10^6$	4.493 ± 0.450	$1.05 \cdot 10^{-3}$ ± 9.46	$-1207 \pm 1.08 \cdot 10^7$
32	2.637 ± 0.780	$2.72 \cdot 10^{-3}$ ± 433.7	$-300.9 \pm 4.81 \cdot 10^7$	2.638 ± 0.422	$3.71 \cdot 10^{-4}$ ± 15.10	$-2216 \pm 9.02 \cdot 10^7$

List of Abbreviations

asl . . . . . Above sea level

BP . . . . . Before present

CE . . . . . Closed-system equilibration (model)

GR . . . . . Gas diffusion relaxation (model)

LGM . . . . . Last glacial maximum

MAAT . . . . . Mean annual air temperature

MC . . . . . Monte Carlo

NGT . . . . . Noble gas temperature

OD . . . . . Oxygen depletion (model)

PD . . . . . Partial degassing (model)

PR . . . . . Partial re-equilibration (model)

STP . . . . . Standard temperature and pressure

UA . . . . . Unfractionated excess air (model)

List of Figures

List of Tables

Bibliography

Aeschbach-Hertig, W., Beyerle, U., Holocher, J., Peeters, F., and Kipfer, R. (2002a). "Excess air in groundwater as a potential indicator of past environmental changes". In: *Study of Environmental Change Using Isotope Techniques*. Vol. 13/P. Vienna: IAEA, 174–183.

Aeschbach-Hertig, W., El-Gamal, H., Wieser, M., and Palcsu, L. (2008). "Modeling excess air and degassing in groundwater by equilibrium partitioning with a gas phase". In: *Water Resour. Res.* 44, W08449, doi:10.1029/2007WR006454.

Aeschbach-Hertig, W., Peeters, F., Beyerle, U., and Kipfer, R. (1999). "Interpretation of dissolved atmospheric noble gases in natural waters". In: *Water Resources Research* 35.9, 2779–2792.

Aeschbach-Hertig, W., Peeters, F., Beyerle, U., and Kipfer, R. (2000). "Palaeotemperature reconstruction from noble gases in ground water taking into account equilibration with entrapped air". In: *Nature* 405, 1040–1044.

Aeschbach-Hertig, W. and Solomon, D. K. (2013). "Noble gas thermometry in groundwater hydrology". In: *The noble gases as geochemical tracers*. Ed. by P. Burnard. Advances in Isotope Geochemistry. Springer Berlin Heidelberg, 81–122. DOI: 10.1007/978-3-642-28836-4_5.

Aeschbach-Hertig, W., Stute, M., Clark, J. F., Reuter, R. F., and Schlosser, P. (2002b). "A paleotemperature record derived from dissolved noble gases in groundwater of the Aquia Aquifer (Maryland, USA)". In: *Geochimica et Cosmochimica Acta* 66.5, 797–817. ISSN: 0016-7037. DOI: 10.1016/S0016-7037(01)00804-3.

Andrews, J. N. and Lee, D. J. (1979). "Inert gases in groundwater from the Bunter Sandstone of England as indicators of age and palaeoclimatic trends". In: *Journal of Hydrology* 41.3–4, 233–252. ISSN: 0022-1694. DOI: 10.1016/0022-1694(79)90064-7.

Ballentine, C. J. and Hall, C. M. (1999). "Determining paleotemperature and other variables by using an error-weighted, nonlinear inversion of noble gas concentrations in water". In: *Geochim. Cosmochim. Acta* 63.16, 2315–2336.

Bard, Y. (1974). *Nonlinear parameter estimation*. Academic Press New York. ISBN: 978-0-120-78250-5.

Barlow, R. J. (1989). *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. John Wiley & Sons, Inc., Chichester, West Sussex, England, 222. ISBN: 978-0-471-92295-7.

Benítez, J. (1948). *Principles and modern applications of mass transfer operations*. 2nd ed. John Wiley & Sons, Inc., Hoboken, New Jersey. 620 pp. ISBN: 978-0-470-18178-2.

Benson, B. B. and Krause, D. (1976). "Empirical laws for dilute aqueous solutions of nonpolar gases". In: *The Journal of Chemical Physics* 64.2, 689–709. DOI: 10.1063/1.432215.

Benson, B. B. and Krause, D. (1980). "Isotopic Fractionation of Helium During Solution: A Probe for the Liquid State". In: *Journal of Solution Chemistry* 9.12, 895–909. ISSN: 0095-9782. DOI: 10.1007/BF00646402.

Beyerle, U., Aeschbach-Hertig, W., Imboden, D. M., Baur, H., Graf, T., and Kipfer, R. (2000). "A Mass Spectrometric System for the Analysis of Noble Gases and Tritium from Water Samples". In: *Environmental Science & Technology* 34.10, 2042–2050. DOI: 10.1021/es990840h.

Beyerle, U., Purtschert, R., Aeschbach-Hertig, W., Imboden, D. M., Loosli, H. H., Wieler, R., and Kipfer, R. (1998). "Climate and Groundwater Recharge During the Last Glaciation in an Ice-Covered Region". In: *Science* 282.5389, 731–734. DOI: 10.1126/science.282.5389.731.

Beyerle, U., Rueedi, J., Leuenberger, M., Aeschbach-Hertig, W., Peeters, F., Kipfer, R., and Dodo, A. (2003). "Evidence for periods of wetter and cooler climate in the Sahel between 6 and 40 kyr BP derived from groundwater". In: *Geophysical Research Letters* 30.4. ISSN: 1944-8007. DOI: 10.1029/2002GL016310.

Blaser, P. C., Kipfer, R., Loosli, H. H., Walraevens, K., Van Camp, M., and Aeschbach-Hertig, W. (2010). "A 40 ka record of temperature and permafrost conditions in northwestern Europe from noble gases in the Ledo-Paniselian Aquifer (Belgium)". In: *J. Quaternary Sci.* 25.6, 1038–1044.

Brandt, S. (2013). *Datenanalyse für Naturwissenschaftler und Ingenieure*. German. Springer Berlin Heidelberg, 1–6. ISBN: 978-3-642-37663-4. DOI: 10.1007/978-3-642-37664-1_1.

Buizert, C., Baggenstos, D., Jiang, W., Purtschert, R., Petrenko, V. V., Lu, Z.-T., Müller, P., Kuhl, T., Lee, J., Severinghaus, J. P., and Brook, E. J. (2014). "Radiometric ^{81}Kr dating identifies 120,000-year-old ice at Taylor Glacier, Antarctica". In: *Proceedings of the National Academy of Sciences* 111.19, 6876–6881. DOI: 10.1073/pnas.1320329111.

Burnard, P., ed. (2013). *The Noble Gases as Geochemical Tracers*. Advances in Isotope Geochemistry. Springer Berlin Heidelberg. 391 pp.

Castro, M. C., Hall, C. M., Patriarche, D., Goblet, P., and Ellis, B. R. (2007). "A new noble gas paleoclimate record in Texas — Basic assumptions revisited". In: *Earth Planet. Sci. Lett.* 257, 170–187.

Clarke, W., Jenkins, W., and Top, Z. (1976). "Determination of Tritium by Mass Spectrometric Measurement of ^{3}He". In: *The International Journal of Applied Radiation and Isotopes* 27.9, 515–522. ISSN: 0020-708X. DOI: 10.1016/0020-708X(76)90082-X.

Clever, H. L., ed. (1979a). *Helium and neon — gas solubilities*. Vol. 1. Solubility data series. Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt.

Clever, H. L., ed. (1979b). *Krypton, xenon and radon — gas solubilities*. Vol. 2. Solubility data series. Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt.

Clever, H. L., ed. (1980). *Argon*. Vol. 4. Solubility data series. Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt.

Corcho Alvarado, J. A., Purtschert, R., Barbecot, F., Chabault, C., Rueedi, J., Schneider, V., Aeschbach-Hertig, W., Kipfer, R., and Loosli, H. H. (2007). "Constraining the age distribution of highly mixed groundwater using ^{39}Ar: A multiple environmental tracer (^{3}H/^{3}He, ^{85}Kr, ^{39}Ar, and ^{14}C) study in the semiconfined Fontainebleau Sands

Aquifer (France)". In: *Water Resources Research* 43.3. ISSN: 1944-7973. DOI: 10.1029/2006WR005096.

Faybishenko, B. A. (1995). "Hydraulic behavior of quasi-saturated soils in the presence of entrapped air: Laboratory experiments". In: *Water Resour. Res.* 31.10, 2421–2435.

Fayer, M. J. and Hillel, D. (1986). "Air encapsulation: 1. Measurement in a field soil". In: *Soil Sci. Soc. Am. J.* 50, 568–572.

Freundt, F., Schneider, T., and Aeschbach-Hertig, W. (2013). "Response of noble gas partial pressures in soil air to oxygen depletion". In: *Chemical Geology* 339, 283–290. ISSN: 0009-2541. DOI: 10.1016/j.chemgeo.2012.07.026.

Gill, A. (1982). *Atmosphere-Ocean Dynamics*. Vol. 30. International Geophysics Series. Academic Press New York. ISBN: 978-0-122-83522-3.

Hall, C. M., Castro, M. C., Lohmann, K. C., and Ma, L. (2005). "Noble gases and stable isotopes in a shallow aquifer in southern Michigan: Implications for noble gas paleotemperature reconstructions for cool climates". In: *Geophys. Res. Lett.* 32.18. ISSN: 1944-8007. DOI: 10.1029/2005GL023582.

Hall, C. M., Castro, M. C., Lohmann, K. C., and Sun, T. (2012). "Testing the noble gas paleothermometer with a yearlong study of groundwater noble gases in an instrumented monitoring well". In: *Water Resources Research* 48.4. ISSN: 1944-7973. DOI: 10.1029/2011WR010951.

Heaton, T. H. E. and Vogel, J. C. (1981). ""Excess air" in groundwater". In: *J. Hydrol.* 50, 201–216.

Ingram, R. G. S., Hiscock, K. M., and Dennis, P. F. (2007). "Noble gas excess air applied to distinguish groundwater recharge conditions". In: *Environ. Sci. Technol.* 41, 1949–1955.

Jähne, B., Heinz, G., and Dietrich, W. (1987). "Measurement of the diffusion coefficients of sparingly soluble gases in water". In: *Journal of Geophysical Research* 92.C10, 10767–10776.

Jung, M., Wieser, M., von Oehsen, A., and Aeschbach-Hertig, W. (2013). "Properties of the closed-system equilibration model for dissolved noble gases in groundwater". In: *Chemical Geology* 339, 291–300. ISSN: 0009-2541. DOI: 10.1016/j.chemgeo.2012.08.006.

Kipfer, R., Aeschbach-Hertig, W., Peeters, F., and Stute, M. (2002). "Noble gases in lakes and ground waters". In: *Noble gases in geochemistry and cosmochemistry*. Ed. by D. Porcelli, C. Ballentine, and R. Wieler. Vol. 47. Rev. Mineral. Geochem. Washington, DC: Mineralogical Society of America, Geochemical Society, 615–700.

Kreuzer, A. M., von Rohden, C., Friedrich, R., Chen, Z., Shi, J., Hajdas, I., Kipfer, R., and Aeschbach-Hertig, W. (2009). "A record of temperature and monsoon intensity over the past 40 kyr from groundwater in the North China Plain". In: *Chemical Geology* 259.3–4, 168–180. ISSN: 0009-2541. DOI: http://dx.doi.org/10.1016/j.chemgeo.2008.11.001.

Kulongoski, J. T., Hilton, D. R., Izbicki, J. A., and Belitz, K. (2009). "Evidence for prolonged El Nino-like conditions in the Pacific during the Late Pleistocene: A 43 ka noble gas record from California groundwaters". In: *Quaternary Sci. Rev.* 28, 2465–2473.

Kulongoski, J. T., Hilton, D. R., and Selaolo, E. T. (2004). "Climate variability in the Botswana Kalahari from the late Pleistocene to the present day". In: *Geophysical Research Letters* 31.10. ISSN: 1944-8007. DOI: 10.1029/2003GL019238.

Lippmann, J., Stute, M., Torgersen, T., Moser, D., Hall, J., Lin, L., Borcsik, M., Bellamy, R., and Onstott, T. (2003). "Dating ultra-deep mine waters with noble gases and ^{36}Cl, Witwatersrand Basin, South Africa". In: *Geochimica et Cosmochimica Acta* 67.23, 4597–4619. ISSN: 0016-7037. DOI: 10.1016/S0016-7037(03)00414-9.

Ma, L., Castro, M. C., and Hall, C. M. (2004). "A late Pleistocene–Holocene noble gas paleotemperature record in southern Michigan". In: *Geophysical Research Letters* 31.23. ISSN: 1944-8007. DOI: 10.1029/2004GL021766.

Marquardt, D. W. (1963). "An algorithm for least-squares estimation of nonlinear parameters". In: *Journal of the Society for Industrial & Applied Mathematics* 11.2, 431–441. ISSN: 0368-4245. DOI: 10.1137/0111030.

Matsumoto, M. and Nishimura, T. (1998). "Mersenne Twister: A 623-dimensionally Equidistributed Uniform Pseudo-random Number Generator". In: *ACM Trans. Model. Comput. Simul.* 8.1, 3–30. ISSN: 1049-3301. DOI: 10.1145/272991.272995.

Moré, J. J. (1978). "The Levenberg-Marquardt algorithm: Implementation and theory". In: *Numerical Analysis*. Ed. by G. Watson. Vol. 630. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 105–116. ISBN: 978-3-540-08538-6. DOI: 10.1007/BFb0067700.

Nocedal, J. and Wright, S. J. (2006). *Conjugate Gradient Methods*. Springer Series in Operations Research and Financial Engineering. Springer New York. ISBN: 978-0-387-30303-1. DOI: 10.1007/978-0-387-40065-5_5.

Peeters, F., Beyerle, U., Aeschbach-Hertig, W., Holocher, J., Brennwald, M. S., and Kipfer, R. (2003). "Improving noble gas based paleoclimate reconstruction and groundwater dating using ^{20}Ne/^{22}Ne ratios". In: *Geochimica et Cosmochimica Acta* 67.4, 587–600. ISSN: 0016-7037. DOI: 10.1016/S0016-7037(02)00969-9.

Ritterbusch, F., Ebser, S., Welte, J., Reichel, T., Kersting, A., Purtschert, R., Aeschbach-Hertig, W., and Oberthaler, M. K. (2014). "Groundwater dating with Atom Trap Trace Analysis of ^{39}Ar". In: *Geophysical Research Letters*. ISSN: 1944-8007. DOI: 10.1002/2014GL061120.

Schlosser, P., Stute, M., Dörr, H., Sonntag, C., and Münnich, K. O. (1988). "Tritium/^{3}He dating of shallow groundwater". In: *Earth and Planetary Science Letters* 89.3–4, 353–362. ISSN: 0012-821X. DOI: 10.1016/0012-821X(88)90122-7.

Schneider, T. (2014). "Eine Paläoklimastudie an einem Grundwasseraquifersystem in der Nordchinesischen Ebene". PhD thesis. Germany: Heidelberg University.

Seber, G. and Wild, C. (2003). *Nonlinear Regression*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, New Jersey. ISBN: 978-0-471-47135-6.

Smith, S. and Kennedy, B. (1983). "The solubility of noble gases in water and in NaCl brine". In: *Geochimica et Cosmochimica Acta* 47.3, 503–515. ISSN: 0016-7037. DOI: 10.1016/0016-7037(83)90273-9.

Solomon, D. K. and Cook, P. G. (2000). "^{3}H and ^{3}He". In: *Environmental Tracers in Subsurface Hydrology*. Ed. by P. G. Cook and A. L. Herczeg. Springer US, 397–424. ISBN: 978-1-4613-7057-4. DOI: 10.1007/978-1-4615-4557-6_13.

Sturchio, N. C., Du, X., Purtschert, R., Lehmann, B. E., Sultan, M., Patterson, L. J., Lu, Z.-T., Müller, P., Bigler, T., Bailey, K., O'Connor, T. P., Young, L., Lorenzo, R., Becker, R., El Alfy, Z., El Kaliouby, B., Dawood, Y., and Abdallah, A. M. A. (2004). "One million year old groundwater in the Sahara revealed by krypton-81 and chlorine-36". In: *Geophysical Research Letters* 31.5. ISSN: 1944-8007. DOI: 10.1029/2003GL019234.

Stute, M., Forster, M., Frischkorn, H., Serejo, A., Clark, J. F., Schlosser, P., Broecker, W. S., and Bonani, G. (1995). "Cooling of tropical Brazil (5 ℃) during the Last Glacial Maximum". In: *Science* 269, 379–383.

Stute, M. and Schlosser, P. (1993). "Principles and applications of the noble gas paleothermometer". In: *Climate Change in Continental Isotopic Records*. Ed. by P. K. Swart, K. C. Lohmann, J. McKenzie, and S. Savin. Vol. 78. AGU Geophysical Monograph Series. Washington, DC: American Geophysical Union, 89–100.

Sun, T., Hall, C. M., and Castro, M. C. (2010). "Statistical properties of groundwater noble gas paleoclimate models: Are they robust and unbiased estimators?" In: *Geochem. Geophys. Geosyst.* 11.2, Q02002. DOI: 10.1029/2009GC002717.

Sun, T., Hall, C. M., Castro, M. C., Lohmann, K. C., and Goblet, P. (2008). "Excess air in the noble gas groundwater paleothermometer: A new model based on diffusion in the gas phase". In: *Geophysical Research Letters* 35.19. ISSN: 1944-8007. DOI: 10.1029/2008GL035018.

Von Oehsen, A. (2008). "Parameter Estimation and Model Validation for Models of dissolved Noble Gas Concentrations in Groundwater". Diploma thesis. Heidelberg, Germany: Heidelberg University.

Weiss, R. F. (1970). "The solubility of nitrogen, oxygen and argon in water and seawater". In: *Deep Sea Research and Oceanographic Abstracts* 17.4, 721–735.

Weiss, R. F. (1971). "Solubility of Helium and Neon in Water and Seawater". In: *Journal of Chemical and Engineering Data* 16.2, 235–241.

Weiss, R. F. and Kyser, T. K. (1978). "Solubility of Krypton in Water and Seawater". In: *Journal of Chemical and Engineering Data* 23.1, 69–72.

Weyhenmeyer, C. E., Burns, S. J., Waber, H. N., Aeschbach-Hertig, W., Kipfer, R., Loosli, H. H., and Matter, A. (2000). "Cool glacial temperatures and changes in moisture source recorded in Oman groundwaters". In: *Science* 287.5454, 842–845.

Wieser, M. (2011). "Imprints of climatic and environmental change in a regional aquifer system in an arid part of India using noble gases and other environmental tracers". PhD thesis. Germany: Heidelberg University.

Zartman, R., Wasserburg, G., and Reynolds, J. (1961). "Helium, argon, and carbon in some natural gases". In: *Journal of Geophysical Research* 66.1, 277–306.

Acknowledgments

Zum Schluss möchte ich mich bei all jenen bedanken, die durch ihre direkte oder indirekte Hilfe zur Entstehung dieser Arbeit beigetragen haben.

Mein besonderer Dank gilt meinem Betreuer Prof. Dr. Werner Aeschbach-Hertig für das Ermöglichen dieser interessanten Arbeit, die mir zahlreiche Möglichkeiten gegeben hat, mich persönlich und professionell weiterzuentwickeln. Trotz deines meist vollen Terminkalenders hast du immer Zeit gefunden, mir mit hilfreichen Ratschlägen zur Seite zu stehen.

Prof. Dr. Kurt Roth danke ich für die Übernahme des Zweitgutachtens.

Tim Schneider möchte ich danken für die Zeit während der beiden Messkampagnen in China. Auch wenn die Aufenthalte oft geprägt waren von langer Warterei und dem Totschlagen von Zeit, waren sie insgesamt doch eine Erfahrung, die ich nicht missen möchte.

Tillmann Kaudse möchte ich danken für die zahlreichen Bugreports und hilfreichen Beiträge zur Weiterentwicklung von *PANGA* aber auch für die vielen interessanten Gespräche.

Mein Dank geht auch an Florian Freundt fürs Korrekturlesen dieser Arbeit und auch für Tipps und Inspirationen zum Thema Fotografie.

Und auch der restlichen Arbeitsgruppe möchte ich meinen Dank aussprechen für die gute Zeit in den letzten Jahren, die Diskussionen beim Mittagessen und in Büro und Labor, die Grillabende (wenn ich denn mal dabei war), die Gruppenausflüge und -feiern.

Außerdem danke ich Christoph, Sebastian, Tilman, Johanna und Marcel, sowie den anderen, die sich mit mir im Laufe der letzten Jahre das Büro geteilt haben, für die gute Atmosphäre und die zahlreichen Diskussionen über Gott und die Welt, Politik und Linux. Mit euch ist es nie langweilig geworden!

Meiner Frau Anne möchte ich von ganzem Herzen danken! Du hast mich immer unterstützt und aufgemuntert wenn es mal nicht so gut lief. Danke auch dafür, dass du mir immer den Rücken freigehalten hast und für deine Geduld, besonders jetzt, in den letzten Wochen dieser Arbeit.

Zuletzt möchte ich meiner Familie danken, für all die Unterstützung und die guten Ratschläge, die ihr mir immer gegeben habt. Ohne euch wäre ich heute nicht da wo ich bin!