Stable Isotope Fractionation in Antarctic Precipitation

Theory of Isotope Thermodynamics, Parameterization of Empirical Models, and Insights from New Data

Parker Liautaud

Submitted as an Essay (G&G 492) in partial fulfillment of the degree of Bachelor of Science in Geology & Geophysics

Yale University, New Haven, CT
May 2016

Abstract

Shallow cores were extracted at 10 sites across Antarctica, and tested for their $\delta^{18}O$ and $\delta^{2}H$ values. From this information, and an aggregate dataset of samples from 1,280 sites across Antarctica (the Masson Data), I present an analysis of stable isotope effects in precipitation across Antarctica. I find that there is a statistical bias in using distance-from-the-coast as an indicator of isotopes in precipitation; that elevation decreases on land are not accounted for in existing statistical models for Antarctica; and that isotope-temperature relationships overestimate observed temperatures near the South Pole while underestimating them closer to the coast. I also propose an Antarctic Meteoric Water Line and new parameters for predicting the composition of Antarctic near-surface snow. The discussion also includes a summary of the theory underlying stable isotope fractionation.

Contents

1 Acknowledgments

1.1 Funding & Support

This study is made possible thanks to funding generously provided by:

- Willis Towers Watson (formerly Willis Group)
- EMC Corporation
- One Young World
- Agility Logistics
- Algebris Investments
- Amcor
- Initiative Supply
- Analog Devices

As well as in-kind and analysis resources provided by the Institute of Geological and Nuclear Sciences (GNS) of New Zealand, and Instituto Antartico Chileno (INACH). I also thank the National Science Foundation (NSF), the U.S. Environmental Protection Agency (EPA), and the U.S. Department of State for granting the necessary permissions and permits.

Interests are listed at the end of the paper.

1.2 Research Advisory Committee

The following individuals were part of a committee that helped to design the research objectives of the Willis Resilience Expedition, whose data are partially reported in this paper.

- Dr. Uwe Morgenstern, Team Leader, Isotope Hydrology & Water Dating Lab, GNS New Zealand

- Dr. Nancy Bertler, Associate Professor, Antarctic Research Centre, Victoria University of Wellington
- Mr. Rowan Douglas CBE, Chairman, Willis Research Network, Member of Natural Environment Research Council (NERC) and the UK Prime Minister's Council on Science & Technology.

2 Introduction

For decades, researchers have used the ratios of atomic weights in stable isotopes found in water to record environmental change. The development of these techniques was a major factor in scientists' abilities to precisely reconstruct climates into the very distant past. Early influencers such as Willi Dansgaard helped establish critical theory in fractionation processes and other paradigms that have formed the bedrock for stable isotope analysis in natural systems.

When scientists began to have reliable access to the Antarctic continent in the mid-1900s, and after the initial theory on using stable isotope ratios in paleothermometry, they turned their attention to the potential for reconstructing climate change through ice cores. Now, cores from Antarctica record climatic shifts to hundreds of thousands of years in the past.

However, while large global shifts can be recorded through these techniques fairly easily, the short-term usefulness of stable isotope data remains hindered by the difficulty of accessing the continent, even though analyzing water isotopes across this unique continent has significant potential in helping us improve our understanding of how fractionation processes change in extreme conditions, and in more precisely mapping water isotopes across the world.

Antarctica remains exceedingly difficult to access for research purposes. Antarctica's research bases are widely dispersed. Activities require delicate diplomacy, as Antarctica is patched with contested borders and claims over valuable wedges of land (the result of the vaguely worded Antarctic Treaty System). Even those expeditions that do manage to take place are always subjected to the harsh Antarctic climate and frustratingly little control over the quality of the data, especially when day-to-day weather conditions often dictate the progress of a program.

When researchers do manage to retrieve them, stable isotope records from cold regions are valuable. Studying the changes in water isotope ratios over time allows us to reconstruct temperature and other characteristics of the climate, hundreds of thousands of years in the past. Not subjected to the biological and other degradation processes of temperate regions, ice records are unique in their preservation, with many ice cores preserving bubbles of atmospheric gas [2].

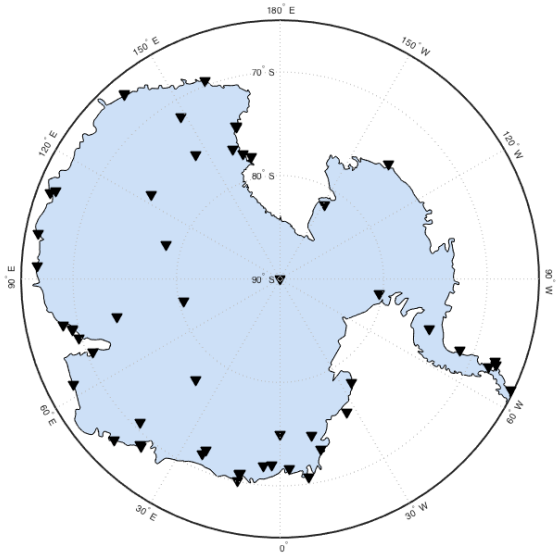

Figure 1: Locations of Research Stations & Bases in Antarctica [1]

Increasingly sophisticated tools have allowed researchers to simulate the meteorological conditions over Antarctica. These models have helped us gain an increasingly robust understanding of isotope fractionation processes that take place in a parcel of moist air along a vector toward precipitation. For example, the 3-stage Mixed Cloud Isotopic Model (MCIM), developed by Ciais and Jouzel (1994) [3], used mathematical relationships grounded in the Rayleigh distillation paradigm to account for mixed-phase clouds when modeling effect of phase changes in clouds over cold regions.

Significant progress has been made in attempting to build precise predictive models. The MCIM was itself an improvement over the models developed by Jouzel and Merlivat (1984) [4]/ Merlivat and Jouzel (1979), the "RMK" Model [5], which only modeled two stages (before and after cooling beyond a pre-determined temperature threshold) and two phases before precipitation (vapor and liquid). The 'Jouzel Set' of models and the expansive literature from well-known atmospheric scientists (Petit, Jouzel, Dahe, Fisher, Masson-Delmotte, and others) have contributed to improvements in our understanding of the dynamics of present-day Antarctic climate, with significant consequences for modeling into the near and distant past.

Atmospheric Global Circulation Models (AGCMs) such as the University of Utrechts Regional Atmospheric Climate Model (RACMO) are also used to model the characteristics of precipitation and moisture in three dimensions above the Antarctic boundary [6].

Models like the MCIM require initial assumptions about source temperature and isotopic composition, air parcel characteristics such as relative humidity and pressure and phase changes, and the moisture trajectory and speed at which the path is followed [7]. Masson-Delmotte (2008) showed that the assumptions required for input into the MCIM show divergent results from some of the more modern models, but also that the relationships inferable from empirical evidence (directly measured snow and ice samples) have different structure from the model simulations as well (but structure nonetheless) [7]. One example would be the consistent overestimation (less negative than the real data) of near-surface δD values from accumulation by three AGCMs: the Melbourne University GCM (MUGCM), NASAs GISS-E, and ECHAM4 [7].

Such differences reflect a crucial component of the study of isotope geochemistry in the context of scarce data, which Antarctica embodies perhaps more than any other region on earth. The scientific community has developed a robust understanding of isotope fractionation in the hydrological cycle at high latitudes, but in studying Antarctica, finds itself constrained by both limited and geographically skewed data sources. Such sources can also be unavoidably noisy in a place as meteorologically chaotic as Antarctica.

Nonetheless, these differences are cause for reflection. Two paradigms seem to have emerged in the literature on the present-day and recent-past composition of Antarctic precipitation. One focuses on the output of AGCMs and the ability to produce detailed three-dimensional mappings of isotopic composition, temperature, accumulation, and other variables at any point on the earth and in an atmospheric column. These models generate results in eye-opening detail, but come at the expense of real-world precision. Both the model outputs themselves and any accusation of imprecision are difficult to verify given the lack of ground truth samples from the continent. Sophisticated AGCMs therefore have the potential to distract us if their use cannot be reconciled with real-world data.

If we think of GCM-centric analysis as a top-down approach, the bottom-up one is rooted in what little data are available. Analysis from approxi-

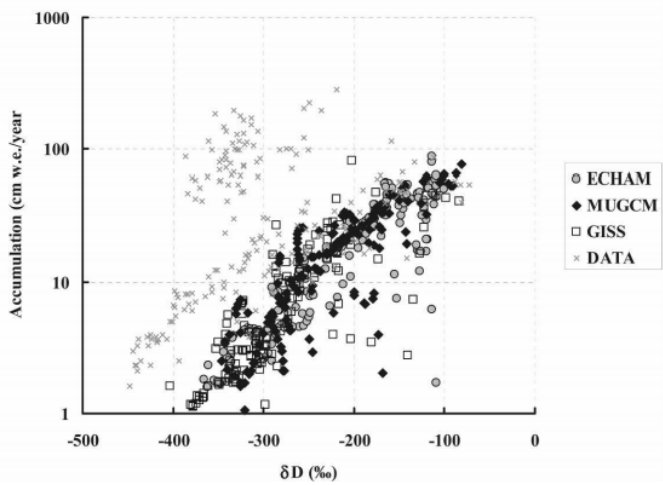

Figure 2: δ D vs. Accumulation rate, AGCM Simulations vs. Antarctic Data [7]

mately 1,280 sites in the last half-century are available through a dataset compiled by Valrie Masson-Delmotte et al [7]. It is possible to construct regression models to explain much of the variance in D and 18O, but traditional OLS approaches have their drawbacks, including an inability to account for geophysical processes which can cause structural deformation of the dataset that cant be explained by a model linear in only a few parameters (in other words, atmospheric processes that impact all observations in a dataset, rather than just throwing off a few individual points). This is especially the case because the variables included in these models are geographic or climatological. Furthermore, this approach has the serious drawback of producing models where the link between correlation and causation is difficult to draw, and functional form in regression relationships could be considered entirely subjective [8].

This results in difficult questions around what 'should' be observed given a certain set of initial conditions and what observations end up showing, in the event there is a discrepancy between the two. How do we reconcile two approaches when neither are perfect?

This leads us to the central question of this paper:

How can we best harness empirical evidence to predict the stable-isotope composition of Antarctic precipitation?

There are many reasons to model near-surface isotopic composition after deposition. First, its easy to forget that as sophisticated as they are, simulated results are hypothetical until proven, and any rift between the ground data and the models needs to be explained. Second, simulations cant account for every micro-level occurrences and disturbances that do have an impact on the fractionation path.

While it's difficult to model water isotopes on a large scale, without computer simulations, attempting to do so allows us to get the most out of empirical evidence, and allows us to find explanations for variations not found in simulated results.

However, questions remain about how we can best describe and project the relationship between the stable-isotopic composition of Antarctic precipitation and the factors affecting it. Challenges include deciphering the best functional form in regression analysis, the effect of simplifying relationships between variables on the accuracy of prediction, and the level of precision to which we should (and can reasonably) ultimately aspire, especially with severely limited access to the continent and geographically skewed samples. Despite these questions, there exists large quantities of data from nearly sixty years of survey expeditions, and the opportunity to harness this data to better understand this extraordinary and remote continent is exciting.

Access to these data also allow for the opportunity to assess the precision of continent-wide predictive models on a micro-scale, and potentially improve the models to reflect the heterogeneity of Antarctic geographic features. This is especially the case where surveys have included closely spaced sampling locations, as was the case with the 2006 Gooseff et al. study in the McMurdo Dry Valleys [7].

This paper therefore also aims to use statistical theory and real analysis to (1) assess the effectiveness of existing models describing the variability of stable isotope composition of precipitation in Antarctica, (2) to propose adjustments, and (3) to derive tools for quantifying the effect of using datasets that are highly constrained in quality and representativeness. Therefore, I will analyze published data from all reported studies of near-surface stable

isotope composition in Antarctica up to 2007 [7].

Lastly, this paper discusses for the first time the results a 2013 research expedition across Antarctica which I created and led, during which samples were taken at ten sites for analysis of stable isotope composition and Tritium composition.

3 Stable Isotope Fractionation in Water

3.1 Fractionation Theory

In the natural systems, the two elements in water occur on average with the abundances listed in Table 1. Both Hydrogen and Oxygen have one overwhelmingly abundant isotope, O^{16} and H^1, and two much rarer ones. Natural processes (whether in the atmosphere, in biological systems, or otherwise) can cause small deviations from these averages as a result of the quantum-scale properties of atoms, and the differences in thermodynamic behavior between atoms of differing masses. A brief summary of the theory behind the fractionation processes relevant to Antarctica is useful, as it will inform the analysis of both new and existing data collected from the continent.

On a fundamental level, mass-dependent isotope effects are dictated by the laws of quantum physics, which state that particles can only take discrete energy levels [11], which within an atom or molecule sum to its energy state [11]. These energy states for the basis for diverging behaviors of the Oxygen and Hydrogen atoms of different masses at the macro level.

As it relates to the hydrological system and our study of Antarctic stable isotopes, mass-dependent isotope fractionation is caused by three processes [12]:

- Thermodynamic effects: effects from equilibria found in natural systems, including from phase changes due to temperature increases/decreases.
- Kinetic processes: isotope effects in chemical reactions which go to completion.
- Large-scale mobility: fractionation occurring during the movement of air parcels laterally and vertically, movement over or against topography, continentality, and others.

Together, these three categories make up the fractionation processes experienced by a quantity of water as it moves through the hydrological cycle. Mass-independent phenomena do exist [12], and are relevant to techniques

Table 1: Relative Abundance of Water Isotopes

Oxygen			Hydrogen		
O^{16}	O^{16}	O^{17}	H^1	H^2	H^3
99.76%	0.201%	0.039%	99.9885%	0.0115%	∼0

for studying the climate through, for example the use of clumped isotopes (which involves the abundance ofO^{17} [12]), but will not be discussed in this paper, as they are not relevant to its main analysis.

The vast majority of the variation we see in Antarctica are the result of a combination of equilibrium and large-scale moisture transport effects. Kinetic effects are mostly the domain of biological systems, which are not relevant to Antarctic precipitation (not to mention that the Antarctic surface is notably devoid of life!), so this discussion will be limited to equilibrium and transport effects.

Thermodynamic and Mechanical Basis for Isotope Effects Potential energy governs the behavior of individual atoms and therefore the fractionation of different isotopic varieties in the aggregate (on the scale we see in Antarctic precipitation) [13]. Mass differences between isotopes result in several important energy-related behavioral differences which become the basis for equilibrium and transport fractionation.

The kinetic energy of a molecule is a simple function of temperature, scaled by the Boltzmann constant (k) [11]:

$$kT = \frac{1}{2}mv^2 \tag{1}$$

Where v is the average molecular velocity, k is approximately $1.38 \times 10^{-23} m^2 kg s^{-2} K^{-1}$, m is the molecular mass, and T is absolute temperature in Kelvin. Critically, as the IAEA points out [11], $k \times T$ is determined entirely externally; it is a constant multiplied by the environmental temperature. This means that a molecule with a larger mass (a heavier isotopic composition) must, by the fundamental laws of physics, have a lower velocity. This becomes relevant during (1) phase changes, (2) interactions between identical molecules and with different substances, and (2) chemical

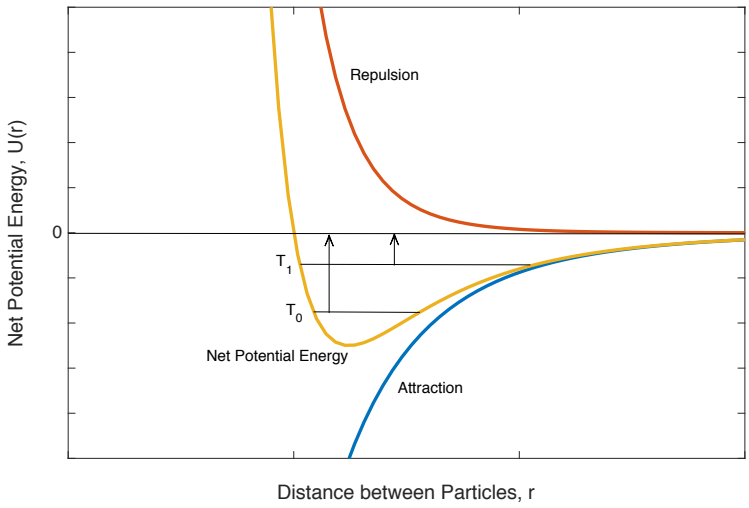

Figure 3: Lennard-Jones Potential [17]

reactions.

Molecules with different masses experience different interatomic potentials [11]. The potential energy between molecules is governed by relationships between repulsive and attractive forces, which change with isotopic composition of the molecule. In a sample of pure water, at very short distances, overlapping electronic orbitals cause a strong repulsion [15]. At larger distances, the repulsive forces decrease sharply and attractive forces dominate specifically van der Waals forces, which are temporary molecular dipoles averaging out to a permanent, significant effect (depending on the molecule) [16].

The Lennard-Jones Potential maps the net potential energy between two molecules [17]:

$$U(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{2}$$

Where the r is the distance between atoms, U is net potential energy, and ε and σ are parameters specific to the atoms in question, which are computed to be $\sigma = 3.02\mathring{A}$ and $\varepsilon = 21.7cm^{-1}$ for H_2O [68].

This relationship governs the particle's binding energy, which is the energy required to overcome the net attractive forces in the negative portion of the curve [17]. As it relates to the isotopic composition of precipitation, which is effectively the subject of this paper, the Lennard-Jones equation equally applies to phase changes, such as evaporation [17]. It can be seen from the figure that at higher temperatures, the energy gap that must be overcome is smaller.

This phenomenon leads to a differentiated isotope effect. Heavier isotopes require more energy to overcome the bonds between molecules and, for example, evaporate less easily but condensate and precipitate more easily than lighter isotopes [11]. This is one of the effects that on a macro-scale causes the moisture in a cloud to become progressively lighter as it rains out.

Equilibrium Isotope Fractionation Many reversible processes experience isotopic fractionation; this is especially the case for phase changes in water. Principles of statistical mechanics tell us that we can compute the equilibrium constant of an isotope exchange reaction by mapping the probability of a molecule being a particular energy state[18] [21]:

$$P_r = \frac{n_r}{n_0} = \frac{e^{-E_r/kT}}{q} \tag{3}$$

Where E_r is the energy level, k is the Boltzmann Constant, T is thermodynamic temperature, n_r/n_o is the ratio of the number of molecules with energy E_r to the number of molecules with zero-point energy.

From this we derive the Boltzmann Distribution Function, which gives the average energy per molecule [18] [21]:

$$\frac{\sum_{r=0}^{\infty} n_r E_r}{\sum_{r=0}^{\infty} n_r} = \frac{\sum_{r=0}^{\infty} E_r e^{-E_r/kT}}{q} \tag{4}$$

We can see from this function that:

$$q = \frac{\sum_{r=0}^{\infty} E_r e^{-E_r/kT}}{E_r}$$

Thus:

$$q = \sum_{r=0}^{\infty} e^{-E_r/kT} \tag{5}$$

q gives us the partition function, which sums the energy levels of the components of the molecule to gives us the overall molecule's energy state [19]. This function is the basis for the derivation of K, the equilibrium constant, whose derivation is critically important to the theory of stable isotope fractionation.

In isotope fractionation processes, what we really seek to know is how the various isotopic compositions of molecules affect the *"mobility"* of the molecule, such that they behave differently in the same conditions. The partition function q thus can be used to describe the different types of movement undertaken by the molecule, of which there are three types: vibrational, rotational, and translational. Thus the net energy system can be described by the combined partition function for each of these types of motion, given by IAEA [21]:

$$Q = q_{trs} q_{rot} q_{vib} \tag{6}$$

We are fortunate in that H_2O is a simple diatomic molecule, which makes it significantly easier to approximate q, since external impacts are negligible (since water is not bound by a macro-structure or lattice, its movement and rotation is not affected by (and does not affect) the translation, vibration, or rotation of other water molecules) [21].

Without deriving them, the partition functions for each type of motion are given by [21]:

$$q_{trs} = \left[\frac{2\pi mkT}{h^2}\right]^{\frac{3}{2}} V$$

$$q_{rot} = \frac{8\pi^2 \mu r_0^2 kT}{s \cdot h^2}$$

$$q_{vib} = \frac{e^{-h \cdot \nu / 2kT}}{1 - e^{h \cdot \nu / 2kT}}$$

Where V is the volume of the ambient space, m is molecular mass, h is Planck's constant ($6.626 \times 10^{-34} m^2 kg s^{-1}$) [20], $\mu = (m_1 m_2)/(m_1 + m_2)$, μr_0^2 is the moment of inertia, $s = 2$ for water, ν is the frequency of vibration for a water molecule, T is temperature.

It is already apparent from these relationships that the mobility is partially dependent on Temperature, which is the root cause of some of the meteorological effects on stable isotopic composition of precipitation. ν is equal to [69]:

$$\nu = \frac{1}{2/\pi} \cdot \left[\frac{c}{\mu}\right]^{1/2} \tag{7}$$

Where c is a force constant. This shows how the mass of the molecule (i.e. the isotope), represented by μ, affects the vibrational frequency, and thus affects q_{vib}, and thus affects overall Q.

The goal is to decipher the effect on mobility of an isotopic substitution, e.g. $H_2O^{18} \rightarrow H_2O^{16}$. The changes in translational and rotational q are temperature-independent [11], but critically are *mass-dependent*.

In vibrational mobility, the force constant doesn't normally change with an isotopic substitution [11], which means that the substituted Q and original Q are related directly by the change in ν [11].

$$\frac{q_{vib}^{(s)}}{q_{vib}} = e^{h(\nu - \nu^{(s)})/2kT} = \exp\left[\frac{h \cdot \nu}{2kT}\left(1 - \frac{\mu}{\mu^{(s)}}^{0.5}\right)\right]$$

Substituting in IAEA formulations for the mass-dependent translational and rotational effects in isotopic substitution, the total change in Q is [21]:

$$\frac{q_{vib}^{(s)}}{q_{vib}} = \left[\frac{M^{(s)}}{M}\right]^{0.5} \cdot \frac{m^{(s)}}{m} \cdot \exp\left[\frac{h \cdot \nu}{2kT}\left(1 - \frac{\mu}{\mu^{(s)}}^{0.5}\right)\right] \tag{8}$$

Thermodynamic Relationship between Q and K The partition function Q is connected to K via the relationship between energy and entropy, and the change in Gibbs Free Energy associated with equilibrium reactions [18].

The assumptions of the canonical ensemble of statistical mechanics [22] tell us that entropy, S, is a function of the number of possible configurations of a system (essentially, its phase) [23]:

$$S = a \cdot \ln(\Omega)$$

Where a is a constant. This is the basis for temperature-dependent Helmholtz Free Energy [23], which is a function of the partition function Q, derived in the previous section:

$$A = -a \cdot T \cdot \ln(Q) \tag{9}$$

Since $A = U - TS$ [22], where U is internal energy and S is entropy, this gives us: $U - TS = -a \cdot T \cdot \ln(Q)$, where aT is the gas constant R, $8.314 Jmol^{-1}K^{-1}$[70] [71]. Thus:

$$\begin{aligned} \Delta U - T\Delta S &= -R\ln[Q] \\ \equiv \Delta U - T\Delta S &= -R\ln[q_{trs}^{\gamma_1} q_{rot}^{\gamma_2} q_{vib}^{\gamma_3}] \\ \equiv \Delta U - T\Delta S &= -R\ln\left[\prod q_n^{\gamma}\right] \end{aligned}$$

Where γ refers to a stoichiometric adjustment [21], the sign of which on whether the molecules in question are reactants or products [18]. We know that $\Delta U - T\Delta S$ is equivalent to Gibbs Free Energy, which is given by $\Delta G = -RT\ln(K)$, where K is the equilibrium constant. Thus [18]:

$$
\begin{aligned}
-RT\ln(K) &= -R\ln\left[\prod q_n^{\gamma}\right] \\
&\equiv K\cdot\exp\left(\Delta T\right) = \left(\prod q_n^{\gamma}\right)
\end{aligned}
$$

However, the equilibrium condition is given by [23]:

$$
dA = dT\left(\frac{\partial A}{\partial T}\right)_{V,n_i} + dV\left(\frac{\partial A}{\partial V}\right)_{T,n_i} + \sum_i dn_i\left(\frac{\partial A}{\partial n_i}\right)_{T,V,n_j} = 0
$$

Which means that the change in free energy is dependent on temperature. For the second and third terms, T is held constant; so $dT\left(\frac{\partial A}{\partial T}\right)_{V,n_i}$ is 0 if the sum of the second and third terms is zero. At equilibrium, dT (i.e. marginal ΔT) should be equal to zero since the forward and backward reactions are at proceeding at equal rates, and therefore the temperature should stabilize. Thus $dT = 0$, so the sum of the second and third terms should equal zero. The same logic applies to the change in volume dV, as well as $\sum_i dn_i$. This implies that at equilibrium the temperature term can be dropped, leaving us with [18]:

$$
K_{eq} = \prod q_n^{\gamma} \tag{10}
$$

The relationship between K and Q is fundamental to stable isotope geochemistry because Q is **mass-dependent**. Thus, the equilibrium constant is a function of the mass of the molecule, and equilibria found in natural systems produce natural stable isotope fractionation with a change in external conditions.

Pressure and temperature are two of these conditions, as Q and K are tied by the ideal gas law. This explains why different isotopes of Hydrogen and Oxygen respond differently to the same conditions in the hydrological cycle, and why the movement of a parcel of moist air vertically in the atmosphere and laterally toward or away from the poles causes a fractionation effect.

Rayleigh Processes We are interesting in determining how (1) phase changes and (2) movement of air affects the stable isotopic composition of the precipitation that reaches Antarctica. To study this, we rely on the Rayleigh framework, which describes how isotopes become distributed between two reservoirs as one gets smaller [24]. The Rayleigh equations can describe fractionation between mixed phases and thus makes it ideal for studying the atmosphere, especially in movements of air masses into regions [24].

Definitions From our previous derivation of K from the mass-dependent partition function, we can say that for phase changes in the atmosphere [12]:

$$K_X(T) = \frac{R_{P1}}{R_{P2}} = \alpha_{1/2}(T) \tag{11}$$

This is the standard description used by the IAEA in the Global Network of Isotopes in Precipitation and by scientists worldwide [12]. Note that R is no longer the gas constant, but in the Rayleigh framework refers to the ratio of the abundance of the rare to the more abundant isotope.

α is the fractionation factor, which itself is not used in analysis, because isotope effects are so small that *alpha* is usually incredibly close to 1. Therefore, we define:

$$\varepsilon = \alpha - 1$$

And the isotope effect is ε expressed in ‰ [12].

We are again fortunate in that we are studying reservoirs and effects with just a single compound, water, so that the isotopic composition of a reservoir is given simply by $R_r = N_i/N$ [12], and where the average isotopic composition of the molecules being removed from the reservoir (which is different from that of the reservoir itself) is given by $R_e = dN_i/dN$. This implies a fractionation factor $\alpha_{e/r}$ of [12]:

$$\alpha_{e/r} = \frac{R_e}{R_r} = \frac{dN_i}{dN} \Big/ \frac{N_i}{N}$$

Where under the assumptions of Rayleigh processes, $\alpha_{e/r}$ remains constant [12]. Algebraic manipulation gives us: $d\ln R/d\ln f = (\alpha_{e/r} - 1)$, where f is the proportion not yet removed, such that a form for Rayleigh fractionation in the general IAEA framework is given by [12]:

$$R = R_0 f^{\alpha_{e/r}-1} \tag{12}$$

Lastly, isotope ratios are expressed in δ notation, relative to a standard [12]:

$$\delta^{18}O = \left(\frac{\frac{O^{18}}{O^{16}}sample}{\frac{O^{18}}{O^{16}}standard} - 1 \right) \times 1000 \tag{13}$$

All fractionation is measured and reported relative to Vienna Standard Mean Ocean Water, a standard maintained by the IAEA, which was taken from distilled seawater. VSMOW has H and O values of $\frac{\delta^{18}O}{\delta^{16}O} = (2005.45 \pm 0.45) \times 10^{-6}$ and $\frac{\delta^{2}H}{\delta^{1}H} = (155.76 \pm 0.05) \times 10^{-6}$ [72]. The δ value for VSMOW is therefore $\delta \sim 1 - 1 = 0$ [12].

This is all the theory and notation that is needed to analyze stable isotope effects in precipitation over Antarctica.

3.2 Earth System Influencers

As previously discussed, the mass-dependent stable isotope effects are caused by fundamental properties of the molecule at a quantum level, but these manifest themselves in measurable ways across large distances in Earth's atmosphere and in natural systems.

An air parcel moving around the globe is constantly subjected to changing conditions. Mostly, these conditions are functions of changes in two dimensions: latitude and elevation. Both of these have impacts on the characteristics of the air, especially temperature. Earth system effects on stable isotopic composition of atmospheric water vapor can be traced back to the molecular level, with the mobility of rotation, translation, and vibration all tied to temperature and a set of constants (since we are only assessing one

type of molecule, T is the only variable in the equilibrium constant for each possible molecular mass of water).

The combined effects of factors like continentality and latitude eventually determine the δ values we see in samples, so a brief discussion of these effects is included below.

The δ values for H and O in a water sample depend to a large extent on the isotopic composition of the liquid source from which it originates. This is because, by definition, the stable-isotopic composition of the entire hydrological system is measured relative to VSMOW.

One of the largest sources of global precipitation is evaporation near the equator [12]. As one moves away from the equator towards the poles, average temperatures decrease, resulting in condensation and rainout, which causes more negative values.

The ocean is well-mixed, and therefore, δ values for ocean surface water don't vary too widely with latitude. However, the latitude effect at the surface is nonetheless not insignificant, which is why we can't use VSMOW as the basis for any marine source, but an additional fractionation must be computed in to take into consideration the source of atmospheric moisture.

Temperature On a global scale, the weighted average of the relationship between temperature and stable isotope composition of water follows the fitted relationship reported by the IAEA [12]:

$$\delta^{18}O = (0.521 \pm 0.014)T - (14.96 \pm 0.21)‰ \tag{14}$$

With T representing ground-level temperature. This effect is the combined result of location-specific effects, which are discussed below.

Latitude The IAEA reports a different latitude effect for different regions, with temperature regions seeing on average more significant with distance from the equator than cold regions [12]: $\Delta^{18}O \approx -0.6‰/\deg$ (degree of latitude, not temperature) for North America and Europe, compared to $\Delta^{18}O \approx -0.2‰/\deg$ for Antarctica [12].

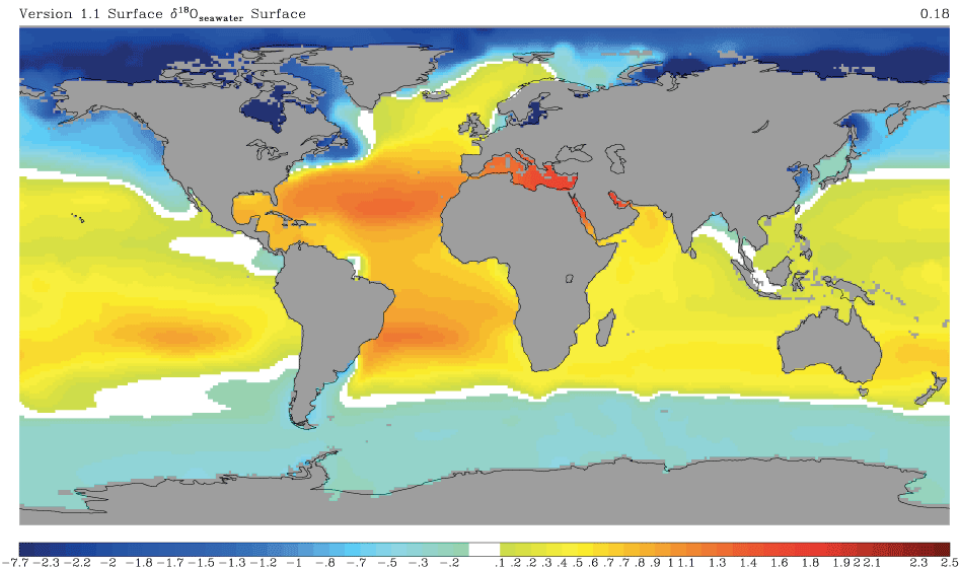

Figure 4: Ocean Surface $\delta^{18}O$ (NASA), Gridded Data Construction

Altitude & Terrain Altitude-induced stable isotope effects are the result of decreases in temperature and pressure with elevation [12]. Due to the pressure-temperature relationship between K and Q previously described (where the isotope effects were partially a function of the gas constant R, which itself is a function of temperature and pressure), a decrease in pressure implies a larger ΔT is needed to reach saturation [25]: $PV = nRT = NkT$ (where k is the Boltzmann Constant and N is the number of molecules, n is the number of moles of the gas) [25]. This effect was quantified early in the 1980s with a series of studies by Siegenthaler & Oeschger [27], as well as many other studies since then.

Altitude effects are different for $\delta^{18}O$ and $\delta^{2}H$, with the IAEA reporting effects of $-0.2‰/100m$ and $-1.5‰/100m$ respectively [12]

This is relatively well-supported in the literature. Liu et al. (2014) found an effect of $-0.13‰/100m$ in China, with this effect rising to $-0.30‰/100m$ in some areas [27]. Yu et al. (1980) found an effect of $-0.126‰/100m$ in Beijing [28]. Siegenthaler & Oeschger found a European effect of between $-0.16‰/100m$ and $-0.4‰/100m$ [26]. Many other studies have been reported as well.

The altitude effect may change with characteristics of the air parcel, include the amount that has already 'rained out' (this is part of a separate effect). Some studies have shown a non-linear (convex) relationship, with the effect decreasing overall with elevation [12].

Continentality A parcel of air moving over ocean experiences a different isotope effect compared to one moving over land. Air that is constantly in contact with the ocean is also able to consistently evaporate water that is within the ocean range of δ values (-8 to 3 ‰ [12]). For this reason, and since frozen water prevents the evaporation of ocean-composition water into the atmosphere (so there's a latitudinal limit), precipitation over the ocean tends to respond more to the ocean's composition than to temperature effects [12]. The IAEA reports precipitation over ocean as having $\delta^{18}O$ values of only 0‰ to -5‰ [12].

Over the continents, progressive rain-out results in the gradual depletion of heavy elements from the air, so inland atmospheric moisture tends to be lighter. The size and existence of the effect depends on the location, but it may be especially significant over Antarctica, which is the driest continent on earth [29], and where the continentality effect is exacerbated by steep elevation gradients.

Seasonal Variability Due to the temperature effect and others, seasons change the stable isotopic composition of precipitation in a regular, periodic way. The seasonal effect is critical to reconstructing recent patterns in locations where weather station or remotely sensed data are not available. Examples include near-subtropical corals, in which a combination of stable isotope variability and different seasonal growth rates allows for visual counting of annual layers [30]. A second, which is employed in this paper, is the use of seasonal variability to compute accumulation rate in Antarctica.

However, that actual magnitude of the effect is not important to the purpose of this paper, because we are analyzing spatial variability on average. Therefore, a more detailed discussion is omitted.

Wind Vectors One of the most difficult challenges to solve in determining the eventual stable isotopic composition of a precipitation event is the

fact that the composition of a cloud is defined by the path it takes before it reaches the event site. NOAA meteorological projections exist up to 315 hours in the future and past [31], but the unpredictability of the atmosphere makes it difficult to predict air trajectories far in advance, unless we are computing general averages for prevailing winds. Furthermore, it is difficult to model when exactly precipitation will occur during an air parcel's trajectory. In cold regions, this is even more difficult since phase changes in the cloud also need to be modeled [3].

The National Centers for Environmental Prediction [33] and other sources [34] provide datasets that can be used to simulate average wind patterns across Antarctica, which are shown in Figure 5. As expected, simulations show that five elevations, wind patterns are very different. The Antarctic surface is concave and therefore it experiences katabatic winds outward toward the coast under gravitational force [35], and are especially strong in areas of persistent steep terrain, such as glaciers [35]. In the Surface simulation, this can be seen at the edge of the Ross Ice Shelf, the Ronne-Berkner Ice Shelf, and on the other side of the continent, Adlie Land [36]. Most coastal areas can be seen to have outward-pointing wind vectors, often faster than the center [34]. In fact, surface winds more or less follow the contours of the main ice catchment divides in Antarctica [34].

These winds may have an impact on post-depositional processes and accumulation, but are not the drivers of stable isotopic composition in precipitation. The winds that drive Antarctic weather are created by pressure gradients on a global scale. This can be seen in the 500 hPa and 250 hPa simulations, which show the consistent wind patterns driving the Antarctic Circumpolar Current [37].

These simulations show that the trajectory of an air parcel makes it far from simple to predict the isotopic composition of a precipitation event. Many models use distance from the coast to map the continentality impact but, as is demonstrated statistically further in the paper, this is an oversimplification. A path-dependent approach is needed, because the isotopic composition of precipitation is partially dependent on the distance of land over which the air moves, as the continentality, elevation, and amount effects all play a role.

In Antarctica, a mathematical framework for path-dependent (rather than straight-line-distance) computations of stable isotope composition may

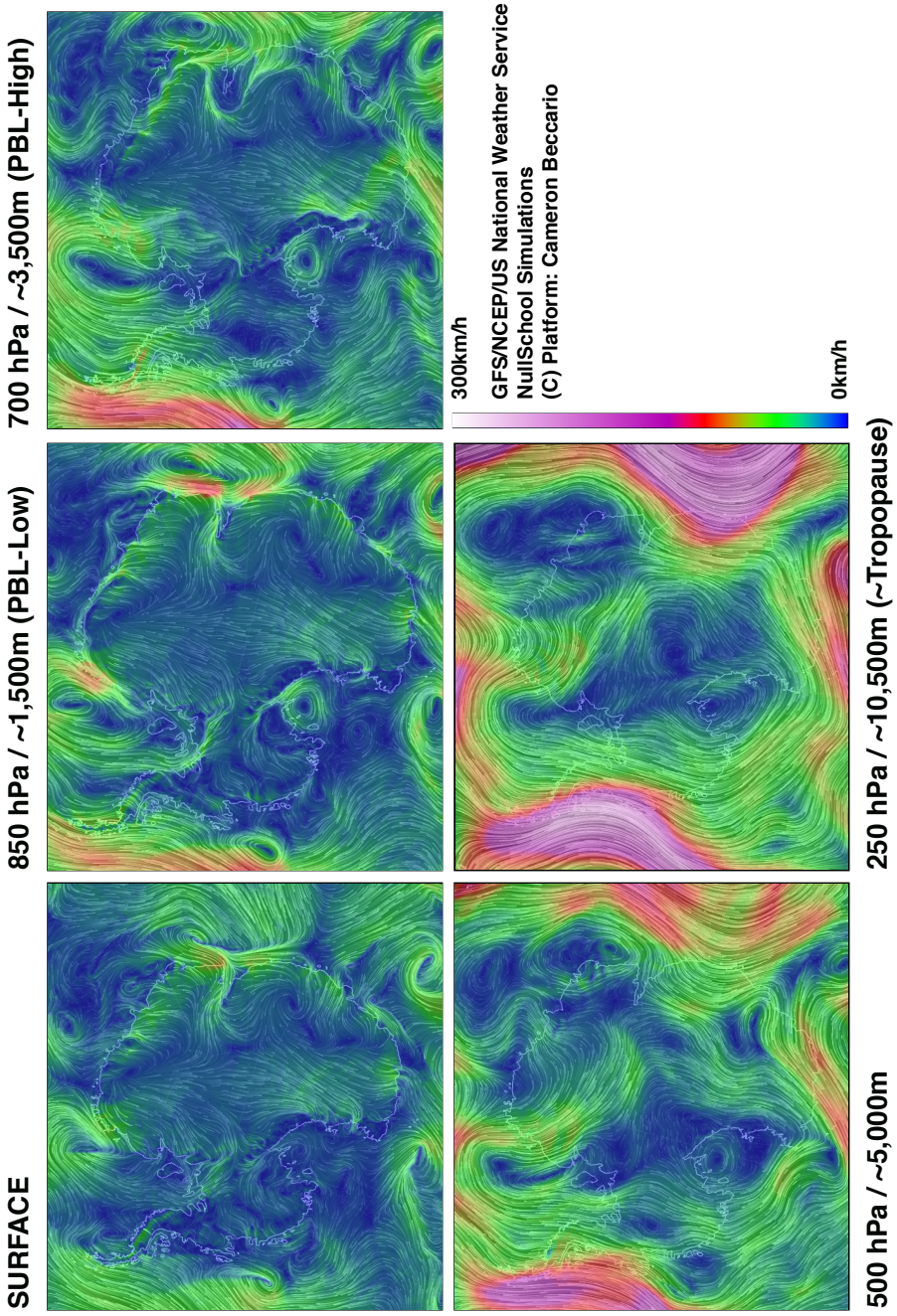

Figure 5: Wind Vectors at Five Tropospheric Heights over Antarctica [34]

Table 2: Geographical Isotope Effect Summary

Effect	Latitude	Altitude	Continentality	Air Traj.
Size of Effect	$-0.2‰/deg$	-0.2 to $-0.6‰/100m$	Varies	Unquant.

be part of the solution, especially if combined with wind patterns generated by AGCMs or other simulations. A starting point for this is developed in Section 6.

Overview A summary of geographic isotope effects can be found in Table 2 [12].

4 Antarctic Stable Isotope Data

4.1 Introduction and Objectives

In this section, I explore all the data that are recently available from research expeditions to Antarctica, which feature information about $\delta^{18}O$ and $\delta^{2}H$, as well as geographic characteristics such as GPS coordinates. From GPS location, we can compute stable isotope influencers, such as elevation and distance from the coast.

A large number of studies was aggregated by Masson et al. (2008) [7]. These studies were analyzed to try to better understand the spatial variability of stable isotopes in Antarctic precipitation. Much of the existing work on continent-wide assimilation of data for this purpose is brought together in that paper alone.

The Masson study makes use of an aggregated dataset featuring an average value for each site (each one of which could feature thousands of observations itself). It will be the basis of this section, since the study's 37 authors [7] have done much of the work in compiling studies of stable isotopic composition. The dataset employed by Masson et al. will be referred to as the "Masson Data". I have separately cleaned and analyzed the raw data myself.

This section will decompose the statistical approach taken by the Masson in the search for a better model in predicting the spatial variability of isotopes.

4.2 Amount of Existing Data

The Masson Data comprises averages from 1,280 locations in Antarctica. Each site involves a snow pit, firn core, or shallow ice core, taken between 1969 and 2007, but the range in depth (and thus, age) varies significantly from a few dozen centimeters to eighty meters (though the average is around 1-2 meters). Some of the datasets included in the study are from one-off coring projects at permanent stations; many others are from long transects of the Antarctic continent punctuated with many sampling sites. For purposes of transparency, the raw dataset is available at this link: Masson Data [38].

Each observation features a quality rating, GPS coordinates, $\delta^{18}O$ and $\delta^{2}H$ mean, min, max, range, and standard deviations, as well as information about the Deuterium excess, and some context about the site (date drilled, expedition name, depth, author of the study, etc.).

Willis Resilience Expedition Data In November-December 2013, I led a research expedition to Antarctica, in which our expedition team extracted samples from cores and snow pits at 10 sites, approximately evenly spaced along two transects. The expedition was called the Willis Resilience Expedition, and is shortened to WRE in this paper. In Section 5, I present the results of the analysis of these data, separately from the Masson Data.

This analysis makes use of several versions of the data, and each version of the dataset is openly available upon request. These datasets are outlined in Table 3.

4.3 Data Cleaning

Since the Masson Dataset is a compilation of many smaller datasets, there are naturally many missing observations, with many of the datasets included missing information on latitude, longitude, elevation, distance from the coast, or other factors. Masson-Delmotte 2008 analyzes correlations

Table 3: Available Data Products from this Analysis [7]

No.	Reference	Includes	n
1	Masson-1	Full Masson dataset	1,280
2	Masson-Water	Masson Data including δ D and $\delta^{18}O$ means.	794
3	WRE	Full Willis Resilience Expedition dataset.	198
4	M-WRE	Combined dataset: Masson-Geo and WRE.	992

only on the subsets of the dataset where the values are reported, which often means the 1,280 sites are reduced to 500-600 observations for many analyses [7].

In this study, analysis is based on an cleaned and enhanced version of the Masson Data. I first cleaned the dataset, removing missing observations that couldnt be replaced and adding in values that could be computed directly from the other variables. The dataset was then combined with data from the 2013 Willis Resilience Expedition (WRE), which adds a further 198 observations taken from ten sites on two different transects across the continent.

Some sites are missing GPS coordinate information, which makes their omission from the models unavoidable. However, some of the omitted observations can be reconstructed using only the latitude and longitude data. This includes 230 observations where elevation data are missing, 174 of which I filled in using the Digital Elevation Models from the Scientific Committee on Antarctic Research (SCAR) [39]. The remaining elevations couldnt be computed since those sites lacked information on latitude and longitude.

Other missing information was recovered through a literature search, including GPS locations for several stations at which firms cores were taken and analyzed by Lorius and Merlivat, 1977 [40].

Masson et al. followed a similar protocol for computing distance to the nearest coast, which can be reconstructed to sufficient precision (within 5km) using GPS coordinates and either any accurate measurement tool (in software such as Google Earth, ArcGIS, or ArcExplorer) or an algorithm in Matlab or Mathematica. The remaining missing values for Distance from the Coast were computed and added to Masson-2.

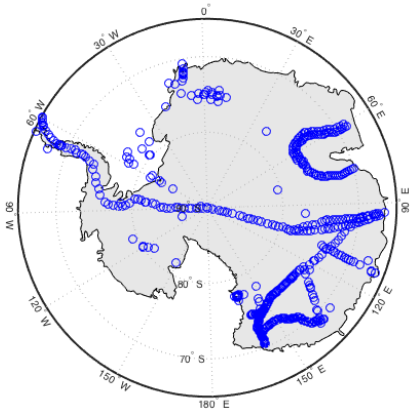

Figure 6: Locations of Masson Data Sampling Sites [38]

4.4 Masson Regressions

The Masson study constructs multivariate regression models linking stable isotopic composition to the factors known to control them.

The wide range in physical characteristics of the sampling locations gives us the opportunity to analyze, for example, isotopic composition of samples at different elevations but similar latitudes, or samples at similar elevations but different distances from the coast. This is especially important for confidence in the results of multivariate regression models, which are based entirely on the *ceteris paribus* effect of a per-unit change in each variable.

Masson proposes the following models:

$$\delta D = 944.40 \sin(\lambda) - 0.057H - 0.034D \tag{15}$$

$$\delta^{18}O = 151.9 \sin(\lambda) - 0.007H - 0.003D \tag{16}$$

where λ is latitude, in radians, H is elevation in meters, and D is distance from the nearest coast in kilometers.

The Masson regressions are among the only attempts to characterize continent-wide water isotope distribution using aggregate datasets, of what I could find. They raise important questions about the use of statistical techniques for inferring large-scale patterns, and these questions are the central purpose of this essay. In addressing these questions, I hope to construct a set of parameters by which stable isotope variability can be more accurately predicted in this region, where very small amounts of data are available, and where these data are not anywhere close to equally distributed across the continent (1,280 sites may seem like a lot, but this is over the course of nearly half a century, in a continent larger than China and India combined).

Questions to be addressed include:

1. Collinearity of Variables: Latitude and Distance from the Coast are in lock-step and this causes bias in the mechanics of least-squares.
2. Distance from the coast: continentality matters, but the real distance that matters is how far an air parcel travels, without the assumption that it takes a straight path from the coast.
3. Elevation: in reality, δ values don't become heavier when elevation drops after a rise; in other words, Masson made the assumption of monotonic elevation increase.
4. Residuals: the error from the regression line at each point is a rudimentary measure of the quality of the model and its explanatory power, even when the model produces a high R^2.
5. Averaging properties: sample properties are different for unrelated observations analyzed together (Masson Data consists of 1,280 rows of unrelated studies).

4.5 Statistical Tests of Data Structure

Residuals The Masson regressions were run on a cleaned dataset with $n = 754$, and the residuals were plotted against $\delta^2 H$. Any correlation between residuals δ values is indicative of a a weakness in the regression parameters [8]. By definition, the residuals are the deviation from the regression (i.e. the error), so correlation is a measure of *less* correlation between the explanatory

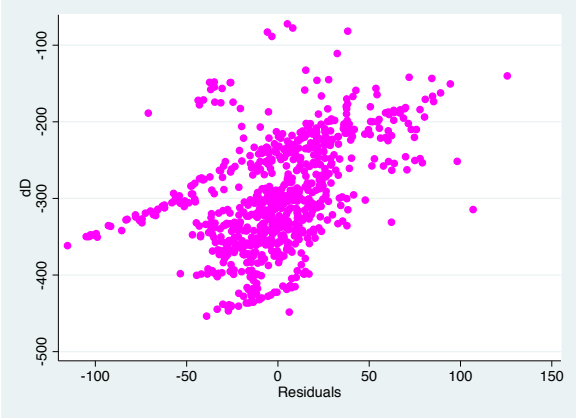

Figure 7: Trend in Masson $\delta^2 H$ Regression Residuals [38]

variables and the dependent one [8]. Ideally, the residuals are distributed normally around zero [8].

The residual plot reveals the heterogeneity of the data. There are clearly substructures correlated with δ values such as the strand starting at around $-370‰$, another along the $-400‰$ line, and another from $-450‰$ to $-400‰$ while other groups within the residuals seem completely uncorrelated. This demonstrates that the Masson model is more valid for some sampling sites and transects than for others. The conclusion is that **the Masson model is not appropriate for continent-wide inference of stable isotopes in precipitation**.

Heteroskedasticity I now test the Masson Dataset for variable-dependence of the variance in the residuals. The theory behind this procedure is based on the construction of ordinary least squares. It can be proven that OLS is the most efficient "BLUE", i.e. Best Linear Unbiased Estimator of the data, as long as the residuals have constant variance σ, though this proof is omitted [41]. If the variance is not constant, it means the data are *heteroskedastic*, and OLS is not the best linear estimator, so that other estimators would be better [41].

Hypothesis 1. *For the parameters suggested by Masson et al., Ordinary Least Squares is not the most efficient mechanism for predicting the spatial variability of stable isotopes in precipitation across Antarctica.*

Proof. **Breusch-Pagan Test** [42]. We test the null hypothesis, H_0, that the residuals are homoskedastic (constant variance), via a χ^2 test statistic.

$$\chi^2(3) = 176.64$$
$$Pr > \chi^2 = 0.0000$$

The test statistic is 176.64. The computed p-value is zero to four decimal places, which means we reject the null hypothesis with the utmost confidence. Therefore, the data are heteroskedastic, so OLS is not the Best Linear Unbiased Estimator, at least with the parameters suggested by Masson. □

This test raises the question of whether heteroskedasticity is inevitable. In other words, can we remove the non-constant variance problem through better parameterization? This will be tested in Section 6.

Averaging Properties Part of the issue with an average-value aggregated dataset is that if the original datasets were homoskedastic and had normally distributed residuals, OLS might have been an appropriate mechanism if all the data had been added together rather than having each site averaged first. But the averaging out meant that each set of residuals was averaged out as well, such that the effect of any distribution in the residuals was lost. Combined with the fact that the 754 observations were independently acquired, I find that any OLS regression loses much of its explanatory power.

We should seek a better model, and I use the WRE data to develop and support it.

5 Willis Resilience Expedition Data

This analysis is a result of a 2013 research expedition, undertaken in partnership with the Institute of Geological and Nuclear Sciences (New Zealand) and the ColdFacts Programme (Netherlands). The expedition had three aims:

- Analyze the factors influencing the spatial variability of stable isotopes in recent Antarctic precipitation.
- Map the rate of deposition of cosmogenic Tritium and its variability with Latitude, with the purpose of refining tools for back-dating Antarctic snow and ice samples to seasonal resolution, up to around 120 years.
- Conduct a 4-week field test of a new automatic weather station, the ColdFacts-3000BX, designed to be especially lightweight and easily deployable in comparison to existing models.

This paper focuses entirely on the stable isotope transects.

5.1 Route

WRE samples were collected along two transects from the coast of Antarctica toward the South Pole. Ten sites were established at which either shallow cores were extracted or snow pits created and samples. The first transect crossed the Polar Plateau from Union Glacier (near the base of the Antarctic Peninsula) to the Geographic South Pole. It comprises a gradual rise from near sea level towards the South Pole, via the Thiel Mountains, one that has been sampled in the past by the ITASE expeditions [38]. The second transect is a unique route that winds from the edge of the Ross Ice Shelf up the steep Leverett Glacier, where 78% of the total rise in elevation occurs in less than one degree of latitude between the 1st and 2nd sampling sites [43]. After the glacier, route takes a dog-leg East toward to the South Pole across the (mostly flat) Plateau. At 560km, the second transect is approximately half the length of the first, while covering the same elevation change.

5.2 Sampling Procedure

The two transects were completed consecutively under the same research permits. The sampling procedure was undertaken by a team of five, but was different for Transect 1 versus Transect 2. The WRE team crossed the continent in a custom-modified Toyota Hilux, which housed all of the sampling equipment as well as essential gear for survival in Antarctica.

Table 4: Willis Resilience Expedition Sampling Sites

Site #	Latitude	Longitude	Elevation (m)	Dist. to Coast (km)	Core Length (cm)
1	-80.1251	-81.11293	884	735	218.0
2	-82.50101	-79.4698	954	817	195.5
3	-83.38184	-80.06116	1163	887	153.0
4	-85.0829	-80.78872	1353	1026	97.0
5	-87.03575	-81.7369	2271	1092	68.3
6	-90.000	NA	2797	1195	39.5
7	-89.15407	-131.9842	2863	1059	71.8
8	-87.55679	-132.0019	2782	953	97.5
9	-86.20755	-136.4033	2602	852	136.5
10	-85.58694	-153.027	419	840	195.0

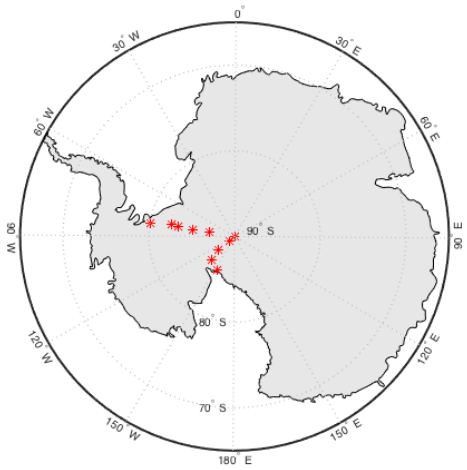

Figure 8: Willis Expedition Sampling Locations

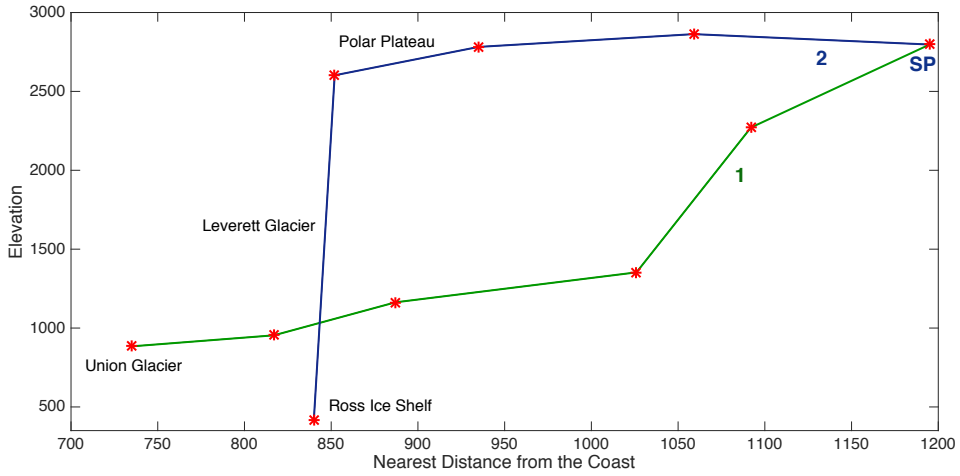

Figure 9: Elevation Cross-Section of WRE Sampling Locations

General Procedures At each site, sampling was carried out at least 100 meters upwind from the truck, to avoid pollution from the exhaust. Samples were stored in polyethylene vials, and protected by foam in 'Pelican' cases (hard-shell cases often used to protect fragile equipment, such as camera lenses). Watches were removed near the sampling site, to avoid contaminating the samples with Tritium, which was being analyzed as part of a separate study related to the expedition (Tritium is found in high concentrations in luminous dials in watches). Approximately 20 samples were collected at each site. The depth and resolution of the samples were chosen based on the estimated accumulation rate at each site, in order to capture approximately the same number of years.

Transect 1 Sampling During the first transect, shallow cores were extracted using a hand auger. This piece of equipment allowed us to drill cores up to 250cm in depth. The core was then separated into individual samples and stored into cases on site. Samples were cut using a stainless steel knife and crushed in a separate container, and the vial filled to the top.

Transect 2 Sampling During the second transect, samples were collected from snow pits rather than cores (the auger was on loan from the Chilean Antarctic Institute (INACH), and was returned at the South Pole Station).

Snow pits were dug such that they had a vertical face, and sample cutoff points were marked using small wooden pegs and a tape measure. Samples were then cut using a stainless steel knife and crushed in a separate container, and the vial filled to the top.

Extraction After extraction from Antarctica, samples were stored in southern Chile for two days, then shipped to New Zealand to be stored at the Institute of Geological and Nuclear Sciences (GNS), awaiting analysis. Stable isotope analysis was conducted approximately two years after the samples were initially taken, and Tritium analysis has not yet been completed. Due to a Customs strike in Chile, equipment was delayed in Santiago. Replacement vials and an auger were borrowed from Instituto Antartico Chileno (INACH), so our sampling procedure varied between the two transects as some of the equipment had to be returned at the Amundsen-Scott South Pole Station.

5.3 Data & WRE Results

5.3.1 Stable Isotope Relationships with Depth

Results for $\delta^{18}O$ are omitted, given the high sampling quality and a close adherence the local meteoric water line, which is discussed later. The relationships would look identical, with the exception of two missing $\delta^{18}O$ values at Site 6, which when computed from the Deuterium values, fit well into the cycle of that year.

The WRE results show strong elevation and continentality signals, with observations becoming more negative with distance from the coast and with elevation. There is also a clear seasonal signal at each site, though it is not sufficiently strong to elucidate accumulation rate visually. Results are summarized in the two figures, which shows the variation in $\delta^{2}H$ by sample number. Site numbers are in bold on the left side of the image, and are in order, so the figure can be seen as a cross-section of Antarctica from coast to coast.

As expected, there are strong but noisy seasonal cycles within each site, with the number of annual cycles decreasing as the sites get closer to the

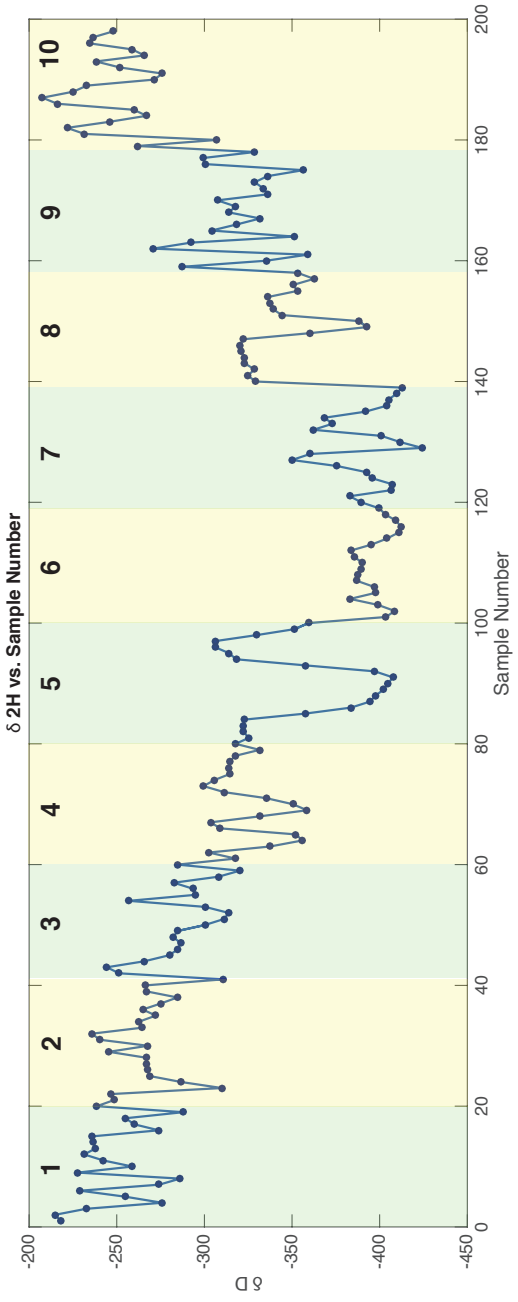

Figure 10: WRE Stable Isotope Results

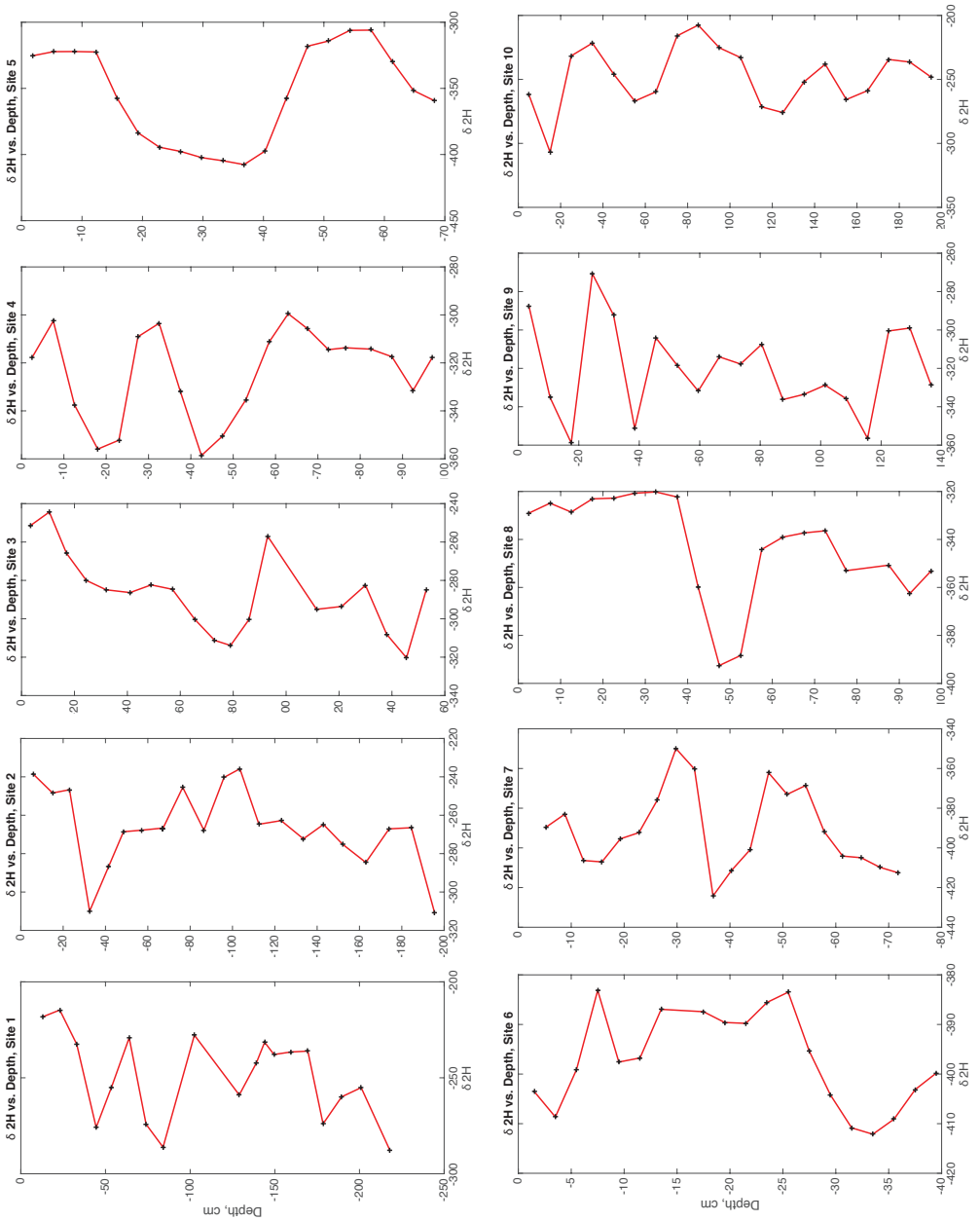

Figure 11: $\delta^2 H$ vs. Depth, Willis Resilience Sampling Locations

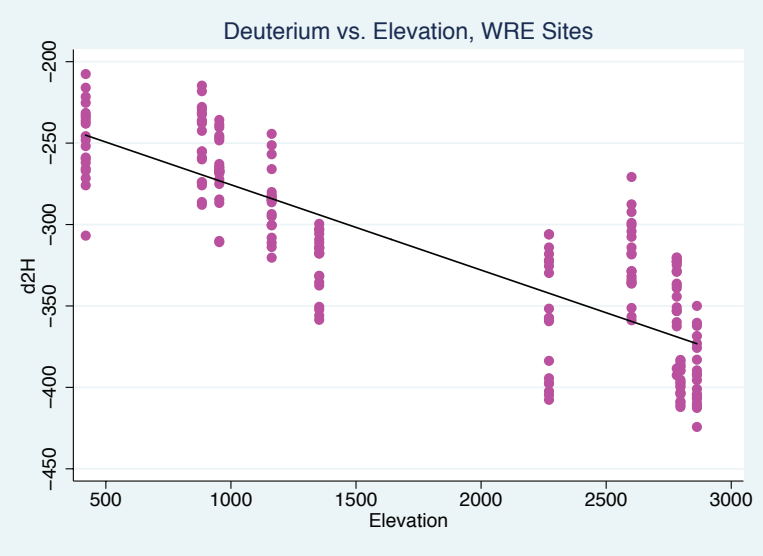

Figure 12: Correlation between Deuterium ratio and Elevation

pole. From these results we can compute accumulation rates as well as temperature, and from the accumulation rates we can attempt to compare our temperature reconstructions to nearby weather stations.

5.3.2 Data Quality & Antarctic Meteoric Water Line

Given the long wait time and the alternate equipment used to store the samples, it is important to ensure the samples have not been compromised between extraction and analysis. The quality and preservation of the samples can be tested by comparison of the dataset to the established global relationship between hydrogen and oxygen isotope ratios, where the Global Meteoric Water Line (GMWL) is given as [44] [45] to be:

$$\delta D = (8.17 \pm 0.06)\delta^{18}O + (10.35 \pm 0.65) \tag{17}$$

Observations from WRE show a close relationship, approximating an offset of the GMWL. A simple OLS regression on the data result in an R-Squared of 0.9968 and with fitted values representing the following regression line:

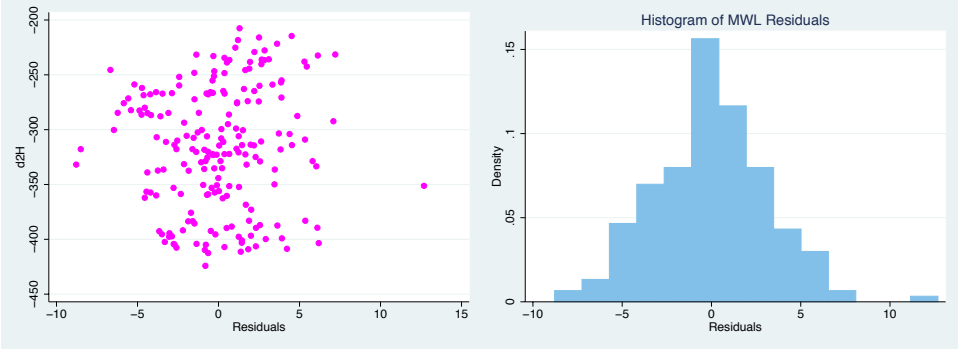

Figure 13: Distribution of LMWL Residuals

$$\delta D = 7.762887\delta^{18}O - 4.77075 \tag{18}$$

Using the definition of Deuterium excess as [45]:

$$d = \delta D - 8\delta^{18}O \tag{19}$$

The fitted values reflects an average D-Excess of:

$$-0.237\delta^{18}O - 4.77075 \tag{20}$$

While it could be assumed that the slope coefficient on the average antarctic Local Meteoric Water Line (LMWL) roughly approximates the GMWL (and that perhaps its good enough to analyze only the offset and not the coefficient), its safer to determine the statistical significance of the $\delta^{18}O$ coefficient.

We first must determine whether the WRE data are suitable for interpretation with parametric tests of significance. Parametric tests generally assume a normal or χ^2 distribution of the residuals. To test for normality, we form a new variable, *mwlhat*, composed of the residuals from the regression of $\delta^{18}O$ on δD. The first instinct is to test visually by plotting a histogram of the residuals.

The residuals appear to be sufficiently close to a normal distribution to test more rigorously. A Skewness-Kurtosis Test of normality [46] provides a

Table 5: Test of Normality [73]

Skewness/Kurtosis Test				Joint	
Var	Obs	Pr(Skewness)	Pr(Kurtosis)	adj $\chi^2(2)$	Prob>χ^2
mwlhat	196	0.2741	0.0445	5.23	0.0730

more rigorous double-check on a null hypothesis H_0 that the residuals are normally distributed. As can be deduced from the name, this tests compares the Skewness and Kurtosis of *mwlhat* to that of the Standard Normal, and generates p-values to determine the probability that the data are normally distributed.

The results of the test show that it may be possible to reject H_0 on the basis that the p-value for Kurtosis (which refers to the shape of the peak) is below the critical value (α¡0.05). However, on the basis of Skewness and the Joint probability tests, it is impossible to reject the hypothesis H_0 that the residuals are normally distributed [46]. Therefore, we may proceed with the parametric tests of significance.

5.3.3 Antarctic Meteoric Water Line

Difference in Slope I argue that the data from the Willis Resilience Expedition, combined with those from the Masson Data, are sufficiently robust and cover a wide enough range of geographic characteristics to propose a Local Meteoric Water Line for Antarctica, which I define here as the *Antarctic Meteoric Water Line (AMWL).*

Hypothesis 2. *The Antarctic Meteoric Water Line is (1) persistently, (2) consistently, and (3) significantly different from the Global Meteoric Water Line.*

Proof. The proof depends on the normality of the residuals and the comparison of both the WRE MWL and the Masson MWL to the Global Meteoric Water Line. In particular, we seek to determine whether the *slope* of the regression line is different from the rest of the world.

1.1 We have already seen that the hypothesis that the WRE residuals are

normally distributed cannot be rejected. Furthermore, the error term (residuals) is independent of the variables. Thus, F-tests are valid.

1.2 See Appendix for construction of the F-statistic.

1.3 Testing the hypothesis that the coefficient on $\delta^{18}O$ is equal to 8.17, the coefficient for the Global Meteoric Water Line, we see the following results:

$$F(1, 194) = 165.51$$
$$Prob > F = 0.0000$$

1.4 Thus, the hypothesis that the coefficients on $\delta^{18}O$ are equal for the local and global meteoric water lines is **rejected** at the 0.0% significance level. In other words, the hypothesis is rejected with the utmost confidence.

1.5 The Masson data are not eligible for the F-test because they fail the Skewness-Kurtosis test. However, their **regression** output are nonetheless valid, which means we can compute a second F-test to demonstrate how closely the WRE data match data from $> 1,200$ other sites across Antarctica. This tests the hypothesis that the coefficient on $\delta^{18}O$ is equal to 7.75, the coefficient on the Masson data MWL.

$$F(1, 194) = 0.17$$
$$Prob > F = 0.6843$$

The test fails to reject the hypothesis, and in fact shows that it is likely ($> \frac{2}{3}$ probability) that the coefficients are equal in real life.

1.6 The sampling sites included in both the WRE and Masson regressions, and the Antarctic continent in general, experience a wide range in physical characteristics, such as elevation, distance from the coast, latitude, and others. However, from a statistical point of view, the WRE and Masson Meteoric Water Lines are identical. The local line is independent of local geographic or meteorological conditions.

1.8 Therefore,the Antarctic Meteoric Water Line is independent of the Global Meteoric Water Line. □

Lemma 1. *The IAEA predicts slightly different temperature effects for Oxygen and Hydrogen. The ratio of the change in $\delta^{18}O$ to the change in $\delta^{2}H$*

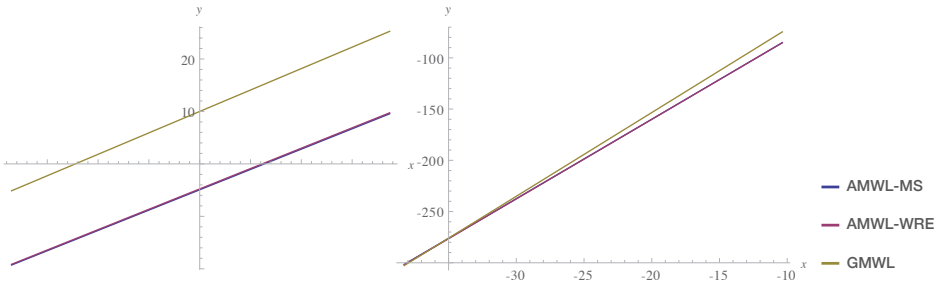

Figure 14: Antarctic vs. Global Meteoric Water Line

is 7.5 [12], which means that as elevation increases, the slope of the MWL decreases from the global average of 8.17; the result is the slightly (but statistically significantly) lower slope of 7.76 in the AMWL, on average across the continent.

The WRE dataset is too small (and the Masson dataset not eligible) to propose a rigorous line, but I propose the following relationship as a starting point:

$$\delta^2 H_{Ant} = (7.76 \pm 0.032) \times \delta^{18} O_{Ant} - (4.77 \pm 1.29) \tag{21}$$

Where Ant applies if $-90 \leq \lambda \leq -66.56$.

5.3.4 Accumulation

The reconstruction of accumulation rates is the most important component of the analysis of the WRE samples. Without accurate accumulation rates we cannot create a chronology of the years covered by the samples at each site, nor can we use temporal data, such as results from automatic weather stations.

Accumulation is impossible to compute directly or predict, because it involves not just precipitation rates (which already are difficult enough to predict), but also local processes that can't be captured by simulations or

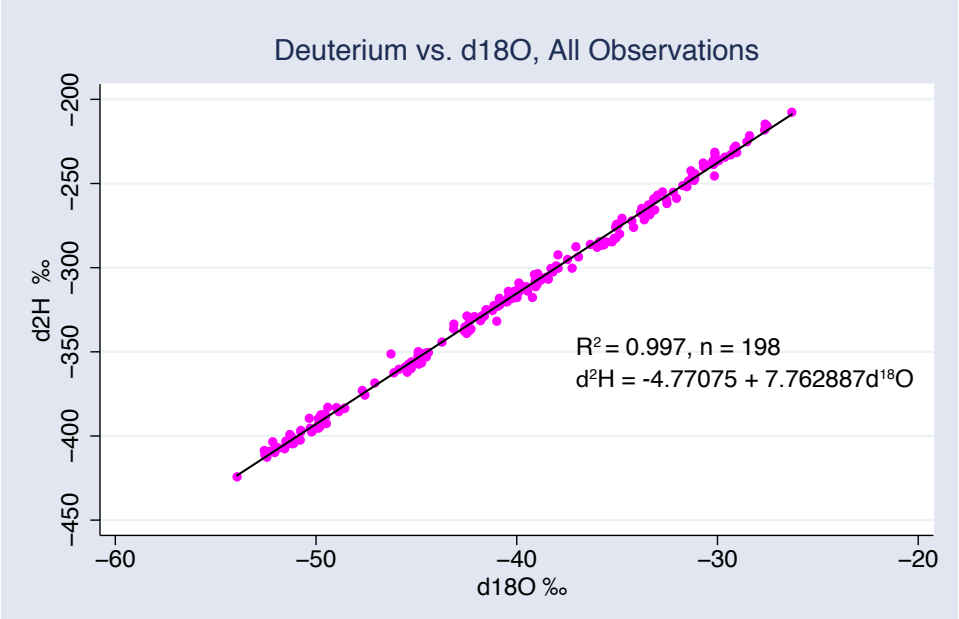

Figure 15: Local Meteoric Water Line, WRE Sites

models, such as sublimation and wind removal. The WRE data are slightly too noisy to reconstruct accumulation rates from seasonal signals visible in the data. In the past, accumulation rates have been computed from seasonal variations in $\delta^{18}O$ and δD [74]; counting of annual layers can be effective where the sampling resolution is sufficiently high to see a clear pattern in other words, where the ratio of samples taken to years in the dataset is high.

In order to resolve annual layers from the isotope record, sample quality must also be adequate. We have already seen that (1) a strong Meteoric Water Line, and (2) no significant outliers from the local regression, demonstrate that the WRE data are high quality and are likely to be reliable for the computation of accumulation rates. Data for δD are likely to be more useful than $\delta^{18}O$, due to a few missing results for both $\delta^{18}O$ and its standard deviation (in separate instances).

With a sample size of only $n = 196$, there are some limitations to the robustness of the data, but this is offset by the fact that we are reconstructing these values at a high number of sites (10) and that the cores are well-resolved. Here, again, the relatively high resolution offsets the concerns of

a smaller sample size.

Layer counting is feasible for sites where accumulation is approximately 1cm or greater [49], which appears to be the case in the WRE data. Casey et al. (2014) reminds us that the South Pole receives an average of around 8 cma^{-1}, which is relatively high compared to some parts of East Antarctica [49], where Dome C, Dome Fuji, and Vostok each receive less than 2.5 cma^{-1} [49]. Critically, though, they each see well over 1 cma^{-1} [49].

Sophisticated methods have recently been developed to resolve annual layers in areas where accumulation rates are much lower, including laser mass spectrometry [49]. However, while data across the whole continent are generally very restricted, the geographic and climatological characteristics of WRE Transects 1 and 2, combined with the data we have collected, make it safe to assert the robustness of using seasonal stable isotope variability to reconstruct accumulation without having to rely on more sophisticated tools.

In more recent studies, a variety of other methods have also been used, including analyses of dust and ion cycles. For instance, it is possible to use the deposition of marine elements on the Antarctic interior [50] or dust [50] to verify seasonal cycles apparent in the stable isotope data. However, these measurements are not available from the WRE study.

Reconstructed WRE Accumulation Rates Reconstructing accumulation rates from a noisy dataset is challenging. It could be done simply visually, but this opens up the analysis to many forms of human error, especially because the relatively small sample size prohibits us from generalizing. Seasons could also be defined from the δD deviation from a straight line that is a function of *all* the points in the core, but the following issues must be taken into consideration:

- There may be an overall trend in stable isotope composition over time (not just seasonal).
- Just one (or more, if they exist) outliers in a set has (or have) the potential to throw off a whole season.
- Missing data points can result in a false impression of the boundaries of seasons, since the highest-resolution picture we have is composed of

straight-line connections between points.

- One strange year might change the whole dataset; these cores were purposefully extracted to capture between 2 and 5 years and not more.

We could use a linear OLS function to find the cutoff for seasons, but this would still have difficulty capturing the effect of outliers or noise, even if we later compare the result to other programs that have taken place in nearby locations in Antarctica to determine the plausibility of our results.

An alternative solution is to use Kernel-Weighted Local Polynomial Smoothing [51], which is a non-parametric approach to estimating a polynomial fit for the data from which we can smooth out noise. This intuitively makes sense, since there is no reason why real-world data should fit the parameters of a generalized polynomial.

Let us define K_α as the Epanechnikov Kernel, as well as a design density function $\hat{f}_{K\alpha}$, and an estimated function $\mathbf{y}$. The technical details of their construction are in the appendix (we use the Epanechnikov type because it is the most efficient of the common kernel designs [52]) [53].

We will choose a function such that the order of the polynomial is sufficiently high for us to define the seasonal cycles, but low enough that the effect of outliers and trend do not disrupt the estimate while also being careful that the polynomial generated does not create 'phantom seasons'. This is simply a matter of trial and error and judgement. Consideration is also made toward the fact that accumulation rates don't very tremendously between years within an overall span of just a few years.

Figure 16 shows the results of the Epanechnikov analysis as well as the established annual layers based on a combination the kernel analysis and the data points themselves. Potential 'problem areas' are encircled in blue, and are explained further below. Note that the far left and far right boxes in each panel do not represent a full year; only the color bars not on the left or right sides of the graphs are used to determine accumulation rates.

- Site 1: Due to the structure of the underlying data, I consider that peak sufficient to round out the year.
- Site 4: Suspected mixing error.

Table 6: Computed Accumulation Rates, WRE Sites

Site	1	2	3	4	5	6	7	8	9	10
Ave. A (cm)	63.5	98.0	40.5	32.0	52.5	11.3	24.5	35.0	38.5	53.3

- Site 8: Suspected mixing error.
- Site 9: (a) Erroneous peak caused by polynomial function, disregarded; (b) suspected mixing error.
- Site 10: Erroneous peak caused by polynomial function.

"Mixing Error" refers to the possibility that snow from above or below was accidentally mixed into the sample, causing the erasure of the seasonal peak. Though we can't be sure that it was the cause of the noise in the observations, this was occasionally an issue during the expedition, and the years above and below can be used to contextualize the estimate of accumulation from the data.

Accumulation rates are within the expected range, comparable with the range from different parts of Antarctica in the Masson Data [38]. The average rates were computed from 1-3 complete years for each site, with the South Pole station having the lowest accumulation of any site. Sites closer to the coast had higher accumulation rates, since lower elevation and and less distance-travelled makes for higher precipitation rates.

As previously mentioned, Transect 2 is unique and therefore there are no previously collected samples in the Masson Data to which we can compare sites 7 to 10. The Masson data features seven sites near the route of Transect 1, from along the ITASE expedition [38]; however, these were published in 1994 so are not useful for quality comparison to the WRE samples. Dahe et al. (1994) found a South Pole accumulation rate of 8.50cm [54], not too dissimilar to the 11.3cm found in the WRE samples, especially given how long before WRE those samples were taken. However, the Masson Data features very few sites with accumulation rates available beyond -80 degrees of Latitude.

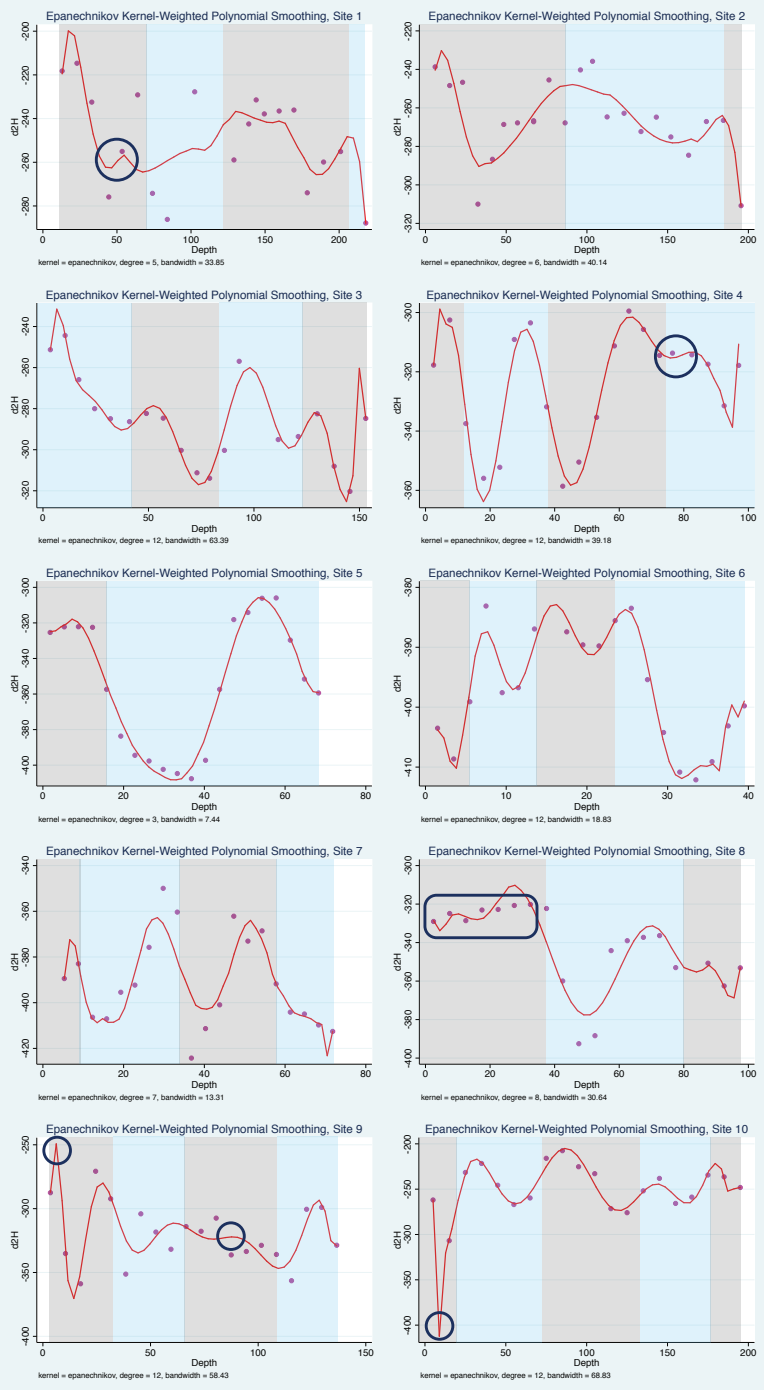

Figure 16: Results of Epanechnikov Kernel Analysis

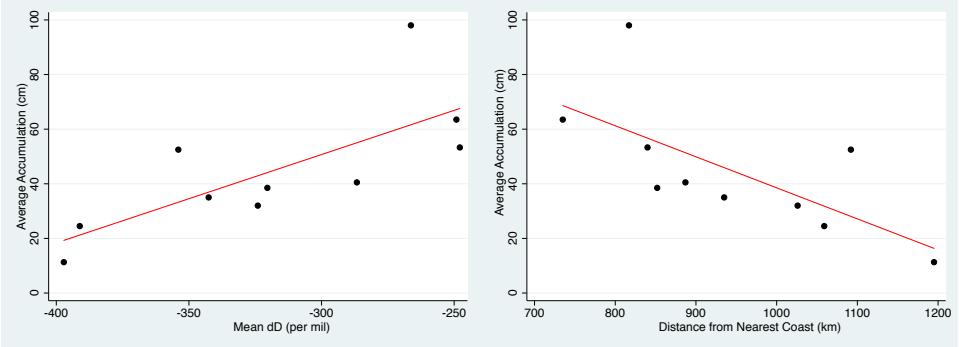

Figure 17: Computed Average Accumulation Rate δD and Distance from the Coast

5.3.5 Temperature

WRE data were compared to British Antarctic Survey data at 114 sites across Antarctica (about half of which were stations, the other half deployed automatic weather stations) [55]. As can be seen in the AWS map, the WRE transects feature effectively no ground meteorological data at all, with 0 weather stations located near either route after about 50km from the South Pole. The National Snow and Ice Data Center didn't feature any additional datasets either [56]. NASA IceBridge data weren't available for the years and locations necessary [56], and firn temperature measurements were listed on NSIDC but not downloadable [56]. Surface temperature datasets were no longer available on the BAS website [55]. NOAA also had no further available data [58], and other NASA tools were too coarse to be useful [59].

It's worth noting that one of the other objectives of the Willis Resilience Expedition was to test a lightweight automatic weather station that could potentially be deployed permanently: the ColdFacts-3000BX, developed specifically for Antarctica as part of this expedition. The lack of surface observations makes the analysis more difficult; these hurdles were the exact motivation behind the development and testing of this weather station. More surface observations are needed.

The closest stations to the WRE sampling sites are: Clean Air (South Pole), Amundsen-Scott (South Pole), Nico, Lettau, Henry, Erin, Elaine, and Elizabeth stations. Of those stations, we can eliminate Clean Air (repetition and because data from the station aren't available after 2009), Erin and

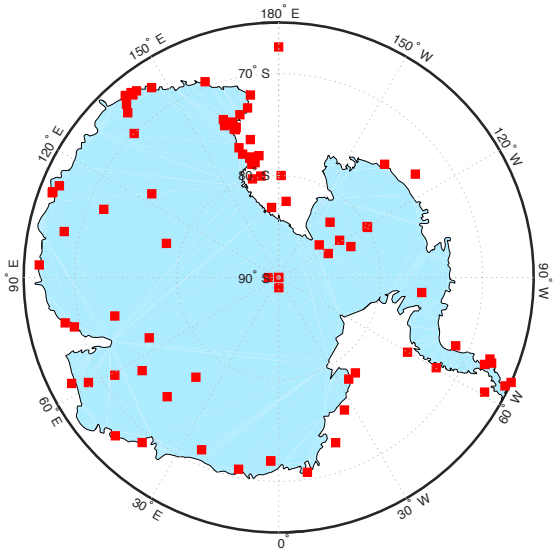

Figure 18: Stations and AWS from which Continuous Weather Data are Available [57]

Table 7: WRE Temperature Reconstructions

Site	1	2	3	4	5	6	7	8	9	10
Global Reconstructed T	-31.55	-32.46	-33.72	-36.36	-38.34	-41.28	-40.93	-37.57	-36.28	-31.19
Surface Observations						-48.35				-22.15

Elizabeth (located in West Antarctica, i.e. too far away), and Henry and Nico (too close to Amundsen-Scott). Those remaining are Amundsen-Scott, Elaine and Lettau. The latter two are within three dozen miles of each other and therefore only one can be used. Elaine, located at $-83.1S$ $174.2E$, has the most complete record so we eliminate Lettau. Thus, we only have two ground-truth datasets to which we can compare the computed temperature values: AS and Elaine.

In the table, the reconstructed temperature uses averages of $\delta^{18}O$ from the WRE record, and the surface observed temperature consists of averages from the years which had full records between 2009 and 2013 (WRE took place in late 2013). For Amundsen-Scott (6), the record overestimates temperature by 7.07^oC, and for Elaine (10), the record underestimates temperature by 9.04^oC.

There do not appear to be Antarctic-specific equations for reconstructing temperature from isotopes in precipitation, but these errors are not unexpected. Like Site 10, Station Elaine is located on the Ross Ice Shelf, but much closer to the coast, subjecting it much more to the moderating influence of the ocean. The fact that Elaine is also located 278.9km North of Site 10 (from latitude GPS, but 442.8km straight-line distance between points). Adjusting for latitude effect improves the effect marginally, but only by 0.2^oC. The remaining error is likely due to a continentality effect that can't be quantified because the 'distance from the coast' effect varies significantly around the world and has not been properly quantified for Antarctica. We will have to be satisfied with the intuitive logic that it makes sense for Station Elaine to have a slightly more elevated average temperature record.

As for the South Pole error, we see another effect that is difficult to quantify. This is the behavior of precipitation that is *extremely* depleted in heavy isotopes.

5.4 Reconstructed Chronology

From the accumulation rates and other computations in this section, we can translate the depth in centimeters to a chronology for δ values, temperature, and more, across the continent. This is shown in the figure. Deuterium values are omitted since they would trace $\delta^{18}O$ values exactly. Where Oxygen values were missing, they were computed from Deuterium using the Antarctic Meteoric Water Line.

We start from the fact that in the Antarctic summer (roughly October through January), δ values will tend to be less depleted in heavy isotopes. In the Epanechnikov analysis, this is clear; the pattern in shallow depths repeats itself across sites. The kernel-weighted polynomial was used to infer the month roughly associated with each data point, using δ peaks and troughs for the Antarctic summer and winter respectively.

The chronology is tabulated in the Appendix.

6 Theoretical Statistics & Mathematical Analysis

6.1 Non-Parametric Representation of Antarctic Surface

The above results show the extent to which the sample properties have been compromised by the effect of averaging out datasets. It also suggests that a new framework is needed to study the distribution of stable isotopes in Antarctic precipitation, one in which we return to the root causes of isotope effects and the importance of factors that are difficult to map in Antarctica, such as wind vectors; furthermore, possible elevation decreases shouldn't compromise the quality of the model.

In determining the best functional form for models describing stable isotope fractionation across Antarctica, the surface of the continent must be appropriately represented (or approximated) mathematically. This is not usually done in studies involving physical sampling and observations, because a wealth of data is available regarding the geographic characteristics of the Antarctic continent. For example, high-resolution DEMs map out the topography of the continent; in fact, this is how the elevations for each WRE sampling were computed.

These databases are useful for providing context to sampling site characteristics, but I am less interested in these datasets for continent-wide inference of stable-isotope composition of surface snow. In order to infer a statistical relationship and 'fill in' the missing points, a continuous approximation of the surface is needed.

Most of the isotope effects don't involve a three-dimensional view of the Antarctic continent, rather a two-dimensional motion vector of a parcel of air; which means we have to hold constant a variable like longitude. On the other hand, parcels of air do not move in a straight line from the coast (another assumption made by the Masson regressions), so a function must be flexible to account for 'wandering' around the continent.

This leads me to propose a theoretical approximation of the Antarctic surface which can be plugged into parameters for least-squares regression, to allow for more accurate model fit, and as a result, more predictive power. This is useful when the available data from the continent are sparse and lopsided.

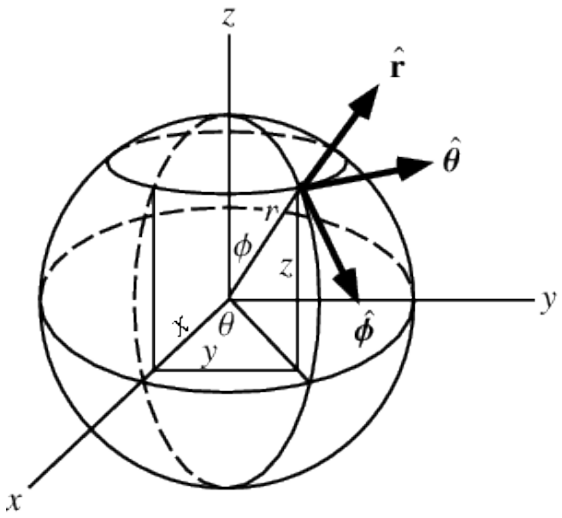

Figure 19: Spherical Interpretation of Antarctic GPS Coordinates [60]

Starting point: Antarctica in Polar Coordinates Antarctica's three dimensions can be considered in polar coordinates, such that:

$$\delta^{18}O, \delta D = f(\phi, \theta, D_p) \tag{22}$$

Where θ is longitude, ϕ is a function of the elevation z where $z = D_p \cdot \tan(90 - \phi)$, and D_p is distance from the pole ('radius'). At the South Pole, $z = 2797$, $D_p = 0$, $\phi = 0$.

We can see immediately that holding θ fixed, D_p is perfectly collinear with latitude, where $\lambda = D_p + c$, where c is a constant. This will be critical in the parameterization of any unbiased estimator to predict the isotopic composition of precipitation at a given location in Antarctica.

There are three dimensions; one along the path from the coast to the pole; a second in elevation; and a third in the longitudinal angle from the 0-degree line. Holding fixed θ, the topography of Antarctica from coast to pole is fixed. The only feature in common that each function has is the starting and end elevation: $z = 0$ at $D_p = \max[g(\theta_i)]$ and $z = 2797$ at $D_p = 0$.

This leads us to the use of a simple but powerful tool of mathematics:

The Calculus of Variations. This field of analysis usually deals with the optimization of functions and finding extremal functions, but in this case, it is a useful framework for parameterizing our model because we can bound the extremes and find local extrema without ending up with far-flung outliers like the Masson regressions did. Instead of optimizing the function to aim for the minimum possible distance, we are looking for a way to harness the chaos.

The theory of variational calculus tells us that a set of continuous functions with the same start and end point can interpreted as a functional with a vector of inputs and adjustment functions [61].

This is useful because for each θ_i, we have a straight-line cross-section that is a variation of a single function of horizontal distance toward the pole vs. elevation, with each beginning and ending at the same point (i.e. all cross sections from the coast can be interpreted as variations of the same function).

It also means we know that the lowest-average and highest-average elevation paths, and that any path taken across the continent must be between them:

Hypothesis 3. *: **Elevation Bounding**.*
Let $J_{min}(D)$ and $J_{max}(D)$ be the two extremal functions of the relationship between distance and elevation from the coast of Antarctica to the South Pole. Then, J(D) takes all possible values between $J_{min}(D)$ and $J_{max}(D)$ for $0 < \theta_i < 360$.

Proof. The proof is simple and relies on the assumption that θ is continuous, which it must be since it spans terrain (and Earth is continuous there is no point on Antarctica at which the surface fails to exist). Since θ is continuous, then $\forall\theta_i$, $\exists$ some form of $J(D)$ for which J_i is either equal to, greater than, or less than J_{i-1}. By the Intermediate Value Property and the continuity of θ, this must be the case for every $J_i(D)$ between $J_{min}(D)$ and $J_{max}(D)$, and thus every J exists. The Antarctic surface approximation is thus continuous in three dimensions: longitude, latitude/distance from the coast, and elevation. □

The reason for the above exercise was to prove continuity for the purpose of approximating stable isotope composition of precipitation at any one

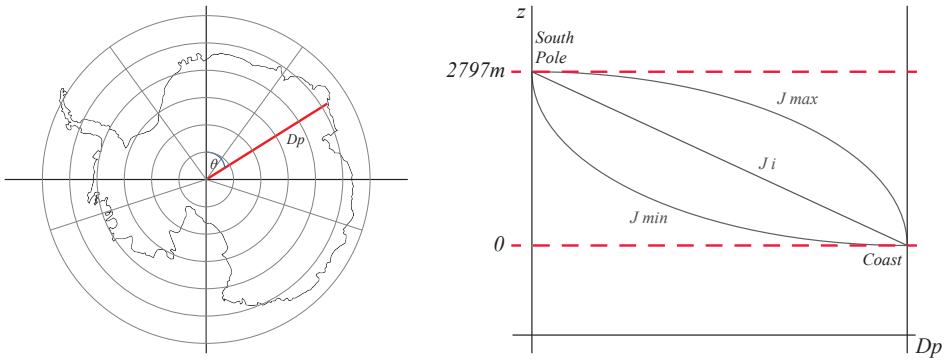

Figure 20: Highly simplified illustration of extremal functions for elevation in cross-sections of Antarctica; a unique J exists for each θ

location.

This is a simple and intuitive hypothesis, but is important because it implies that the parameter that will later be applied to the stable isotope model is dependent on the neighborhood, that is to say, stable isotopic composition of precipitation is path-dependent, and this proof shows we can sidestep the problem of the two-way elevation effect. This is because the Calculus of Variations uses functionals, rather than functions [61], which means the output of the optimization process is a scalar that is dependent on a vector of inputs (i.e. the output scalar is dependent on every component of the vector input, which includes the slope of the elevation).

Defining the Functional Optimization Problem (Reed, M.C. 1998 [62]). Let J be a function of three variables, D_p, $z(D)$, and $z'(D)$ (first derivative of z), such that:

$$J(D_p, z, z') = \int_0^{max_i} l(D_p, z, z')dz \tag{23}$$

is an extremal (maximized or minimized). Note that J depends on z and D_p, and is unique for each θ_i. Now, let $\eta_i(D_p)$ be a continuously differentiable function dependent on distance from the Pole, such that $\eta_i(a) = \eta_i(b) = 0$ (in other words, η_i vanishes at both ends of the extremal terrain function). η represents any variation in the distance-elevation function for different values of θ_i.

This means we can define the elevation transform $I_i(\varepsilon)$ as [62]:

$$I_i(\varepsilon) \equiv J(D_1, z + \varepsilon\eta(D_p), z' + \varepsilon\eta(D_p)). \tag{24}$$

This is the case for the straight-line distance from the coast model. However, air doesn't travel in a straight line from the coast. So, I represents the optimized (the 'ideal') air parcel trajectory which would minimize the isotope effect, but this is not remotely close to the real-world trajectory.

The point of this exercise is to demonstrate that the distance-from-the-coast paradigm of the effect of continentality on stable isotope composition is somewhat flawed. While it is an easy metric correlated with isotopes, it's not a predictor; and we can more effectively project the effect if we can parameterize statistical models with the variation in elevation vertically and laterally across the continent. The lateral variation will be discussed later in the section.

We can't assume that the surface of Antarctica is infinitely differentiable, so we cannot approximate with a Taylor expansion. The surface of Antarctica is only known at the highest resolution as a mesh of more than 56 million data points [63] in three-space. We could approximate the shape with a Fourier transform, but I prefer to approximate with a non-parametric density-weighted function, because this gives us the ultimate flexibility.

From non-parametric polynomial approximations in two directions distance across the surface of Antarctica and elevation at any particular point we solve two problems. First, we can include in the model the total distance traveled by an air parcel rather than the straight-line distance from the coast. Second, we can find a more accurate approximation for the highest elevation over which the air has passed, which is not likely to be the highest elevation on the straight line function I from the nearest coast.

It should be emphasized that this section is a theoretical exercise, designed to demonstrate that it is possible to do better than simple OLS regression. However, by combining these parameters with real-world data on geographic variables such as elevation, we can generate real values each of them.

6.2 Air Trajectories

We can predict moisture trajectories only very roughly, and only within a few days; and we can't predict very well whether precipitation will take place. What is clear, however, is that a new metric is needed that goes beyond the paradigm of 'continentality' or 'distance from the coast'. While straight-line distance from the coast is correlated with an isotope effect, it is a crude measure. What is more important is the path taken by a parcel of air after it leaves the influence of the ocean.

In some circumstances, the distance from the coast may be a good estimator. But as shown in Section 3, Antarctica's wind patterns can be chaotic. At the Polar Front, westerlies drive strong winds clockwise around the continent, while a number of local currents circulate around the interior. While a certain degree of consistency in averages is known to global weather patterns, the actual path taken by a parcel of air which precipitates out onto the continent is theoretically more important to the composition of surface snow.

It is very difficult to compute air trajectories, due to the chaos inherent in the atmospheric system, but the general idea is illustrated in the figure. I ran a NOAA HYSPLIT backward trajectory model for each WRE sampling site [31], computed at 6-hour intervals, to demonstrate how the path taken by air is usually very different to a straight line from the coast. Each path is 315 hours (¿13 days). Since the models are backwards-trajectory, they get more accurate as they approach the sampling site.

The reliability of the physical trajectory is less important than the overarching point: **air doesn't travel in a straight line**, so we should stop modeling it as such.

There are two reasons why this matters:

1. Arc Length: the stable isotope composition of precipitation is more likely a function of distance traveled than straight-line distance to the sampling site.

2. Elevation: δ values are partially a function of the *highest* elevation over which the air has traveled, not the elevation at the sampling site. Unpredictable and inconsistent air trajectories across the continent

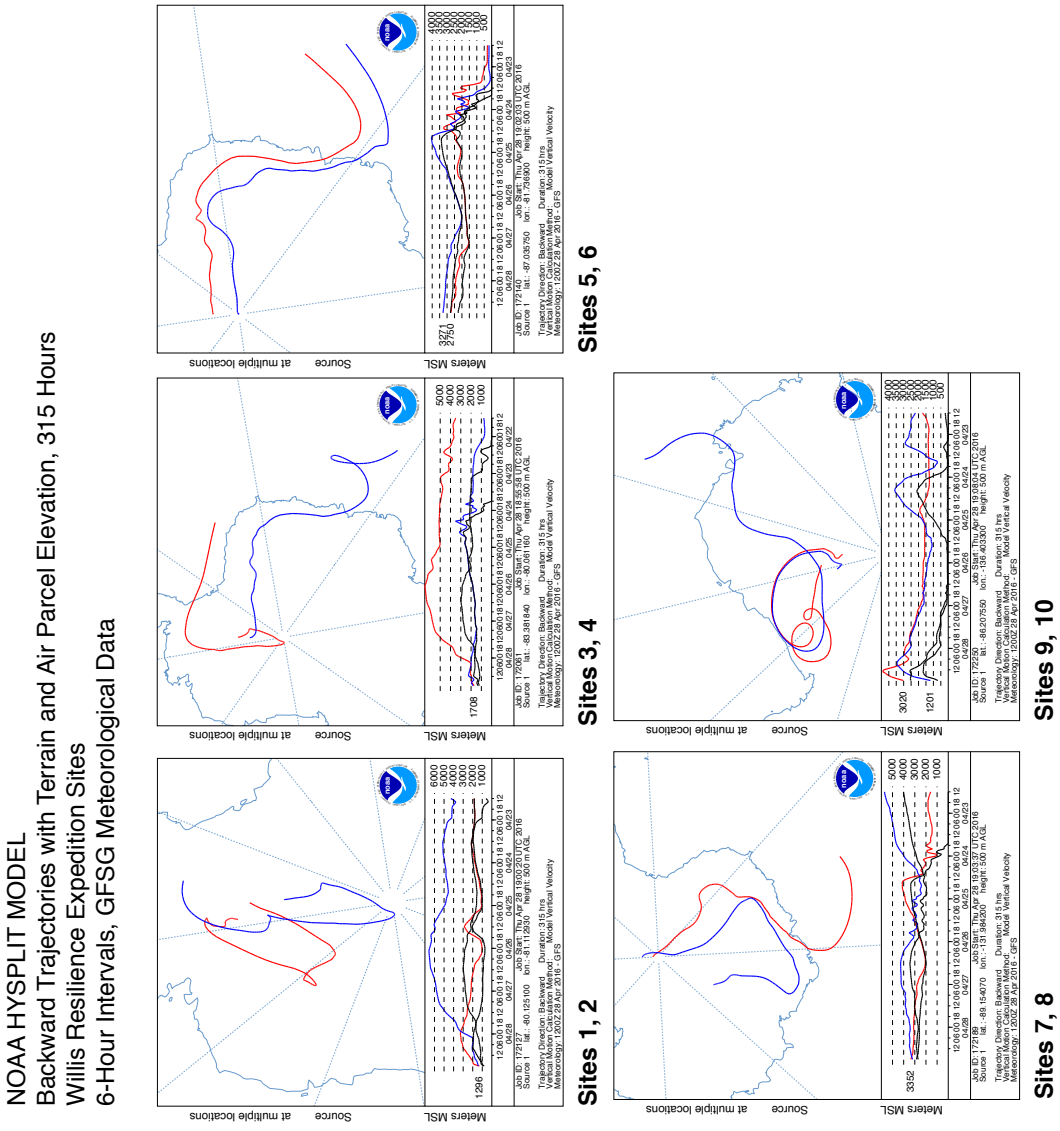

Figure 21: Backward Air Trajectories to each WRE Site

also means that the air travels over both low- and high-elevation areas, sometimes with very large changes; this exacerbates the problem of assuming either monotonicity of Antarctic terrain or that a two-way elevation influence is acceptable. This issue is discussed further later in this section.

Non-Parametric Arc Length I propose that the concept of 'distance from the coast', for Antarctica, be replaced with a new parameterization for the influence of distance traveled over dry land. This parameter is the **arc length** of a function of a two-dimensional path taken by an air parcel over the continent.

Let us define the path taken over the continent as a function in polar coordinates the same coordinate system previously define, but now only assessing the path in two dimensions where longitude is still θ_i and distance from the pole is still D_p. Let a and b be the start and end points for the air parcel, where a is a point on the coast of Antarctica and b is the site of precipitation. In general, we define the arc length [64] as L, such that:

$$L = \int_a^b \sqrt{D_p^2 + \left(\frac{dD_p}{d\theta}\right)^2} \tag{25}$$

However, stochasticity prevents us from finding a parametric solution to this problem, because there is no function that reliably maps the path of an air parcel to GPS points (which is the entire reason for this exercise).

Therefore, we once again rely on an Epanechnikov Kernel Analysis, where K and $\hat{f}_K$ are defined as they were in Section 5. For any trajectory with n data points (which are locations), in a plot between latitude (or distance from the pole) and longitude, the technique creates a smoothed non-parametric polynomial approximation, which we can consider accurate as long as the degree of the polynomial is less than or equal to $n - 1$. Thus the polynomial satisfies:

$$\mathbf{y} = m(x) + \epsilon$$

where m is determined by the kernel, a series of estimates for regression

coefficients, and a Taylor expansion based on the difference between x-values in each band in the path.

The arc length L is therefore:

$$L = \int_a^b \sqrt{1 + \left(\frac{dD_p}{d\theta_i}\right)^2} \tag{26}$$

Here, we have defined a 'non-parametric parameter', in that the scalar output of this non-parametric function becomes the basis for the continentality effect that replaces 'distance from the coast'.

There are obviously some practical limitations to this. Some atmospheric processes are difficult to predict long in advance. But as computing power increases, and models become more sophisticated, it might be possible for this issue to become less problematic.

Furthermore, we only need to compute backwards from the moment the air parcel moves onto land, and no further; the NOAA simulations shown here demonstrate that often this is not a very long period of time. It might also be possible to increase precision by running the same trajectory models many times and use average outputs of L for the parameter.

6.3 Collinearity of Variables

Distance from the coast and latitude are two different influencers on the stable isotopic composition of precipitation, and have different effects. The first is the combined effect of continentality and dryness; as a parcel of air moves across Antarctica, it can only become lighter, since there are no surface water sources. The second is a temperature effect from climates at lower latitudes.

Physically, these are different impacts, so it makes sense to separate them. Statistically, they are also different in most parts of the world. However, this does not apply to Antarctica, where latitude and continentality are in lock-step due to the shape and location (centered around the pole) of the continent. This has an impact on the relevance of a model that includes them both. Suppose you hold longitude, θ, constant. Then for each

θ, latitude and distance from the coast are exactly collinear, such that:

$$\lambda = D + c$$

where c is a constant, and λ and D are in one-to-one units. Masson et al. proposes models that take into account elevation, distance from the coast, and latitude [7]:

$$\delta D = \beta_0 + \beta_1 \sin(\lambda) + \beta_2 H + \beta_3 D$$

This is equivalent to:

$$\delta D = \beta_0 + (\beta_1 \sin(D + c) + \beta_3 D) + \beta_2 H = \beta_0 + (\beta_1 \sin(\lambda) + \beta_3(\lambda - c)) + \beta_2 H$$

which means you could theoretically leave out latitude or continentality from the regression entirely, depending on the parameterization, and get the same results. This does not align with what we know scientifically: latitude and continentality have different physical effects, and should be independent.

Normally, including an irrelevant variable does not bias regression models [8] (which would make either the extra λ or D harmless). This is not the same; the extra variable isn't irrelevant, it is instead given extra weight.

According to the model, the effect of horizontal distance is $\beta_1 \cdot \gamma_1$ where $\gamma_1 = \sin(D+c) + [\beta_3/\beta_1]D$, so that $[\beta_3/\beta_1]D - 1 \leq \gamma_1 \leq [\beta_3/\beta_1]D + 1$, where the inequality is dependent on D and varies with it. This is equivalent to:

$$\beta_3 D - \beta_1 \leq \beta_1 \gamma_1 \leq \beta_3 D + \beta_1$$

In practice, in Antarctica, the range of latitude cannot depart from $-66 \leq \lambda \leq -90$, which is equivalent to $-\frac{11}{30}\pi \leq \lambda \leq -\frac{1}{2}\pi$ in radians. Since:

$$\sin(-11\pi/30) = -0.914,$$
$$\sin(-\pi/2) = -1$$

$$\beta_3 D - \beta_1 \leq \beta_1 \gamma_1 \leq \beta_3 D - 0.914\beta_1 \tag{27}$$

Thus, the bias is formally quantified. This may not seem like a large effect; but it should be pointed out that in the Masson models, the β_1 coefficient was 151.59 for $\delta^{18}O$, and 944.46 for δ^2H; this is a significant bias for δ values.

As such, I argue that only one 'horizontal distance' parameter should be included in the model.

This issue shows that when dealing with real-world parameters in multi-dimensional space, multiple variables for each dimension can cause bias.

6.4 Compromised Elevation Variable

Post-deposition descriptive models, such as the one used in the Masson Data, make the assumption that Antarctica rises monotonically from the coast toward the pole. In the Masson study, a regression of geographic factors onto δ^2H gives [7]:

$$\delta D = 944.46 \sin(\lambda) - 0.057H - 0.0034D \tag{28}$$

The elevation effect is captured, but the *ceteris paribus* effect of a change in elevation implies that a parcel of air can recapture lost heavy isotopes upon a drop in elevation. This is incorrect, as air can only be 'recharged' with heavy isotopes where there is a reservoir from which it can draw, and there is no liquid water on the surface of inland Antarctica. Any model hoping to capture the effect of elevation must also include the non-linearity of the relationship, in that no matter the change in elevation, $\delta^{18}O$ *and* δD *must decrease monotonically* from the coast of Antarctica toward the pole. The Masson explanatory regressions are therefore not suitable where there are local elevation extrema which stray significantly from the average elevation path from the coast to the pole.

This can be demonstrated qualitatively with **Vostok Station**, which lies at 3,488m (691m higher than the South Pole station) [65]: The Masson Dataset reports that Vostok shows mean δD values of $-440.05‰$ and mean $\delta^{18}O$ values of $-56.50‰$. However, the model predicts a δD of $-1152‰$, which is an extraordinary deviation from the truth. The same applies for the predicted value of $\delta^{18}O$ of $-175.6‰$, which is equally incorrect.

Parameterization Using DEM Data NSIDC provides access to Digital Elevation Model data for the Antarctic continent, at resolutions of 1km, 400m, and 200m [63]. Software-generated mappings are useful for visualization but not useful in data analysis; in order to assess the practicality of an approach that uses a function to describe a regression parameter (in this case elevation), I generated a CSV file from the raw DEM data at 1-km resolution. A cross-section can be generated by plotting latitude against elevation for each unique θ_i. This is possible at higher resolutions but this is not necessary for the purposes of this analysis. The critical purpose of this section is the use of data points to approximate a function to describe elevation.

For a third time, we can do this using kernel-weighted polynomial smoothing, and again, to optimize efficiency, we use Epanechnikov. I generated examples from three different parts of the continent are shown in Figure 22.

One issue that is is immediately apparent is that the surface of Antarctica is indeed not monotonic, especially given that the Geographic South Pole is not at the pinnacle of an idealized cone but offset from the highest points on the continent.

δ values should more or less track elevation as it rises, and stay flat when elevation decreases. As previously discussed, however, the elevation that matters is not a cross-section from the coast to the pole, but a cross section of Antarctica in the path taken by the air parcel containing the moisture. In the NOAA HYSPLIT simulations, terrain and air elevation cross-sections are shown below the path taken for each site [31]. In most cases, the air passes over terrain that is higher than the eventual site of deposition, which is to be expected.

We can define a new parameter E which is a function of longitude, distance from the pole, and Elevation (H), and where the value of E is simply the elevation at a specific time in the trajectory taken by the air parcel. The black lines in theH HYSPLIT cross-sections represent E. The parameter for our use is therefore $max\{E\}$.

This is relatively easy to compute because Digital Elevation Models are available which can provide a value of elevation at 200m resolution; unlike the arc of the air trajectory, there is no uncertainty or even a need to approximate a function, we simply have to choose the highest value.

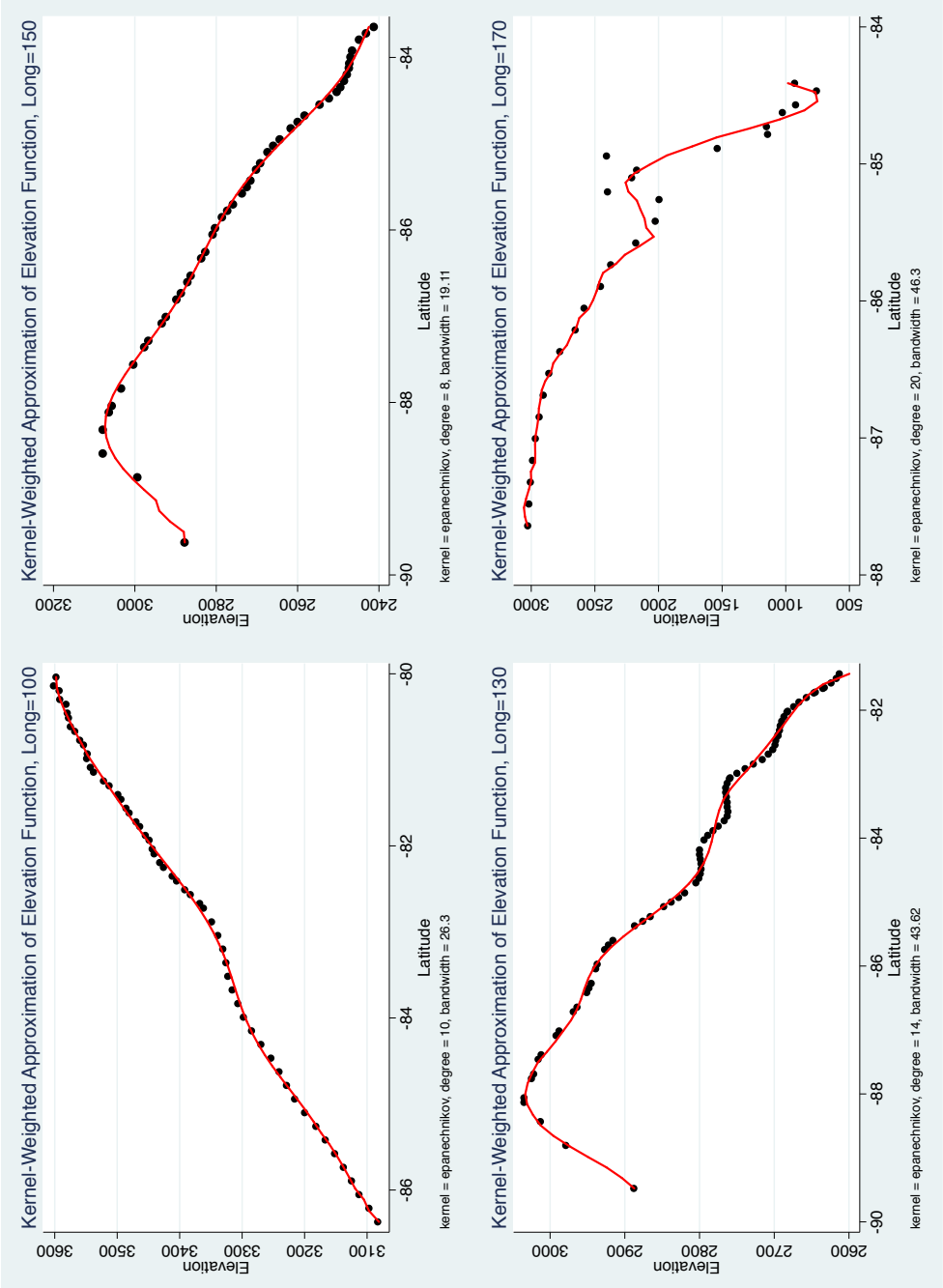

Figure 22: Four Elevation Cross-Sections of Antarctica with different θ_i, Epanechnikov Kernel-Weighted Polynomial Approximation [63]

Table 8: p-values for Local Meteoric Water Line Coefficients, WRE Sites

Site	1	2	3	4	5	6	7	8	9	10
β_0 P>$\mid t \mid$	0.00	0.053	0.545	0.838	0.848	0.007	0.031	0.753	0.772	0.653

One caveat is the practical consideration of computing power. The highest-resolution DEM available for Antarctica has more than 56 million observations [63]. Computation speeds are slow on modern ordinary computers for even the 1km-resolution model (14 million observations), so the kernel-weighted approximation may be better for some users, or for computing many outputs simultaneously (where computation time adds up).

Error in Old Parameters With the new parameters, the new optimization function becomes:

$$M_i(\varepsilon) \equiv J(L, E + \varepsilon\eta(D_p), E' + \varepsilon\eta(D_p)).$$

Such that the error of the old 'ideal' distance parameter is $M_i(\varepsilon)/I_i(\varepsilon)$, which by construction is always greater than 1 because I approximates the most efficient elevation and the most efficient distance from the coast.

6.5 Matrix Form for Geo-specific Analysis

We should also aim to answer the question of whether developing a vector approach which separates sites and regresses 'sets of variables' on location. To analyze this, we can only use the WRE data, since the test requires a full dataset for each site.

Site-specific analysis returned mixed results. While it's true that an array of variables is a more powerful basis for statistical analysis than a single functional form for the whole continent, this study is limited by small sample properties not in the sense that 'small' has been referred to throughout this paper (i.e. smaller than the Masson Dataset), but in the technical sense. With a sample size of 20 per site, OLS isn't reliable and the 95% confidence interval often crosses 0 (i.e. the parameter could be negative or positive).

An example is shown in the table, which displays the p-values for Local Meteoric Water Lines at each site. Only three sites out of ten showed statistically significant results, i.e. where $p < 0.05$. Even regardless of this number, the technical details of small sample properties prohibit us from relying on the standard errors of datasets with so few samples [66] (standard errors are the basis for confidence intervals and some tests of significance).

With full datasets for each of the sites in the Masson Data, this would be less of a problem, but since these data are not available, regressions of vector-valued functions will have to be omitted from this paper.

6.6 Model Assimilation

We have now deconstructed the paradigm for statistical inference of the spatial variability of stable isotopes in Antarctica, and suggested new methods of parameterization. Each of these can be summarized below.

6.6.1 New Parameters

Elevation The elevation parameter becomes the maximum elevation over which the air has traveled, $max\{E\}$.

Distance from the Coast The continentality parameter becomes Arc Length of the 2-Dimensional function of distance from the pole and longitude: L.

Latitude & Collinearity Problem From our construction using the calculus of variations, the Latitude factor becomes absorbed as an exacerbating factor in the continentality variable, to avoid the problem of perfect collinearity.

Estimator Construction I concluded in Section 4 that OLS was not efficient because of heteroskedasticity. To obtain heteroskedasticity-consistent regression outputs, we use Weighted Least Squares; in the event of homoskedasticity, it may sacrifice some efficiency, but in general we should

operate under the assumption that the results on a continent-wide scale will be heteroskedastic, as we know them to be in the Masson Data. The WLS estimator is similar to OLS, but in addition to the general Residual Sum of Squares minimization problem solved by OLS, WLS uses a weighting multiple that is the inverse of the variance at each observation i: $\min \sum_i^n \frac{1}{\sigma_i^2}\varepsilon^2$ [67].

6.6.2 New Model

After assimilation, the new model for predicting the spatial variability of stable isotopes in Antarctic precipitation becomes:

$$\delta^{18}O = \beta_0 + \beta_1 L + \beta_2 \max E + \varepsilon \tag{29}$$

Where $\beta = \hat{\beta} + u_i$, and where u_i is an error, and $\hat{\beta}$ is the Weighted Least Squares estimator.

This simple equation summarizes the purpose of this analysis.

7 Conclusions

There were several purposes to this analysis. One was to report new data from Antarctica, test their quality, and use them to show the unique properties of Antarctic precipitation, where δ values are extremely depleted compared to the global hydrological system. A second was to explore the possibility of using statistical methods to more effectively predict water isotope in Antarctica while harnessing empirical evidence as much as possible, and *without* using integrated Rayleigh-simulating models, which can't take into consideration the chaotic nature of meteorology or take into consideration specific precipitation events.

In analyzing the WRE samples, statistical tests revealed that the samples are sufficiently high-quality to consider them accurate, but they have limited use on their own. Future research should aim to take a step beyond the

Masson Data and compile complete datasets into a shared database, rather than using averages to map the spatial variability of isotopes. Of course, this would be an arduous and difficult task, but given the usefulness of water isotope data in climate science and meteorology, and the uniqueness of Antarctica, it may be worth the trouble to bring together leading experts to do so.

More broadly, I have learned that it is possible to harness statistical methods and real analysis to attempt an accurate prediction. It would be possible to go much further than I have done in this paper, for example by mapping seasonal or monthly wind patterns to model the trajectory of an air parcel during a specific part of the year (which we can date by computing accumulation rates). The path of the air parcel remains the single most uncertain component of this analysis, but with more focused, more powerful simulation methods and improving computing power, it may be possible to alleviate some of the concerns associated with the approximation uncertainty.

We are limited by the lack of measured meteorological data and also by the low accumulation rates in Antarctica. The lack of surface stations made comparison of the isotope temperature reconstructs against local weather results impossible for most sites, and it was a stretch to compare Site 10 to Station Elaine. The low accumulation rates and high surface wind speeds (katabatic winds) means that the near-surface isotope record is not very well-preserved. Post-depositional processes disturb the layers, and the fact that precipitation rates are low means that even a minor disturbance can remove a whole season or even a year from the record. This problem is worse in East Antarctica, where elevation is higher and the accumulation rates even lower than the two WRE transects.

These are long-term issues that no single research program could solve alone, but as an aside, other aspects of the the Willis Resilience Expedition's research program aimed to take steps towards solving these problems. We tested a new version of a lightweight automatic weather station, built especially for the Polar Regions by a member of the scientific team, Marc Cornelissen. Secondly, the problems of low accumulation can partially be mitigated by having more accurate recent dating techniques that aren't prohibitively expensive. We are analyzing the samples for Tritium content in order to map its deposition rate across Antarctica. This would allow us to reconstruct the age of samples to within seasonal resolution without having

to rely on counting annual layers alone.

Overall, however, the Masson Data provides a large amount of information from which we can derive very useful information about the behavior of Antarctica's meteorological system. Datasets will only become more easily available and analyzable as the digital access challenges faced by the last generation of scientists fade away. We should aim to harness statistical methods to their maximum potential in analyzing future empirical results.

8 End Matter

8.1 Statement on Funding & Interests

Funding was provided by the private-sector third parties listed in the Acknowledgements, but this funding was independent of the goals of the research, and changes to the program were made entirely at the discretion of the expedition team and the Research Advisory Committee, without the approval of non-expert funders.

Scientists and lab technicians at GNS New Zealand processed the samples and generated the WRE dataset. Beyond that, however, this paper was written without their help or input.

8.2 Appendix

8.2.1 Construction of the F-Statistic

The F-statistic used for the test is computed from a Wald test. Let $\mathbf{V}$ be the variance-covariance matrix, $\mathbf{b}$ the estimated coefficient vector, and $\mathbf{Rb} = \mathbf{r}$ the set of jointly-tested hypotheses [47]. Then the Wald statistic is given by [48]:

$$W = (\mathbf{Rb} - \mathbf{r})'(\mathbf{RVR}')^{-1}(\mathbf{Rb} - \mathbf{r})$$

And the F-statistic is given by [**?**]:

$$F = \frac{1}{q} \cdot W$$

Where q is the numerator degrees of freedom, which in this case is 1.

8.2.2 Epanechnikov Kernel

The kernel and related functions are defined as follows [53].

$$K_\alpha[z] = \frac{\frac{3}{4}(1 - \frac{1}{5}z^2)}{\sqrt{5}}$$

if $|z| < \sqrt{5}$, and 0 otherwise; where z is a quantile of the standard Gaussian Distribution [75]. Now let us define a kernel density estimate $\hat{f}_{K\alpha}$ such that [**?**]:

$$\hat{f}_{K\alpha} = \frac{1}{h}\sum_{i=1}^{n} w_i K\left(\frac{x - X_i}{h}\right)$$

where x is the variable in question (depth), n is the number of observations, h is the 'window width', and w_i is a weighting factor [76]. Note that f is real-valued: $f : \mathbb{R} \longrightarrow \mathbb{R}$. The eventual model estimates [77]:

$$\mathbf{y} = m(x) + \epsilon$$

where $\hat{m}(x_0)$ is a local approximation of polynomial of order p, which is equal to a vector of locally-weighted regression coefficients: $\hat{\beta} = [\hat{\beta}_0, \hat{\beta}_1, ...\hat{\beta}_p]^T$. The weighting depends on K, and the overall function depends on m and f.

8.2.3 WRE Chronology

Tables 9 and 10 show the chronology of the WRE samples.

Table 9: Chronology of Stable Isotopes in Antarctic Precipitation, WRE

Time		$\delta^{18}O$, (‰)									
Month	Year	1	2	3	4	5	6	7	8	9	10
December	2013	-27.65	-30.2	-31.72	-39.98	-41.21				-37.05	-32.52
November	2013					-40.97		-50.34	-42.1		
October	2013	-27.62	-31.43	-31.11	-38.19	-41.14		-49.41			
September	2013		-31.13			-44.88	-52.16			-42.59	
August	2013	-30.12	-38.99	-33.57	-42.62	-48.57			-41.53		
July	2013		-35.78	-34.87		-50.22	-52.57	-52.03		-45.53	-38.42
June	2013	-35.07	-33.39		-45.24	-50.79		-51.87	-41.73		
May	2013			-35.52		-51.17		-50.3			
April	2013	-32.74	-33.34		-44.92	-51.56			-40.96		
March	2013			-35.66		-50.19	-51.3	-49.86			
February	2013	-29.19	-33.37		-39.89	-45.39			-40.95		
January	2013		-30.15	-35.06		-40.86		-47.57			-29.04
December	2012	-35.03	-33.8		-38.96	-40.04	-48.98				
November	2012		-30.68	-35.25		-38.73		-44.91	-40.86	-34.76	-28.4
October	2012		-30.16		-41	-38.96			-40.55		
September	2012		-33.52	-37.24		-41.72	-50.76	-45.87	-40.95	-37.95	
August	2012	-36.33	-33.44		-45.28	-45.59			-45.26		-31.25
July	2012		-34.27	-39.06			-50.75	-53.93	-49.48	-46.27	
June	2012		-33.78		-44.41						
May	2012		-34.97	-39.47					-49.52	-39.12	-33.68
April	2012	-29.09			-42.55			-52.55	-43.72		
March	2012		-35.89	-37.94			-49.56			-40.28	-32.53
February	2012		-33.84		-39.52			-51.22	-42.49		
January	2012		-33.67	-32.98					-42.35	-42.35	-27.53
December	2011	-33.17	-39.76	-37.47	-37.99						
November	2011				-38.51		-49.76	-45.45	-42.28	-40.42	-26.29
October	2011	-31.32			-39.9			-47.7	-44.5		
September	2011	-30.13			-40.05			-49.57	-44.82		-28.54
August	2011	-30.71		-36.93	-40.15		-49.87	-51.45		-39.22	
July	2011	-30.22			-40.42			-52.45			-29.35
June	2011				-41.81		-49.66				
May	2011	-30.14								-38.79	-33.64
April	2011	-34.91		-35.15							
March	2011	-33.19					-48.86				-34.18
February	2011									-43.15	
January	2011	-32.2									-31.52
December	2012	-36		-39.09	-40.12		-48.58				
November	2012									-43.13	-30.31
October	2012						-49.87			-42.48	
September	2012						-51.1			-42.54	
August	2012						-53.56				-33.12
July	2012			-40.47			-53.72			-44.73	
June	2012						-52.32				
May	2012										
April	2012						-51.5				-32.05
March	2012						-51.26				
February	2012										
January	2012										-29.63
December	2009									-38.29	
November	2009			-35.67						-38.04	
October	2009										-29.93
September	2009									-41.62	
August	2009										-31.16

Table 10: Chronology of Reconstructed Monthly Temperature Averages, WRE Sites

Time		Surface T (^{o}C)		Reconstructed T (^{o}C)									
Month	Year	AS	Elaine	1	2	3	4	5	6	7	8	9	10
December	2013	-26.2	-7.1	-29.36796866	-30.6942	-31.48612359	-35.78707607	-36.42946875				-34.26400197	-31.90289998
November	2013	-35.6	-14.7					-36.30650437		-41.18490569	-36.89333538		
October	2013	-51.2	-22.7	-29.35109407	-31.33503	-31.16955833	-34.85913874	-36.39621902		-40.70215728			
September	2013	-51.2	-30.9		-31.17873			-38.34418942	-42.13505648			-37.14890193	
August	2013	-53.3	-34.6	-30.65303358	-35.27379	-32.44919912	-37.16594923	-40.26540174			-36.59745317		
July	2013	-58.6	-32.9		-33.60138	-33.12838769		-41.12251866	-42.35035519	-42.06929294		-38.68344612	-34.97780856
June	2013	-52.6	-35	-33.23346126	-32.35619		-38.53178957	-41.41909376		-41.98366955	-36.70378583		
May	2013	-62.4	-29.3			-33.46509333		-41.61973897		-41.16801368			
April	2013	-57.6	-35.3	-32.01797157	-32.33014		-38.36496938	-41.82057308			-36.299583		
March	2013	-50.4	-20.3			-33.53744649		-41.10999944	-41.68662071	-40.93470179			
February	2013	-43.8	-15.4	-30.1680047	-32.34577		-35.74304464	-38.60828701			-36.29440506		
January	2013	-26.9	-6		-30.66815	-33.22850038		-36.24814037		-39.74339851			-30.08742986
December	2012	-25.3	-5.8	-33.20902564	-32.5698		-35.2589864	-35.81852615	-40.4765874				
November	2012	-36.2	-15.3		-30.94428	-33.32629213		-35.13804569		-38.35666677	-36.24960065	-33.0684101	-29.75707199
October	2012	-51.5	-22.9		-30.67336		-36.3188485	-35.26008337			-36.08836126		
September	2012	-60.7	-28.9		-32.42392	-34.36204872		-36.69530199	-41.40492377	-38.85741179	-36.29482973	-34.73331584	
August	2012	-62.7	-34.6	-33.88927438	-32.38224		-38.55286302	-38.71321642			-38.53878068		-31.24028875
July	2012	-60.6	-30.4		-32.81467	-35.3097645			-41.40161992	-43.05936651	-40.7386437	-39.06832178	
June	2012	-58	-33.7		-32.55938		-38.09718151						
May	2012	-58.7	-22.5		-33.17937	-35.52290207					-40.76181531	-35.3435598	-32.50889745
April	2012	-59.5	-25.7	-30.11690707			-37.12670307			-42.33654389	-37.74062342		
March	2012	-51	-22.3		-33.65869	-34.72745919			-40.7811524			-35.94420351	-31.91066148
February	2012	-40.2	-13.5		-32.59064		-35.54878725			-41.64353489	-37.09665976		
January	2012	-28.6	-8.1		-32.50207	-32.14034752					-37.02231359	-37.02565943	-29.30182723
December	2011	-27.4	-5.9	-32.24046242	-35.67496	-34.48209612	-34.75332516						
November	2011	-34.9	-11.8				-35.02151448		-40.88377461	-38.64202402	-36.98976402	-36.01703133	-28.65822403
October	2011	-48.7	-20.2	-31.27876621			-35.74738177			-39.81088806	-38.14279226		
September	2011	-59.6	-34.2	-30.65788817			-35.82440069			-40.78342983	-38.31293929		-29.82704319
August	2011	-56.5	-25.4	-30.95981249		-34.1995176	-35.87724797		-40.94193183	-41.76310968		-35.39559033	
July	2011	-58.2	-23	-30.70567475			-36.01817406			-42.28724759			-30.25061497
June	2011	-61.5	-25.9				-36.7410727		-40.83281887				
May	2011	-62.6	-25.9	-30.66205863								-35.17058952	-32.48847942
April	2011	-59.9	-27.8	-33.1472278		-33.27360264							
March	2011	-56.2	-25.5	-32.25187746					-40.41413116				-32.76650273
February	2011	-37.1	-15.4									-37.4410396	
January	2011		-8.6	-31.73712748									-31.38316399
December	2012	-26.7	-7.5	-33.71441411		-35.32588122	-35.86315536		-40.26842046				
November	2012	-38.1	-13.4									-37.43198695	-30.75213606
October	2012	-49.6	-22.6						-40.94250875			-37.09137534	
September	2012	-61.7	-32.9						-41.58428061			-37.12322017	
August	2012	-57.7	-30.6						-42.86476				-32.21423451
July	2012	-58.7	-33.9			-36.04744436			-42.94812			-38.26296998	
June	2012	-60.5	-30.8						-42.21932266				
May	2012	-54.1	-30.9										
April	2012	-62	-25.7						-41.79123925				-31.65840155
March	2012	-50.4	-26.3						-41.66584108				
February	2012	-37.9	-14.1										
January	2012	-25.5											-30.39637037
December	2009											-34.90683565	
November	2009					-33.54638977						-34.78028455	
October	2009												-30.55371243
September	2009											-36.64529043	
August	2009												-31.1953787

References

[1] Scientific Committee for Antarctic Research (2015). Antarctic Digital Database. Retrieved from http://bit.ly/1SZxScC

[2] University of Washington (NA). Atmospheric greenhouse gas concentrations from the ice core record. Retrieved from http://bit.ly/1OcpRRw.

[3] P. Ciais, J. Jouzel (1994). Deuterium and oxygen 18 in precipitation: Isotopic model, including mixed cloud processes. Journal of Geophysical Research: Atmospheres. Vol. 99, Iss. D8.

[4] J. Jouzel, L. Merlivat (1984). Deuterium and oxygen 18 in precipitation: Modeling of the isotopic effects during snow formation. Journal of Geophysical Research: Atmospheres. Vol. 89, Iss. D7.

[5] L. Merlivat, J. Jouzel (1979). Global climatic interpretation of the deuterium-oxygen 18 relationship for precipitation. Journal of Geophysical Research: Oceans. Vol. 84, Iss. C8.

[6] Utrecht University (2015). Institute for Marine and Atmospheric Research (IMAU): Ice and Climate Regional Modeling. Retrieved from http://bit.ly/1T1EHan.

[7] V. Masson-Delmotte et al. (2008). A Review of Antarctic Surface Snow Isotopic Composition: Observations, Atmospheric Circulation, and Isotopic Modeling. Journal of Climate. Vol. 21, Iss. 13.

[8] J. Wooldridge (2012). Econometrics: A Modern Approach. ISBN 9781111531041.

[9] NA

[10] Scientific Committee for Antarctic Research (2015). Antarctic Digital Database. Retrieved from http://bit.ly/1rMt4gj.

[11] International Atomic Energy Agency Water Resources Programme (NA). Environmental Isotopes in the Hydrological Cycle: Principles and Applications. Volume 1: Introduction, Theory, Methods, Review. Chapter 3: Abundance and Fractionation of Stable Isotopes.

[12] International Atomic Energy Agency Water Resources Programme (NA). Environmental Isotopes in the Hydrological Cycle: Principles and Applications. Volume 2: Atmospheric Water.

[13] Rabia Naeem (NA). Lennard-Jones Potential. UC Davis. Retrieved from http://bit.ly/1Tov9Fs.

[14] University of Colorado Boulder (NA). Interatomic and Intermolecular Forces. Retrieved from http://bit.ly/1SZxXx1.

[15] University of Colorado Boulder (NA). Interatomic and Intermolecular Forces. Retrieved from http://bit.ly/1SZxXx1.

[16] University of Colorado Boulder (NA). Interatomic and Intermolecular Forces. Retrieved from http://bit.ly/1SZxXx1.

[17] Theochem.org (NA). EFN316G: Introduction to Quantum Chemistry. Retrieved from http://bit.ly/1Tyni8W.

[18] W. M. White (2013). Stable Isotope Geochemistry. Retrieved from http://bit.ly/24HfuJD.

[19] W. M. White (2013). Stable Isotope Geochemistry. Retrieved from http://bit.ly/24HfuJD.

[20] National Institute of Standards and Technology (2014). Fundamental Physics Constants. Retrieved from http://1.usa.gov/1oJxUV5.

[21] International Atomic Energy Agency Water Resources Programme (NA). Environmental Isotopes in the Hydrological Cycle: Principles and Applications. Volume 1

[22] University of Oxford (NA). Statistical Mechanics. Retrieved from http://bit.ly/1SZy1wP.

[23] University of Oxford (NA). Statistical Mechanics. Retrieved from http://bit.ly/1SZy1wP.

[24] C. Kendall, E. A. Caldwell (1998). Fundamentals of Isotope Geochemistry. United States Geological Survey. Retrieved from http://on.doi.gov/1rMtGmb.

[25] R. Nave (NA). Ideal Gas Law. Georgia State University. Retrieved from http://bit.ly/1XhJauM.

[26] U. Siegenthaler, H. Oeschger (1980). Correlation of 18O in precipitation with temperature and altitude. Nature, 285, pp. 314-317.

[27] Z. Liu et al. (2014). Stable Isotopic Compositions of Precipitation in China. Tellus B, Chemical and Physical Meteorology. Vol 66.

[28] Yu, J. et al (1980). Oxygen isotopic composition of meteoric water in the eastern part of Xizang. Geochimica. 2, 113121. (in Chinese) [referenced by Liu et al. 2014].

[29] American Museum of Natural History (NA). Retrieved from http://bit.ly/1TynmFG.

[30] T. Felis, J. Ptzold (2004). Climate Reconstructions from Annually Banded Corals. Global Environmental Change in the Ocean and on Land, Eds. M. Shiyomi et al., pp 205 − 227.

[31] National Oceanic & Atmospheric Administration (2015). HYSPLIT Hybrid Single Particle Lagrangian Integrated Trajectory Model. Retrieved from http://1.usa.gov/1SZy8s4.

[32] NA

[33] NOAA National Centers for Atmospheric Prediction (NCEP) (2016). Retrieved from http://1.usa.gov/1T7C4pJ.

[34] C. Beccario (2016). Visualization of Global Weather Patterns. Retrieved from http://earth.nullschool.net/.

[35] American Meteorological Society (2012). Meteorological Glossary. Retrieved from http://glossary.ametsoc.org/

[36] World Meteorological Organization (NA). Antarctic Region. Retrieved from http://bit.ly/1T7C74J.

[37] Intergovernmental Panel on Climate Change (2007). The Southern Hemisphere and Southern Annular Mode. Retrieved from http://bit.ly/1rMu9ES.

[38] V. Masson-Delmotte (2016). Laboratoire des Sciences du Climate et de l'Environnement. Retrieved from http://bit.ly/24HfCbX.

[39] Scientific Committee for Antarctic Research (2015). Antarctic Digital Database. Retrieved from http://bit.ly/1rMt4gj.

[40] C. Lorius, L. Merlivat (1977): Distribution of mean surface stable isotope values in East Antarctica, Observed changes with depth in a

coastal area. Isotopes and Impurities in Snow and Ice, Proceedings of the Grenoble Symposium, Vol. 118, pp.125137. [Referenced by Jouzel 2013, http://bit.ly/1rDpl54]

[41] J. Wooldridge (2012). Econometrics: A Modern Approach. ISBN 9781111531041.

[42] St. Francis Xavier University (NA). Violations of OLS Assumptions: Heteroskedasticity. Retrieved from http://people.stfx.ca/tleo/econ370term2lec1.pdf.

[43] Google Earth.

[44] K. Rozanski et al. (1993). Isotopic Patterns in Modern Global Precipitation. Retrieved from http://bit.ly/1QWKcVj.

[45] J. Hoefs (2015). Stable Isotope Geochemistry, 7th Edition. Springer.

[46] Stata (2016). Skewness and kurtosis test for normality. Retrieved from http://www.stata.com/manuals13/rsktest.pdf.

[47] Stata (2016). Test linear hypotheses after estimation. Retrieved from http://www.stata.com/manuals13/rtest.pdf.

[48] Stata (2016). Test linear hypotheses after estimation. Retrieved from http://www.stata.com/manuals13/rtest.pdf.

[49] K. Casey et al. (2014). The 1500 m South Pole ice core: recovering a 40 ka environmental record. Annals of Glaciology. Vol. 55, Iss. 68.

[50] A. Svensson et al. (2015). On the occurrence of annual layers in Dome Fuji ice core early Holocene ice. Clim. Past. Vol 11, pp. 1127-1137.

[51] Stata (2016). Kernel-weighted local polynomial smoothing. Retrieved from http://www.stata.com/manuals13/rlpoly.pdf.

[52] Max Planck Institute (NA). Introduction to Kernel Smoothing. Retrieved from http://bit.ly/1q9MgDA.

[53] Stata (2016). Kernel-weighted local polynomial smoothing. Retrieved from http://www.stata.com/manuals13/rlpoly.pdf.

[54] Masson-Delmotte (2016). Laboratoire des Sciences du Climate et de l'Environnement. Retrieved from http://bit.ly/24HfCbX.

[55] British Antarctic Survey (NA). Met READER. Retrieved from https://legacy.bas.ac.uk/met/READER/data.html.

[56] National Snow & Ice Data Center (NA). Retrieved from http://nsidc.org/data/search/.

[57] NA.

[58] National Oceanic & Atmospheric Administration (NA). Retrieved from https://www.ncdc.noaa.gov/data-access.

[59] NASA Goddard Institute for Space Studies (2016). GISS Surface Temperature Analysis (GISSTEMP). Retrieved from http://data.giss.nasa.gov/gistemp/stdata/.

[60] Wolfram (2016). Spherical Coordinates. Retrieved from http://mathworld.wolfram.com/SphericalCoordinates.html.

[61] C. Rousseau, Y. Saint-Aubin (2008). Calculus of Variations and Applications. University of California, Berkeley. Retrieved from https://math.berkeley.edu/ strain/170.S13/cov.pdf.

[62] M. Reed (1998). Fundamental Ideas of Analysis. ISBN 0471159964.

[63] National Snow & Ice Data Center (2016). Antarctic 1km Digital Elevation Model (DEM) from Combined ERS-1 Radar and ICESat Land Satellite Altimetry. Retrieved from http://bit.ly/1Tyntkz.

[64] Lamar University (NA). Arc Length with Polar Coordinates. Retrieved from http://tutorial.math.lamar.edu/Classes/CalcII/PolarArcLength.aspx.

[65] Russian Antarctic Expedition (2016). Station Vostok. Retrieved from http://bit.ly/1YiqCZx.

[66] University of California, Berkeley (NA). Small-Sample Properties of IV and OLS Estimators. Retrieved from http://bit.ly/1T71arR.

[67] J. Wooldridge (2012). Econometrics: A Modern Approach. ISBN 9781111531041.

[68] A. Jasper, J. Miller (2013). Lennard-Jones parameters for combustion and chemical kinetics modeling from full-dimensional intermolecular potentials. Combustion and Flame. Retrieved from http://1.usa.gov/1ZylFfw.

[69] International Atomic Energy Agency Water Resources Programme (NA). Environmental Isotopes in the Hydrological Cycle: Principles and Applications. Volume 1: Introduction, Theory, Methods, Review. Chapter 3: Abundance and Fractionation of Stable Isotopes.

[70] Cornell University (2003). Stable Isotope Theory: Equilibrium Fractionation. Retrieved from http://bit.ly/1TAxuxl.

[71] L. Lee (NA). California Polytechnic, Pomona. Values of R (Gas Constant). Retrieved from http://www.csupomona.edu/ lllee.

[72] Columbia University (NA). Stable Isotopes. Retrieved from http://bit.ly/1T71dUr.

[73] Stata (2016). Skewness and kurtosis test for normality. Retrieved from http:// www.stata.com/manuals13/rsktest.pdf.

[74] A. Svensson et al. (2015). On the occurrence of annual layers in Dome Fuji ice core early Holocene ice. Clim. Past. Vol 11, pp. 1127 − 1137.

[75] Stata (2016). Test linear hypothesis after estimation. Retrieved from http:// www.stata.com/manuals13/rtest.pdf.

[76] Stata (2016). Univariate kernel density estimation. Retrieved from http://www.stata.com/manuals13/rkdensity.pdf.

[77] Stata (2016). Kernel-weighted local polynomial smoothing. Retrieved from http:// www.stata.com/manuals13/rlpoly.pdf#rlpoly.