



# Reconstructing atmospheric H<sub>2</sub> over the past century from bi-polar firn air records

John D. Patterson<sup>1</sup>, Murat Aydin<sup>1</sup>, Andrew M. Croftwell<sup>2,3</sup>, Gabrielle Pétron<sup>2,3</sup>, Jeffery P. Severinghaus<sup>4</sup>,  
5 Paul B. Krummel<sup>5</sup>, Ray L. Langenfelds<sup>5</sup>, Vasilii V. Petrenko<sup>6</sup>, and Eric S. Saltzman<sup>1</sup>

<sup>1</sup>Department of Earth System Science and Chemistry, University of California, Irvine, Irvine, CA 92697, USA

<sup>2</sup>Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Boulder, CO 80309, USA

<sup>3</sup>Global Monitoring Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO 80305, USA

<sup>4</sup>Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093, USA

10 <sup>5</sup>Climate Science Centre, Commonwealth Scientific and Industrial Research Organisation, Environment, Aspendale, Victoria  
3195, Australia

<sup>6</sup>Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY 14627, USA

*Correspondence to:* John D. Patterson (jdpatter@uci.edu)

## Abstract

15 Historical hemispheric atmospheric H<sub>2</sub> levels since 1930 were reconstructed using the UCI\_2 firn air model and firn air  
measurements from three sites in Greenland: (NEEM, Summit, and Tunu) and two sites in Antarctica (South Pole and  
Megadunes). A joint reconstruction based on the two Antarctic sites yields H<sub>2</sub> levels monotonically increasing from about  
350 ppb in 1900 to 550 ppb in the late 1990's, levelling off thereafter. These results are similar to individual reconstructions  
published previously (Patterson et al., 2020; 2021). Reconstruction of the Greenland data is complicated by a systematic  
20 bias between Tunu and the other sites. The Tunu reconstruction shows substantially lower historical H<sub>2</sub> levels than the other  
two sites, a difference we attribute to possible bias in the calibration of the Tunu measurements. All three reconstructions  
show a late 20<sup>th</sup> century maximum in H<sub>2</sub> levels over Greenland. A joint reconstruction of the Greenland data shows H<sub>2</sub>  
levels rising 40% from 1930-1990, reaching a maximum of 550 ppb. After 1990, reconstructed atmospheric H<sub>2</sub> decrease by  
6% over the next 20 years. The reconstruction deviates by at most 4% from the few available surface air measurements of  
25 atmospheric H<sub>2</sub> levels over Greenland from 1998-2004. However, the longer instrumental records from sampling sites  
outside of Greenland show a more rapid decrease and stabilization after 1990 compared to the reconstruction. We explore  
the possibility that this difference is an artefact caused by the firn air model underestimating pore close-off induced  
enrichment, evidenced by a mismatch between measured and modelled Ne in firn air. We developed new parameterizations  
which more accurately capture pore close-off induced enrichment at the Greenland sites. Incorporating those  
30 parameterizations into the UCI\_2 model yields reconstructions with lower H<sub>2</sub> levels throughout the mid-late 20<sup>th</sup> century and  
more stable H<sub>2</sub> levels during the 1990's, in better agreement with the flask measurements.



## 1 Introduction

Molecular hydrogen ( $H_2$ ) is the second most abundant reduced gas in Earth's atmosphere after methane. Atmospheric  $H_2$  levels are linked to Earth's radiative budget, air quality, and the atmospheric oxidative capacity (eg. Ehhalt & Rohrer, 2009; Novelli et al., 1999; Paulot et al., 2021). As  $H_2$  becomes a more important component of the energy sector, anthropogenic emissions of  $H_2$  are expected to rise substantially (Derwent et al., 2020; Prather, 2003; Wang et al., 2013a, 2013b). Projecting the atmospheric response to increased anthropogenic emissions in a changing climate requires a comprehensive understanding of the biogeochemical cycle of  $H_2$ . Studying past changes in atmospheric  $H_2$  can inform predictions of the effects of future perturbations by providing insight into the relationship between atmospheric  $H_2$ , human activities, and climate.

The globally averaged mixing ratio of  $H_2$  in the modern atmosphere is roughly  $530 \text{ nmol mol}^{-1}$  (ppb). Major sources of atmospheric  $H_2$  include direct emissions from fossil fuel combustion and biomass burning and photochemical production from the oxidation of methane and non-methane hydrocarbons (NMHCs). Direct emission as a by-product of  $N_2$  fixation is a minor source of atmospheric  $H_2$ . The major sink of atmospheric  $H_2$  is consumption by soil microbes, accounting for about 70% of total losses and the remainder of the losses are to oxidation by the OH radical. Based on the atmospheric burden and estimates of the microbial and OH sinks, the lifetime of  $H_2$  is estimated at 2 years (Novelli et al., 1999; Pieterse et al., 2011; 2013; Paulot et al., 2021).

Increasing atmospheric  $H_2$  could influence the climate system in several ways. The reaction of  $H_2$  with OH represents a loss for OH and therefore increases the atmospheric methane lifetime and associated radiative forcing. This reaction leads to the production of  $HO_2$ , which can influence tropospheric ozone levels. In the stratosphere, the reaction of  $H_2$  with OH leads to the production of water vapor. Increased stratospheric water vapor cools the stratosphere and warms the troposphere. Paulot et al. (2021) estimated the effective radiative forcing of atmospheric  $H_2$  at  $0.13 \text{ mW m}^{-2} \text{ ppb}^{-1}$  due to its effects on the methane lifetime and stratospheric water vapor content.

Systematic measurements of atmospheric  $H_2$  levels began in the late 1980's (Khalil & Rasmussen, 1990). The NOAA Global Monitoring Laboratory (NOAA/GML), CSIRO GASLAB (CSIRO), and the Advanced Global Atmospheric Gases Experiment (AGAGE) started to monitor atmospheric  $H_2$  levels at several sites around the world in the early 1990's (Novelli et al., 1999; Novelli, 2006; Prinn et al., 2019; Langenfelds et al., 2002; Pétron et al., 2023). Integration of records produced by the different groups has been complicated by calibration issues. Earlier use of different scales was addressed by implementation of the internationally accepted MPI09 scale established by the Max Planck Institute for Biogeochemistry (MPI; Jordan & Steinberg, 2011). Measurements from CSIRO and AGAGE, have been transferred to the MPI09 scale. Additionally, NOAA/GML measurements from 2009 onward were recently revised to the MPI09 scale (Pétron et al., 2023). Broadly, the instrumental record shows northern hemispheric  $H_2$  levels rising during the late 1980's to a maximum near 1990. There is no discernible trend in Northern Hemisphere  $H_2$  levels during the 1990's and the first decade of the 2000's



(Figure 3). The Southern Hemisphere instrumental record shows H<sub>2</sub> levels rising until 1999, then plateauing for the next 10-15 years.

Longer-term changes in atmospheric H<sub>2</sub> levels during the 20<sup>th</sup> century have been reconstructed using polar firn air from Antarctica (South Pole, Megadunes) and Greenland (NEEM; Petrenko et al., 2013; Patterson et al., 2020; 2021). Firn is the compacted snow layer that forms the upper 40 – 120 m of ice sheets and its interconnected porosity contains an atmospheric archive that can extend up to ≈100 years back in time. For a comprehensive introduction to firn and firn air we refer the interested reader to Buizert (2013). The Antarctic reconstructions show that H<sub>2</sub> levels increased by 60% over the 20<sup>th</sup> century, consistent with the instrumental record and with increasing anthropogenic emissions and photochemical production from methane (Patterson et al., 2020). The Greenland reconstruction shows an increase in atmospheric H<sub>2</sub> levels during the 1960's to a peak in the late 1980's or early 1990's, followed by a small decline from 1990-2009 (Petrenko et al., 2013). This recent decline is inconsistent with northern hemisphere modern flask measurements that show roughly constant H<sub>2</sub> levels after the mid-1990's (Novelli, 2006; Prinn et al., 2019). Petrenko et al. (2013) noted that the firn air model used for that reconstruction did not include pore close-off fractionation, a process that enriches H<sub>2</sub> in the lock-in zone where vertical diffusion of trace gases effectively goes to zero. Patterson et al. (2020) suggested that the inferred peak could be an artifact of ignoring this process.

In this work, we reassess historical Northern Hemisphere atmospheric H<sub>2</sub> levels over the 20<sup>th</sup> century using firn air from NEEM and two additional Greenland sites with previously unpublished H<sub>2</sub> measurements (Summit and Tunu). Reliable bipolar records are useful for constraining changes to H<sub>2</sub> cycling because the two hemispheres differ in their sensitivity to changes in the soil sink and anthropogenic emissions. Understanding past variability in the soil sink is particularly important because the response of the soil sink to future changes in climate is the largest source of uncertainty in projecting the radiative consequences of increasing anthropogenic H<sub>2</sub> emissions (Warwick et al., 2022).

## 2. H<sub>2</sub> measurements in Greenland firn air

### 2.1 Firn air sampling

The sampling methods used in the Summit, Tunu, and NEEM campaigns were similar to techniques from previous published firn air studies (e.g. Battle et al., 1996; Severinghaus et al., 2001; Severinghaus & Battle, 2006). Briefly, a borehole was drilled into the firn, pausing the drilling at each desired sampling depth. The borehole was sealed above each sampling depth with an inflatable rubber packer to prevent contamination from the modern atmosphere. Synflex 1300 (formerly known as “Decabon”) tubing for both waste air and sample air extends from below the packer to the surface. The waste air intake was positioned directly below the rubber packer. The waste air intake was separated from the sample air intake by a stainless-steel plate (baffle) with a diameter slightly less than the borehole. Air was pumped from the waste air intake 2-5x faster than from the sample air intake to ensure that no air that had been in contact with the rubber packer was



95 sampled. Sampled air was dried using  $\text{Mg}(\text{ClO}_4)_2$  and stored in 2.5 L glass flasks. More detailed information regarding the sampling at each site may be found in the references listed in Table 1.

100

**Table 1-** Site characteristics and data references for the three Greenland firn air sampling campaigns.

Site	Lat/lon	Date	T (°C) <sup>1,4</sup>	Accum. <sup>2,4</sup>	Ref.	Firn air model parameters <sup>3</sup>
NEEM	77°N, 51°W	7/2008	-29	21.6	Petrenko et al., 2013	Buizert et al., 2012
Summit	73°N, 38°W	5/2013	-31	23.4	Hmiel et al., 2020	C. Buizert, personal communication
Tunu	78° N, 34° W	5/1996	-29	10.0	Butler et al., 1999	This study

<sup>1</sup>Mean annual surface temperature

<sup>2</sup>Average modern accumulation (cm ice equivalent per year)

105 <sup>3</sup>Modeling parameters include density profile, diffusivity profile, partitioning between open and closed porosity, and the lock-in depth (Section 3.1).

<sup>4</sup>Temperature and accumulation for NEEM and Tunu are from Buizert et al. (2012) and Butler et al. (1999) respectively. Temperature and accumulation for Summit were supplied by Christo Buizert (Personal Communication).

## 110 2.2 Firn air measurements and calibration

Firn air samples from Tunu were analyzed for  $\text{H}_2$  by the NOAA/GML Carbon Cycle Group using a reducing gas analyzer (RGA) with  $\text{HgO}$  detection. At the time of analysis, the measurements were calibrated using a single 530 ppb  $\text{H}_2$  standard assuming a linear detector response. We subsequently applied an empirical correction to account for the non-linear response of the  $\text{HgO}$ -RGA. The non-linearity bias is estimated to be -3 ppb at 453 ppb and -12 ppb at 312 ppb based on  
115 laboratory measurements made at NOAA/GML in 2009. The biases were linearly interpolated onto the firn air measurements (i.e. a measurement of 300 ppb was corrected to 312 ppb). Jordan & Steinberg (2011) found a substantially larger non-linearity bias for a similar RGA detector. The sensitivity of our results to uncertainty in the non-linearity correction is minimal as demonstrated in the Supplement (Figures S3 and S4). The Tunu data were originally calibrated on the NOAA96 calibration scale, that is known to have drifted due to  $\text{H}_2$  increasing in the GML gravimetric standards stored in aluminum  
120 cylinders. We adjusted the Tunu firn air measurements to the MPI09 scale based on a matched flask intercomparison between NOAA/GML and CSIRO conducted using atmospheric samples from Cape Grim Observatory, Tasmania during June and July 1996. The intercomparison gives an average offset (CSIRO-NOAA/GML) of 4.7 ppb. A constant correction of 4.7 ppb was added to the Tunu measurements to account for the offset. The Summit firn air measurements were made by NOAA/GML using a He pulsed discharge detector with linear response (HePDD; Novelli et al., 2009). The Summit



125 measurements have been formally revised to the MPI09 calibration scale. For both the Tunu and Summit data, the  
corrections described above are below 5% and do not significantly influence the main conclusions of this research.

Firm air samples from NEEM were analyzed for H<sub>2</sub> at CSIRO using gas chromatography coupled to an HGO-RGA.  
These measurements were calibrated using standards between 340 and 1000 ppb. The CSIRO measurements have been  
revised to the MPI09 scale. CSIRO estimates the analytical uncertainty to be  $\pm 0.2\%$ . However, there are other sources of  
130 uncertainty in the true H<sub>2</sub> content in the firm air. For example, tests conducted during the NEEM campaign demonstrated a  
small procedural blank (4-6 ppb). No such tests were conducted at the other sites. To account for the possibility of such  
blanks, we assign an overall uncertainty of  $\pm 2\%$  to the data from all sites.

The firm air H<sub>2</sub> measurements from all three sites were corrected for gravitational fractionation (equation 2; Section  
3.1), and duplicate measurements at the same depth averaged to generate the finalized depth profiles (Figure 1b-d; 2b-d).  
135 Lock-in depths for the sites are given in Table 2. The H<sub>2</sub> depth profiles at the three Greenland sites display strong gradients  
above lock-in that reflect the fast gas phase diffusivity of H<sub>2</sub> and intense seasonality of atmospheric H<sub>2</sub> in the northern  
hemisphere. The H<sub>2</sub> gradients observed in the upper firm at NEEM are different than the gradients at Summit and Tunu  
because NEEM was sampled in July while the other two sites were sampled in May. The depth profiles from NEEM and  
Summit both display sharp maxima of 530 and 560 ppb respectively in the lock-in zone. In contrast, Tunu H<sub>2</sub> levels do not  
140 display a peak in the lock-in zone. Instead, the Tunu H<sub>2</sub> measurements are approximately constant for the 10 m just above  
lock-in and then decrease significantly to a minimum of 420 ppb at the bottom of lock-in. The NEEM measurements also  
show a sharp decrease to 490 ppb at the bottom of lock-in. At Summit, the decrease towards the bottom of lock-in is less  
dramatic than at the other sites because only one sample depth is deeper than the observed maximum. The qualitative  
differences in the depth profiles reflect both the different sampling times ranging from 1996-2013 and the different site  
145 physical characteristics. For example, the firm air at the bottom of the lock-in zone at Summit is younger than the two other  
sites because: 1) Summit was sampled later than the other sites and 2) Summit has a much higher accumulation rate than  
Tunu and a thinner lock-in zone than NEEM (Table 1).

**Table 2-** Summary of firm air H<sub>2</sub> measurements from NEEM, Summit, and Tunu.

Site	Samples	Unique Depths	Lab	Detector	Lock-in depth (m)
NEEM	23	18	CSIRO	HgO-RGA	63
Summit	37	19	NOAA	HePDD	68.5
Tunu	48	22	NOAA	HgO-RGA	58



### 150 3. Firn air modelling and inversions

#### 3.1 Firn air model

The UCI\_2 firn air model is a 1-dimensional finite-difference advective-diffusive model that is used to simulate the evolution of H<sub>2</sub> levels in firn air. The model domain is divided into an upper “diffusive zone” and lower “lock-in zone.” In the diffusive zone, vertical gas transport occurs via wind-driven convective mixing in the shallowest ~5 m and via molecular diffusion throughout. Diffusive mixing decreases with depth due to the increasing tortuosity of the firn. In the lock-in zone, vertical molecular diffusion ceases due to the presence of impermeable ice layers. Gas transport in the lock-in zone occurs primarily due to advection with a small non-fractionating mixing term. The model uses a forward Euler integration scheme and a time step of 324 s. The model is largely based on Severinghaus et al. (2010), with two important differences: 1) thermal diffusion is neglected, as it is unimportant for H<sub>2</sub>, and 2) our model parameterizes pore close-off fractionation differently than the Severinghaus model (see below). The model tracks the air content and composition in both open pores and closed bubbles as a function of time and depth. The model code is written and executed in MATLAB R2022a (Mathworks Inc.).

The site-specific bulk density profile ( $\rho_{firn}$ ; kg m<sup>-3</sup>) is calculated from an empirical fit to density measurements of the firn core. Total porosity ( $s_{total}$ ; dimensionless) is estimated from the density profile:

$$165 \quad s_{total} = 1 - \frac{\rho_{firn}}{\rho_{ice}} \quad (1)$$

Where  $\rho_{ice}$  (kg m<sup>-3</sup>) is the temperature dependent density of ice from Bader (1964). Accumulation rate, temperature, and the depth of the lock-in zone onset are site-specific parameters, and we implement previously published parameterizations for the partitioning between open and closed porosity at each site (Tables 1 and 2).

Model grid spacing is 0.5 m in the diffusive zone. Gas transport is dominated by convective mixing in the upper part of the diffusive zone, and by molecular diffusion in the lower part of the diffusive zone. To simulate non-fractionating convective mixing, a depth-dependent eddy diffusivity term is added to the classical one-dimensional firn air transport equation (Schwander et al., 1993; Severinghaus et al., 2010; Trudinger et al., 1997). Additionally, we neglect the typical gravitational term, and instead empirically correct the firn air measurements using the  $\delta^{15}N$  of N<sub>2</sub> depth profile. The correction is calculated from the  $\delta^{15}N$  data for each borehole according to equation 2:

$$175 \quad Corr_{grav} = \frac{\delta^{15}N}{1000} * \frac{\exp\left(g*z*\frac{\Delta m_g}{RT}\right)-1}{\exp\left(g*z*\frac{\Delta m_{15N}}{RT}\right)-1} \quad (2)$$

180 Where  $Corr_{grav}$  is the depth dependent fractional correction for the gas of interest,  $\delta^{15}N$  is the measured isotopic composition of N<sub>2</sub> (‰) each depth,  $g$  is the gravitational acceleration constant (9.8 m s<sup>-2</sup>),  $z$  is depth (m),  $R$  is the ideal gas constant (8.314



$\text{J mol}^{-1} \text{K}^{-1}$ ),  $T$  is the annual average temperature at the site,  $\Delta m_g$  is the difference in molar mass (in kg) between the gas of interest and air, and  $\Delta m_{15N}$  is the difference in molar mass between  $^{28}\text{N}_2$  and  $^{29}\text{N}_2$  (.001 kg).

Equation 3 governs the evolution of the concentration of the gas of interest in the open pores in the diffusive zone  
185 (Severinghaus et al., 2010):

$$s_o \frac{\partial C}{\partial t} = \frac{\partial}{\partial z} \left( \frac{\partial C}{\partial z} * [s_o D_{mol}(z, T, P) + s_o D_{eddy}(z)] \right) - s_o w \frac{\partial C}{\partial z} \quad (3)$$

Where  $s_o$  is the open porosity,  $C$  is concentration of the gas of interest ( $\text{mol m}^{-3}$ ),  $D_{mol}$  is the gas, depth, temperature,  
190 and pressure dependent molecular diffusivity constant ( $\text{m}^2 \text{s}^{-1}$ ),  $D_{eddy}$  is the depth dependent eddy diffusivity ( $\text{m}^2 \text{s}^{-1}$ ), and  $w$  is  
the downward velocity of the firn column ( $\text{m s}^{-1}$ ). We used  $D_{mol}$  and  $D_{eddy}$  profiles for NEEM and Summit developed  
previously using a suite of measured gases with well constrained atmospheric histories of  $\text{CO}_2$ ,  $\text{CH}_4$ ,  $\text{CH}_3\text{CCl}_3$ , and  $\text{SF}_6$   
(Buizert et al., 2012; Hmiel et al., 2020; C. Buizert, personal communication). The diffusivity profiles were validated by  
forcing the model with the established atmospheric histories for the previously mentioned trace gases and comparing the  
195 model output to the measurements. Diffusivity profiles for Tunu were tuned using  $\text{CO}_2$  and  $\text{CH}_4$  measurements. During the  
tuning process, the ratio between the effective molecular diffusivities of each trace gas at each depth is forced to remain  
identical to the ratio of their gas phase diffusivities (i.e. at each depth  $D_{mol}$  for  $\text{CH}_4$  is 1.4 time that of  $D_{mol}$  for  $\text{CO}_2$ ). After  
tuning  $D_{mol}$  is then scaled using the ratio of the gas phase diffusivity of  $\text{CO}_2$  to the gas phase diffusivity of the gas of  $\text{H}_2$  or  
other trace gas of interest. Temperature and pressure dependent gas phase diffusivities are calculated using Fuller et al.  
200 (1966) as described by Reid et al. (1987; Ch. 11).

Model grid spacing in the lock-in zone is 1 annual layer. There is no molecular diffusivity in the lock-in zone and  
gas transport occurs via downward advection of annual layers once per year. Small values of eddy diffusivity are prescribed  
in the upper part of lock-in. This non-fractionating mixing term is included to account for vertical airflow caused by  
barometric pressure fluctuations. Model parameterizations for vertical mixing in the lock-in zone are validated by forcing the  
205 model with best-estimate Greenland atmospheric histories for various trace gases with well constrained atmospheric histories  
(Buizert et al., 2012).

In addition to convective mixing, diffusion and advection, pore close-off fractionation is parameterized in the firn  
air model. As closed pores (bubbles) are advected downward in the firn, internal pressure increases, creating a partial  
pressure gradient which drives a net diffusive flux of highly mobile trace gases out of the bubbles, through the ice lattice,  
210 and into the open pores. The open pores are therefore enriched in these gases. This phenomenon affects gases with kinetic  
diameters (KD)  $< 3.6 \text{ \AA}$  such as Ne, and He due to their large diffusivity in ice (Severinghaus & Battle, 2006).

Given its small molecular diameter (KD= 2.89  $\text{\AA}$ ) pore close-off fractionation must affect  $\text{H}_2$  in a similar manner to  
Ne (Patterson et al., 2020; Petrenko et al., 2013, Severinghaus & Battle, 2006). Recent laboratory measurements indicate  
that the permeability of  $\text{H}_2$  in ice is sufficiently fast to equalize the partial pressure of  $\text{H}_2$  in the open porosity and closed



215 bubbles (Patterson & Saltzman, 2021). Therefore, in the model, it is assumed that the partial pressure of H<sub>2</sub> is in equilibrium between the closed bubbles and open pores in each layer as described by equations 4-6:

$$P_n = (P_{bubble}x_{n(bubble)}s_c + P_{ambient}x_{n(firn)}s_o)/s_{total} \quad (4)$$

220  $x_{n(bubble)} = P_n/P_{bubble} \quad (5)$

$$x_{n(firn)} = P_n/P_{ambient} \quad (6)$$

Where  $P_n$  (Pa) is partial pressure of gas  $n$  (H<sub>2</sub> or Ne; See section 5),  $x_n$  is mole fraction of gas  $n$ ,  $P_{bubble}$  (Pa) is the total bubble  
225 pressure,  $P_{ambient}$  (Pa) is the ambient pressure in the open pores, and  $s_c$  is closed porosity. The subscripts *bubble* and *firn* distinguish between the closed and open porosity. Equations 4-6 are executed at every time step in each grid cell in the model. Note that the equilibrium parameterization of pore close-off fractionation used here is different from the kinetic parameterization used by Severinghaus & Battle (2006) which was developed assuming a much smaller permeation constant for Ne. Pore close-off fractionation is important in terms of the exchange of H<sub>2</sub> and Ne between closed and open pores.  
230 Permeation of these gases through the ice lattice vertically between model layers can be neglected on time scales relevant to firn air modelling (Patterson & Saltzman 2021).

### 3.2 Inversion methods

The firn air measurements are inverted to recover an atmospheric history of H<sub>2</sub> from each site. The inversion  
235 technique employs depth-dependent age distributions or “Green’s functions” ( $G$ ; (Rommelaere et al., 1997)). The firn air model is initialized with no H<sub>2</sub> in the firn air column, then forced with a transient 1-year pulse of H<sub>2</sub> at the surface. The model is integrated for 300 years, and the evolution of the pulse is tracked as a function of depth and time to produce the Green’s functions. The age distributions grow older and broader with depth (Figure S1). Given an atmospheric history ( $H_{2(atm)}$ ), modelled levels of H<sub>2</sub> in the firn ( $H_{2(firn)}$ ) are calculated according to equation 7:

240 
$$H_{2(firn)} = \sum_{t=0}^{300} H_{2(atm)} * G \quad (7)$$

Using Green’s function allows for rapid iteration over many possible atmospheric histories without running the firn air  
model in forward mode. The inversion problem is under-constrained. Multiple atmospheric histories can provide good  
245 agreement with the measured depth profile, and additional constraints are required to yield a unique solution. In previous studies, the additional constraint has been implemented as a smoothing parameter (Patterson et al., 2020, 2021; Petrenko et



al., 2013; Rommelaere et al., 1997). In this work, we rely instead on the assumption of autocorrelation and utilize the probabilistic modelling software package, Stan (mcstan.org). Stan has been used previously for firm air reconstructions by Aydin et al. (2020). As in that work, we use Stan to implement a Bayesian hierarchical model in which the atmospheric history is treated as an auto-correlated random variable:

$$\mathbf{m}_{atm}[t_i] \sim N(\mathbf{m}_{atm}[t_{i-1}], \beta \mathbf{m}_{atm}[t_{i-1}]) \quad (8)$$

Where  $\mathbf{m}_{atm}$  is a vector of length  $i$  that contains the discrete atmospheric  $H_2$  dry air mole fraction history. The atmospheric mixing ratio at time  $t_i$  is normally distributed around the atmospheric mixing ratio at time  $t_{i-1}$  with a standard deviation of  $\beta \mathbf{m}_{atm}[t_{i-1}]$ .  $\beta$  is a positive scalar which may be specified or varied as a free parameter. The atmospheric history is related to the firm air measurements:

$$\mathbf{h}_{obs} \sim N(\mathbf{G}\mathbf{m}_{atm}, \boldsymbol{\sigma}) \quad (9)$$

Where  $\mathbf{h}_{obs}$  is a vector of length  $n$ , which contains the measured  $H_2$  levels at each depth,  $\mathbf{G}$  is an  $n$  by  $i$  matrix where each row corresponds to the Green's function for that unique depth (referred to as the "kernels" by Rommelaere et al. (1997)), and  $\boldsymbol{\sigma}$  is a vector of length  $n$ , which contains the analytical uncertainties of the firm air measurements. Stan samples the joint posterior probability distribution for the parameters,  $\mathbf{m}_{atm}$  and  $\beta$  using a fast Hamiltonian Markov Chain Monte Carlo algorithm (Carpenter et al., 2017). Parameters are sampled from uniform prior distributions unless otherwise specified. Inversions are carried out using the MatlabStan 2.15.1.0 interface to cmdstan 2.27.0. The advantages of this method over other methods include: 1) No artificial smoothing criteria is imposed (instead the atmospheric history is assumed to be autocorrelated), and 2) Stan allows us to easily explore sensitivity to assumptions about the data by encoding those assumptions as priors.

In Greenland,  $H_2$  levels in the upper firm layers are strongly influenced by seasonal variations and firm air reconstructions cannot accurately reconstruct such high frequency variability. Therefore, measurements from the upper part of the firm were excluded from the reconstructions. A cut-off depth was determined by forcing the firm air model in forward mode with an atmospheric history that consisted of constant annually averaged  $H_2$  levels with a realistic seasonal cycle imposed. Depths where the  $H_2$  concentration varied by more than 1% from the annual average were excluded from the reconstruction. The cut-off depths were 62 m, 66 m, and 50 m for NEEM, Summit and Tunu respectively. Unlike Greenland, Antarctic  $H_2$  reconstructions are not sensitive to seasonality primarily because seasonal variability in atmospheric  $H_2$  levels over Antarctica is less than half that of seasonal variability over Greenland.



#### 4. Atmospheric reconstructions

280 Atmospheric H<sub>2</sub> reconstructions were carried out for each site independently (Figure 1). The NEEM, Summit, and  
Tunu sites have different physical characteristics (mean annual temperature and accumulation rate) and were sampled at  
different times. The firm air age distributions at the three sites are shown in Figure S1. The model inversions yield generally  
good agreement between the modeled and measured depth profiles for NEEM and Tunu (Figure 1b-d). Agreement is also  
good at Summit, except at the bottom of lock-in where the modeled depth profile increases monotonically with depth while  
285 the deepest measurement shows a slight decrease.

Firm air measurements from NEEM constrain atmospheric H<sub>2</sub> levels after 1950. The NEEM reconstruction shows  
atmospheric H<sub>2</sub> increasing from 1950-1989 at an average rate of 3.1 ppb y<sup>-1</sup> (425-545 ppb). After 1989, atmospheric levels  
decrease at an average rate of 1.6 ppb y<sup>-1</sup> reaching 515 ppb in 2008. Summit firm air is younger, yielding an atmospheric  
history beginning in 1975. The Summit reconstruction shows atmospheric H<sub>2</sub> increasing from 505-535 from 1975 to 1994,  
290 an average rate of 1.5 ppb y<sup>-1</sup>. After 1994 there is a slight decrease in atmospheric H<sub>2</sub> levels until 2003. After 2003,  
atmospheric H<sub>2</sub> decreases more rapidly at a rate of 1.5 ppb y<sup>-1</sup>, reaching 515 ppb in 2013. Tunu firm air is considerably older,  
showing a rise in atmospheric H<sub>2</sub> from 370-510 ppb from 1930 to 1992, an average rate of 2.3 ppb y<sup>-1</sup>. After 1992, H<sub>2</sub> levels  
decreased at a rate of 1.5 ppb y<sup>-1</sup> to just over 500 ppb in 1996.

All three reconstructions show H<sub>2</sub> levels rising during the mid-late 20<sup>th</sup> century with a maximum in atmospheric H<sub>2</sub>  
295 near 1990. The maximum is somewhat later and broader in the Summit reconstruction compared to the other two sites. There  
is only one measurement below the observed H<sub>2</sub> maximum at Summit, and that single measurement does not provide a  
sufficiently strong constraint on the reconstruction to generate the sharper maximum observed in the NEEM and Tunu  
reconstructions. We conducted another reconstruction using the Summit data with the uncertainty on the measurements  
reduced by 25%. Reducing the uncertainty on the Summit measurements imposes tighter constraints on the reconstruction  
and forces better agreement between the modeled and measured depth profiles. This reconstruction shows an earlier  
300 maximum (1992) and a significantly more rapid rise in H<sub>2</sub> levels from 1975-1992, in better agreement with the  
reconstructions from NEEM and Tunu (Figure S2). The reconstructed decrease during the mid to late-1990's is still  
somewhat slower than in other two reconstructions.

Although the trends observed in the NEEM and Tunu reconstructions are in good agreement, there is an offset of  
305 roughly 35 ppb in absolute reconstructed H<sub>2</sub> levels (Figure 1a). We suspect that the offset is due to a calibration problem in  
the Tunu data. Those measurements were made in 1996, the same year that NOAA revised their H<sub>2</sub> working standard  
upward by 6.4%. Revising the Tunu measurements upward by 6.4% yields a reconstruction in near perfect agreement with  
the NEEM reconstruction. Furthermore, the probability model suggests that adjusting the Tunu measurements upwards by  
7.1% yields the best agreement between the three sites (see below), in good agreement with the 6.4% calibration adjustment.  
310 The original NOAA laboratory notes for the Tunu samples indicate that the upward revision to the working standard was



applied to the Tunu measurements but the data suggest that this revision was not actually applied. An alternative explanation for the offset between sites could be *in situ* H<sub>2</sub> production (see Section 8).

315 A “joint” reconstruction was carried out in which the most probable atmospheric history given the firm air data from all three sites is calculated (Figure 2). To account for the apparent offset between Tunu and the other sites, we introduced another free parameter,  $\gamma$ , which is a dimensionless scalar multiplier for the data from Tunu.  $\gamma$  is drawn from a uniform prior between 0.9 and 1.1. Equation 9 is adapted for the measurements from Tunu.

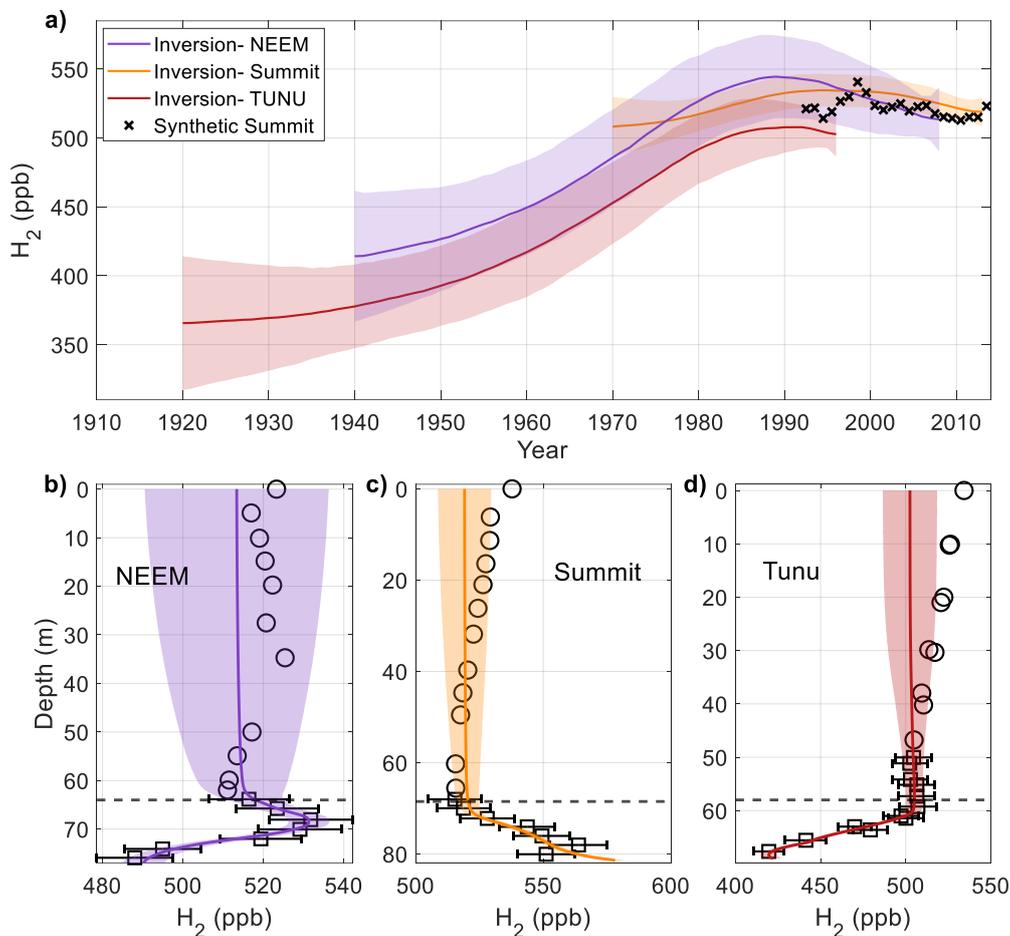
$$\gamma \mathbf{h}_{obs} \sim N(\mathbf{G}\mathbf{m}_{atm}, \sigma) \quad (10)$$

320 The posterior distribution for  $\gamma$  is approximately normally distributed around 1.071 with a standard deviation of 0.012, implying an offset of about 7.1% for the Tunu measurements. Atmospheric H<sub>2</sub> levels increase from 390 ppb in 1930 to 545 ppb in 1990 at an average rate of 2.6 ppb y<sup>-1</sup>. After 1990, atmospheric H<sub>2</sub> starts to decrease at an average rate of 1.4 ppb y<sup>-1</sup>, reaching 515 ppb in 2013. There is good agreement between the modeled and measured depth profiles for the joint reconstruction.

325

**Figure 1: Atmospheric histories reconstructed independently from firn air profiles at three Greenland sites. a) purple line and shading- result from NEEM and associated  $\pm 1\sigma$  uncertainty; orange line and shading- result from Summit and associated  $\pm 1\sigma$  uncertainty; dark red line and shading- result from Tunu and associated  $\pm 1\sigma$  uncertainty; black x's- annual mean synthetic Summit H<sub>2</sub> history (Section 5; Pétron et al., 2023; Langenfelds et al., 2002); b) black markers- measured H<sub>2</sub> depth profile at NEEM; squares with error bars are measurements used in the reconstruction, and circles are measurements excluded from the reconstruction because of seasonality; purple line and shading- modeled depth profile using the atmospheric history plotted in purple in a) with the propagated  $\pm 1\sigma$  uncertainty; the dashed black line indicates the top of lock-in zone c) and d) as in b) for Summit and Tunu respectively**

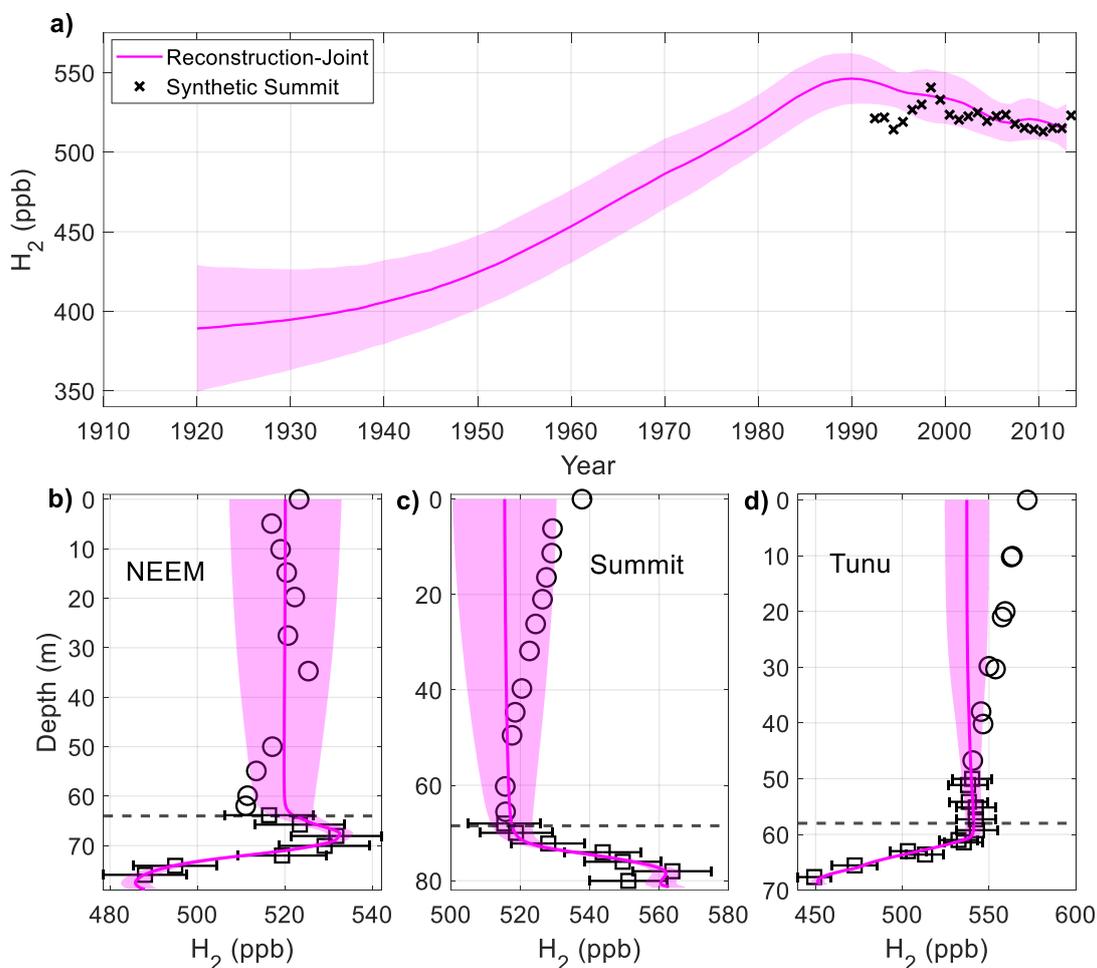
330



335

**Figure 2: Joint atmospheric H<sub>2</sub> reconstruction using firn air from three Greenland sites (NEEM, Summit, and Tunu). The reconstruction was conducted using equation 9 for NEEM and Summit, and equation 10 for Tunu (see text). All lines and markers are the same as in Figure 1. The Tunu measurements are adjusted upwards by 7.1% (see text).**

340



### 5. Comparison of firn air reconstruction with atmospheric flask measurements

The firn air reconstructions are compared to available atmospheric  $H_2$  flask measurements from the high northern latitudes (Figure 3). Direct comparisons between the firn reconstructions and atmospheric measurements at other sites are complicated by spatial variability in  $H_2$  levels in the high northern latitudes. The best temporal coverage for flask measurements is from the Barrow, Alaska and Alert, Nunavut sampling sites. Both sites sample the boundary layer and are influenced by their proximity to seasonally active soils.  $H_2$  levels over Greenland are more representative of the free troposphere due to higher elevation and remoteness from active soils. To make this comparison, we utilize NOAA/GML measurements at Alert, Nunavut and Summit, Greenland from 2010-present (revised to the MPI09 calibration scale; Pétron et al., 2023) and CSIRO measurements from Alert from 1992 onwards (also on the MPI09 scale).

During the 12 years of NOAA/GML measurements at Summit and Alert, the offset between Summit and Alert varies from a minimum of 18.9 ppb to a maximum of 29.4 ppb with a mean offset of  $25.4 \text{ ppb} \pm 3.0 \text{ ppb}$  ( $1\sigma$ ) with Summit



355 levels higher than Alert. We added this offset to the longer CSIRO Alert time series to generate a longer “synthetic Summit”  
flask history, assuming that the offset remains constant through time (Figures 1-4, 8). The most conspicuous feature of the  
synthetic Summit history is a sharp peak in H<sub>2</sub> levels during 1998 which can be attributed to increased biomass burning  
emissions during the strong 1997-1998 ENSO event. Firm air cannot resolve such high frequency variability due to diffusive  
smoothing, so that feature is absent from our reconstructions. Excepting the 1998 peak, synthetic Summit H<sub>2</sub> levels are  
largely constant between 525 and 515 from 1992-2013. In contrast, the joint reconstruction shows a trend with H<sub>2</sub> levels  
360 decreasing from 545 to 515 ppb over the same time period.

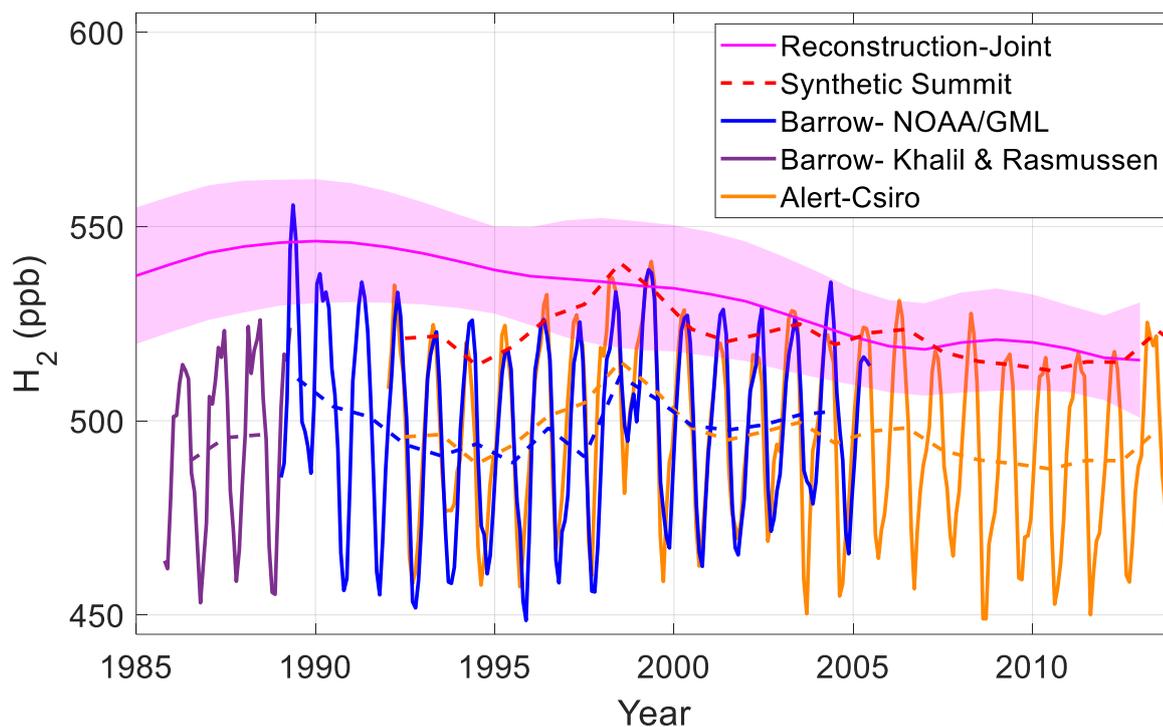
We also examined the atmospheric data for evidence of the H<sub>2</sub> maximum around 1990 indicated by the firm air  
reconstruction. If this peak had occurred in the atmosphere, we would expect similar trends at all of the high latitude flask  
sites despite the differences in absolute H<sub>2</sub> levels. Here we examine flask measurements from Barrow, Alaska (Khalil &  
Rasmussen, 1990; NOAA/GML) and Alert, Canada (CSIRO; Figure 3). The measurements from NOAA/GML from 1992-  
365 2005 were corrected to the MPI09 calibration scale using the matched flask intercomparison project mentioned previously.  
The Khalil & Rasmussen (1990) data are reported on an independent calibration scale which has not been intercompared to  
the MPI09 calibration scale. The Barrow data from Khalil & Rasmussen (1990) shows an increase in atmospheric H<sub>2</sub> of 2.1  
ppb y<sup>-1</sup> from 1985-1989 and the subsequent NOAA/GML data show a decrease in atmospheric H<sub>2</sub> from 1989-1993, at a rate  
of 5.0 ppb y<sup>-1</sup>. Together, the trends in these two data sets imply a maximum in atmospheric H<sub>2</sub> in 1989, in agreement with the  
370 firm air reconstructions. The strength of this conclusion is questionable, because the flask intercomparison project did not  
include data prior to 1992. From 1993-2010, there is no discernible trend in atmospheric H<sub>2</sub> at either Barrow or Alert, while  
the firm air reconstructions show a slow decrease in atmospheric H<sub>2</sub>. The flask measurements suggest a sharper maximum  
and more rapid stabilization than the firm air reconstructions.

Firm diffusion imposes smoothing on atmospheric reconstructions. To test whether the breadth of the atmospheric  
375 H<sub>2</sub> peak and slow subsequent decrease in the in the firm air reconstruction could be an artifact of that smoothing, we forced  
the firm air model with a synthetic atmospheric history. The synthetic history was identical to the joint reconstruction before  
1989, decreasing by 5 ppb y<sup>-1</sup> decrease from 1989- 1993 (as observed in the Barrow flask measurements) and subsequently  
decreases by 0.5 ppb y<sup>-1</sup>, reaching the same atmospheric concentration as the joint reconstruction in 2013. The depth profiles  
generated by the model when forced with the synthetic history show poorer agreement with the measured firm H<sub>2</sub> than the  
380 levels generated by the joint reconstruction (Figure S5). This result demonstrates that the firm air measurements are more  
consistent with a broad peak and slow decrease in H<sub>2</sub> levels than a sharp peak, rapid decrease, and stabilization.

**Figure 3: Comparison of the joint Greenland firm air reconstruction with high northern latitude flask measurements.**  
385 **Magenta line and shading- joint reconstruction as in Figure 2a; purple lines- monthly (solid) and annually-averaged  
(dashed) flask measurements from Barrow, Alaska made by Khalil & Rasmussen (1990); Blue lines- monthly and  
annually-averaged flask measurement from Barrow, Alaska made by NOAA/GML (Novelli et al., 1999; Novelli,  
2006); Red line- annual mean synthetic Summit H<sub>2</sub> history (Section 5; Pétron et al., 2023, Langenfelds et al., 2002);**



390 Yellow lines- monthly and annually-averaged flask measurement from Alert, Nunavut made by CSIRO (Langenfelds et al., 2002). Measurements from NOAA/GML were empirically corrected to the MPI09 calibration scale using a matched flask inter-comparison project between CSIRO and NOAA/GML.





## 6. Comparison to previous firn air reconstructions

395 Petrenko et al. (2013) produced two atmospheric H<sub>2</sub> reconstructions from the NEEM firn air data. The  
reconstructions were generated using two different models: LGGE-GISPA and INSTAAR. Neither of those models include  
pore close-off fractionation, and Patterson et al. (2021) suggested that maximum in atmospheric H<sub>2</sub> inferred from the NEEM  
firn air could be an artifact of ignoring pore close-off fractionation. Petrenko et al (2013) used the matrix inversion method  
of Rommelaere et al. (1997), which requires a prescribed smoothing parameter to fully constrain the inversion. Those results  
400 were corrected for a calibration revision (see Supplement) and are compared to our joint reconstruction (Figure 4a). The  
LGGE-GISPA reconstruction shows a rapid rise in H<sub>2</sub> levels from 1960-1990, and a rapid decrease from 1990-2008. The  
INSTAAR reconstruction shows H<sub>2</sub> levels rising more gradually from 1960-1985 and decreasing gradually thereafter. The  
rate of increase of H<sub>2</sub> levels prior to 1990 in our joint reconstruction falls between the rates of increase from the LGGE-  
GISPA and INSTAAR reconstructions. The timing of the atmospheric peak in our joint reconstruction shows good  
405 agreement with the timing from the LGGE-GISPA reconstruction, but the magnitude of the peak is in better agreement with  
the INSTAAR reconstruction. The rate of the decline in atmospheric H<sub>2</sub> from 1990-2008 from our joint reconstruction is  
nearly identical to the rate from the INSTAAR reconstruction. The late 20<sup>th</sup> century maximum in atmospheric H<sub>2</sub> is a robust  
feature of all Greenland firn air reconstructions. According to our model, pore close-off induced enrichment is <1% at the  
depth of the observed H<sub>2</sub> maximum in the lock-in zone. That is, the observed lock-in maximum is not caused by enrichment,  
410 but by a historical atmospheric maximum.

We conducted a similar joint reconstruction for H<sub>2</sub> firn measurements from two Antarctic sites: South Pole and  
Megadunes. The results of that joint reconstruction are compared to previously published reconstructions from those  
individual sites from (Patterson et al., 2020; 2021; Figure 4b). The differences between the joint reconstructions and the  
previously published reconstructions are relatively minor. The joint reconstruction shows a slightly slower increase in H<sub>2</sub>  
415 levels prior to 1965 and a more rapid increase from 1965-1980. After 1980, the joint reconstruction is in good agreement  
with the previously published reconstructions. Overall, the earlier Antarctic reconstructions exhibit a more uniform rate of  
increase over the past century. This is likely due to the use of an arbitrary smoothing parameter in the earlier studies, while  
an autocorrelation function was used in the joint reconstruction.

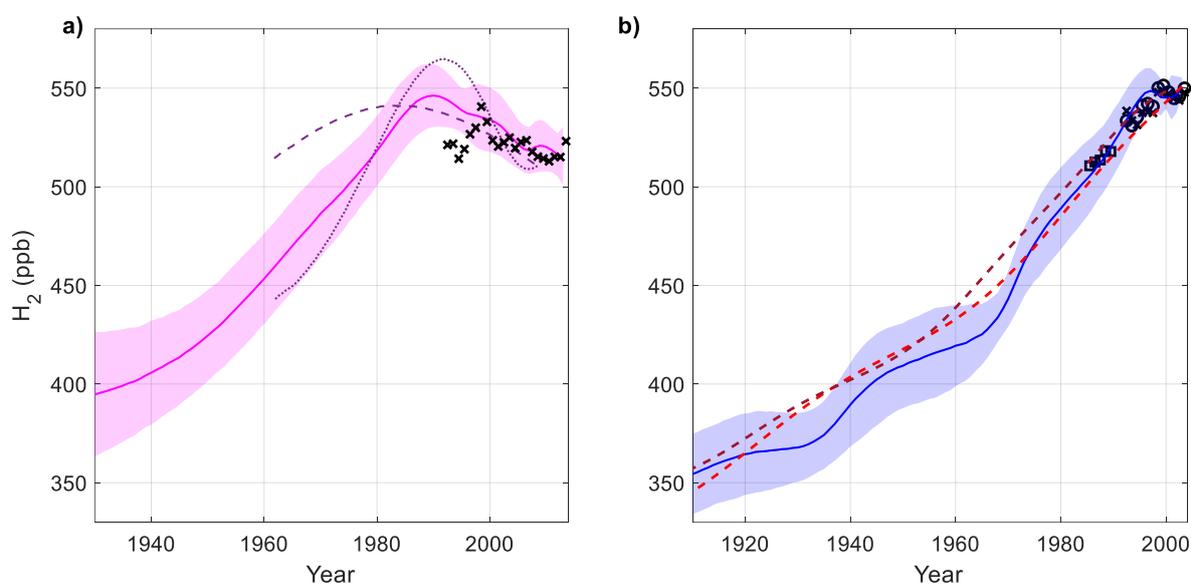
420

425

430



435 **Figure 4: Comparison of our joint reconstructions with previously published firn air reconstructions a) Magenta line and shading- joint Greenland reconstruction as in Figure 2a; purple lines- Reconstructions from Petrenko et al. (2013) using the LGGE-GISPA model (dotted line) and INSTAAR model (dashed line); black x's- annual mean synthetic Summit H<sub>2</sub> history (Section 5; Pétron et al., 2023, Langenfelds et al., 2002). b) Blue line and shading- joint Antarctic reconstruction and associated  $\pm 1\sigma$  uncertainty; dashed maroon line- South Pole firn air reconstruction from Patterson et al., 2020; dashed red line- Megadunes firn air reconstruction from Patterson et al., 2021; black x's- H<sub>2</sub> annual means from measurements made by NOAA/GML at Palmer, Syowa, South Pole, and Halley Stations, Antarctica and Cape Grim Observatory, Tasmania (Novelli et al., 1999; Novelli, 2006); black circles- H<sub>2</sub> annual means from measurements made by CSIRO at Casey, South Pole, and Mawson Stations, Antarctica and Cape Grim and Macquarie Island, Tasmania (Langenfelds et al., 2002); H<sub>2</sub> annual means from measurements made by Khalil & Rasmussen (1990) at Palmer Station, Antarctica, and Cape Grim, Tasmania**



## 7. Investigating the impact of bias in the pore close-off fractionation parameterization

445 Ne and H<sub>2</sub> have similar permeability in ice due to their similar molecular diameters (Patterson & Saltzman, 2021; Satoh et al., 1996). Atmospheric Ne is constant, making it a useful tracer for assessing the effects of pore close-off fractionation. The UCI<sub>2</sub> firn air model underestimates pore close-off induced Ne enrichment at the base of the firn at NEEM and Summit. If a firn air model underestimates Ne enrichment, it is highly likely to also underestimate pore close-off enrichment for H<sub>2</sub>. An underestimate of pore close-off induced enrichment would yield a reconstruction that is biased high, and a high bias could be responsible for the reconstructed 1990's peak in H<sub>2</sub> that is not observed in the surface air flask network measurements and for the differing trends in the reconstruction and the surface air flask network after the peak. Additionally, the bias could account for the unexpectedly high reconstructed Greenland H<sub>2</sub> levels relative to reconstructed Antarctic levels. We consider bias in the Greenland record to be more likely than bias in the Antarctic record because there is good model-measurement agreement for Ne enrichment at the two Antarctic ice core sites. Additionally, the individual



455 Antarctic reconstructions are in good agreement with each other and with high latitude Southern Hemisphere atmospheric  
flask measurements since the late 1980's (Figure 4b; Patterson et al., 2020; 2021).

Pore close-off fractionation is expected to substantially enrich H<sub>2</sub> and Ne levels in base of the firm (Section 3.1).  
δ<sup>22</sup>Ne/N<sub>2</sub> at the base of the firm was measured as 179‰ at NEEM and 112‰ at Summit where δ<sup>22</sup>Ne/N<sub>2</sub> of a sample is  
defined based on the relative mixing ratios of <sup>22</sup>Ne and N<sub>2</sub>:

460

$$(\delta^{22}\text{Ne}/\text{N}_2)_{\text{sample}} = [({}^{22}\text{Ne}/\text{N}_2)_{\text{sample}}/({}^{22}\text{Ne}/\text{N}_2)_{\text{atmosphere}} - 1] * 1000 \quad (11)$$

The UCI<sub>2</sub> firm air model generates enrichments of only 122‰ at NEEM and 90‰ at Summit at the base of the firm,  
significantly lower than the measurements (Figure 5). The Ne content of the firm air at Tunu has not been measured. As a  
465 first attempt to improve agreement between measured and modeled Ne enrichment, we re-tuned the closed porosity profiles  
for NEEM and Summit to optimize model-measurement agreement for Ne enrichment. However, the optimization yielded  
closed porosity profiles that are qualitatively different from previously published parameterizations and are probably  
unrealistic (e.g. Schwander et al., 1989; Goujon et al., 2003). Therefore, the parameterization of some physical process is  
likely incorrect. Possible candidates include the rate of densification, the rate of bubble close-off, and the rate of  
470 pressurization of bubbles. A detailed observation-based investigation of these processes at these sites would require field  
data that does not currently exist. Instead, we explored how altering the model parameterizations of these processes could  
improve the model-measurement agreement for Ne and examined the sensitivity of the firm air reconstruction to various  
physical assumptions.

475

480

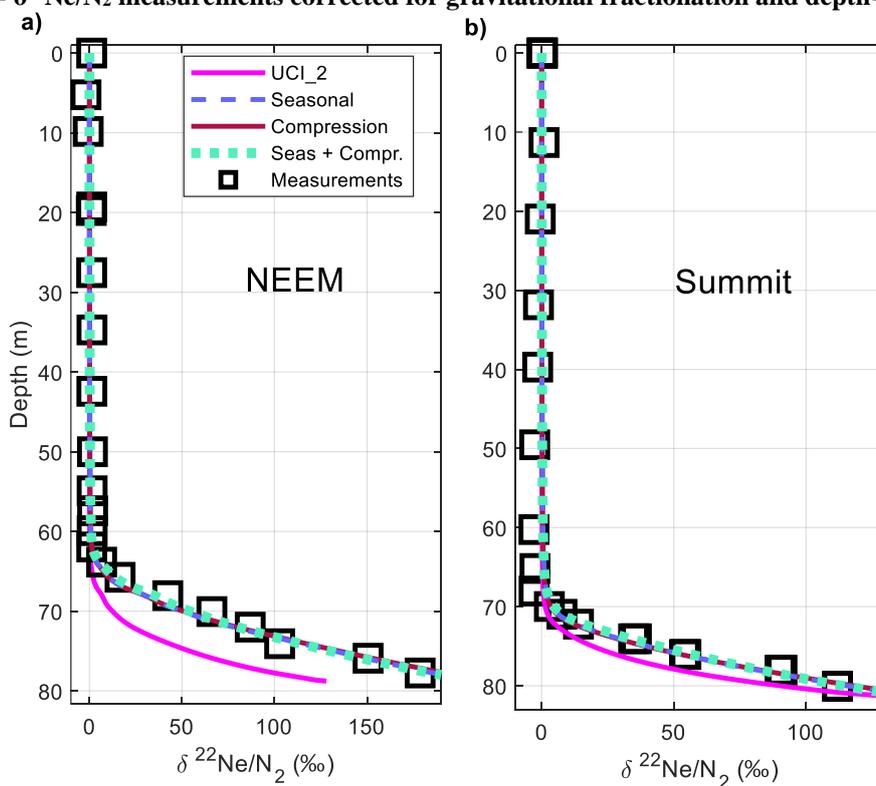
485

490



495

500 **Figure 5: Measured and modelled  $\delta^{22}\text{Ne}/\text{N}_2$  at NEEM (a) and Summit (b). Magenta lines-  $\delta^{22}\text{Ne}/\text{N}_2$  generated by the original UCI\_2 model; dashed lavender lines-  $\delta^{22}\text{Ne}/\text{N}_2$  generated by the model with seasonal layering implemented (Section 7.1); maroon lines-  $\delta^{22}\text{Ne}/\text{N}_2$  generated by the model with reduced bubble compression (Section 7.2); dotted cyan lines-  $\delta^{22}\text{Ne}/\text{N}_2$  generated by the model with both seasonal layering and reduced bubble compression (Section 7.3); black squares-  $\delta^{22}\text{Ne}/\text{N}_2$  measurements corrected for gravitational fractionation and depth-averaged.**



505

### 7.1 Method 1: Seasonal layering

Here we explored whether incorporating seasonal differences in densification could yield improved model-measurement agreement for Ne. Winter layers at both NEEM and Summit densify more rapidly than summer layers (Martinerie et al., 1992; Fujita et al., 2014). Fujita et al. (2014) attribute increased winter densification and deformation to seasonal differences in ion concentrations in the snow. Fujita et al. (2014) further show that bubbles occlude and pressurize preferentially in the dense winter layers before the summer layers. The UCI\_2 model (and most firn air models) utilizes seasonally averaged bulk density and porosity to calculate bubble close-off and pressurization. This should lead to errors because of the non-linearities in the parameterization. The importance of layering in the firn to bubble trapping has been discussed extensively by Mitchell et al. (2015) and Fourteau et al. (2017).



515 In this experiment, we divide each model layer into a winter and summer “sub-layer.” We arbitrarily specify that  
 half of the mass of each layer is winter accumulation and half is summer accumulation. The density of each sub-layer is  
 estimated based on Fujita et al., (2014; Figure S6), and the total porosity of the sub-layers is calculated using equation 1. The  
 bulk density of the firm is not changed from the original model. The original bulk closed porosity profile does not yield  
 agreement between measured and modeled Ne enrichment, so the closed porosity profile is revised. The closed porosity of  
 520 each sub layer ( $s_c$ ) is prescribed as a function of the total porosity and density of that sub layer:

$$s_c = \frac{s_{total}}{1 + \exp(-m(\rho_{firm} - \rho_0))} \quad (12)$$

Where  $m$  and  $\rho_0$  are free parameters that are tuned to yield agreement between measured and modeled Ne enrichment (Table  
 525 3; Figures 5 and 7). It has been demonstrated that winter and summer layers have similar close-off characteristics as a  
 function of density, so the tuneable parameters are the same for both winter and summer sub-layers (Martinerie et al., 1992;  
 Fourteau et al. 2017). Bubble close-off and pressurization is calculated independently for winter and summer sub-layers  
 using the same parameterization as in the UCI\_2 model after Severinghaus & Battle (2006; equations A11-A13). Then, the  
 bulk closed porosity and bubble pressure of each layer is determined from its constituent sub-layers, and the model is run as  
 530 before using the calculated bulk properties. New H<sub>2</sub> age distributions are calculated, and the new age distributions are used to  
 generate a revised joint reconstruction (Section 3.2; Figures 6 and 8).

## 7.2 Method 2: Reduced bubble compression

The bubble pressurization scheme used in the UCI\_2 model assumes that closed bubbles compress at the same rate  
 535 as the total porosity after Severinghaus & Battle (2006). Bubble trapping models that assume iso-compression have been  
 shown to overestimate total air content in bubbles at the Lock-in and Vostok sites in East Antarctica (Fourteau et al., 2017).  
 In that work, better model-measurement agreement was achieved by reducing the compressibility of closed bubbles by 50%.  
 Here, we explored whether reducing the compressibility of bubbles could yield improved model-measurement agreement for  
 Ne enrichment. We modified equation A12 in Severinghaus & Battle (2006), which calculates the incremental new volume  
 540 of air at ambient pressure ( $\Delta V_{b(i)}$ ) occluded in bubbles between layer ( $i-1$ ) and layer ( $i$ ) per unit volume of firm:

$$\Delta V_{b(i)} = \frac{2 \left\{ s_{c(i)} - s_{c(i-1)} \left[ (1 - k_{comp}) + k_{comp} \frac{s_{total(i)}}{s_{total(i-1)}} + \frac{\rho_{firm(i)}}{\rho_{firm(i-1)}} - 1 \right] \right\}}{(1 - k_{comp}) + k_{comp} \frac{s_{total(i)}}{s_{total(i-1)}} + 1} \quad (13)$$

Where  $k_{comp}$  is a dimensionless free parameter that represents the degree of compressibility of the bubbles. In the limit of  
 545  $k_{comp} = 1$ , equation 13 reduces to equation A12 of Severinghaus & Battle (2006), and bubbles compress at the same rate as the  
 open pores. In the limit of  $k_{comp} = 0$ , the bubbles are completely incompressible. Again, the original closed porosity profile



550 does not yield agreement with the measured Ne enrichment, so a new closed porosity profile is prescribed using equation 12. The three tuneable parameters from equation 12 and 13 ( $k_{comp}$ ,  $m$ , and  $\rho_0$ ) are simultaneously optimized to give the best fit between modeled and measured Ne enrichment (Table 3; Figures 5 and 7). Then, the model is used to calculate new age distributions for H<sub>2</sub>, which are used for new reconstructions (Figures 6 and 8).

### 7.3 Method 3: Seasonal layering and reduced bubble compression

555 Lastly, we incorporated both seasonal layering and bubbles with reduced compressibility into the UCI\_2 model. The density of the summer and winter layers is identical to Method 1. Instead of equation A12 from Severinghaus & Battle (2006), we use equation 13 to calculate bubble compression. Again,  $k_{comp}$ ,  $m$ , and  $\rho_0$  are tuned simultaneously to optimize the fit between measured and modeled Ne enrichment (Table 3; Figures 5 and 7). The model is run for H<sub>2</sub> and the new age distributions are used to generate new H<sub>2</sub> reconstructions (Figures 6 and 8). We note that all three new parameterizations have a negligible effect on the modelling of trace gases like CO<sub>2</sub> and CH<sub>4</sub>, which are not affected by pore close-off fractionation (Figure S7)

560

**Table 3-** Optimized tunable parameters for each method described in 7.1-7.3

Site	Method 1			Method 2			Method 3		
	$\rho_0$ (kg m <sup>-3</sup> )	$m$ (m <sup>3</sup> kg <sup>-1</sup> )	$k_{comp}$	$\rho_0$ (kg m <sup>-3</sup> )	$m$ (m <sup>3</sup> kg <sup>-1</sup> )	$k_{comp}$	$\rho_0$ (kg m <sup>-3</sup> )	$m$ (m <sup>3</sup> kg <sup>-1</sup> )	$k_{comp}$
NEEM	827	0.070	1.00	815	0.070	0.75	812	0.138	0.69
Summit	832	0.070	1.00	821	0.070	0.75	818	0.138	0.69

565

### 7.4 Revised reconstructions

570 The reconstructions generated by the three methods are very similar (Figure 8). The most significant consequence of increasing pore close-off fractionation is a reduction in the atmospheric maximum around 1990 and a more modest decrease in atmospheric levels from 1990-2005. That is, higher modeled enrichment due to pore close-off fractionation yields correspondingly lower reconstructed H<sub>2</sub> levels. The 1990 maximum in the revised reconstruction is ~530 ppb, compared to 547 ppb in the original. After 1990, H<sub>2</sub> levels in the revised reconstructions decrease very slowly to 515 ppb in 2013, converging with the original reconstruction. The modest decrease in H<sub>2</sub> levels during the 1990's in the revised reconstructions is in better agreement with high northern latitude flask measurements and synthetic Summit history, which do not show a discernible trend in atmospheric H<sub>2</sub> during the 1990's.

575

Because the older firm air that was sampled at Tunu is excluded from this analysis, the reconstructions are only meaningful after about 1950. Overall, H<sub>2</sub> levels in the revised reconstructions are reduced compared to the original



reconstruction by 15-30 ppb from 1950-1990 (Figure 8). The new reconstructions remain 0-15 ppb higher than the Antarctic reconstructions, so the unexpectedly high reconstructed Greenland levels relative to Antarctic levels can be only partially explained by the underestimate of pore close-off induced enrichment (Figure 9).

580 We can draw several important conclusions from these results: 1) H<sub>2</sub> atmospheric reconstructions are sensitive to pore close-off fractionation, 2) the large 1990's peak in the Greenland reconstructions is likely an artefact of underestimated pore close-off fractionation in the UCI\_2 model, 3) as long as sufficient fractionation is achieved, the reconstructions are insensitive to the specific parameterization and internal model physics. All three of the approaches resulted in an increase in the integrated area of the age distributions and a slightly older shift in mean age (Figure 6). All three methods greatly  
585 improve the fit between modelled and measured Ne enrichment at both sites. The similarity in optimized tuneable parameters suggest that a common mechanism underlies the UCI\_2 model's underestimate of pore close-off induced enrichment at both NEEM and Summit (Table 3). It is likely that the common mechanism is associated with bubble close-off and pressurization. Further investigation of layering and bubble close-off using high resolution measurements of density, porosity, and air content, in conjunction with measurements of highly diffusive trace gases like Ne or He would be useful in constraining pore  
590 close-off fractionation for future Greenland firm air work.

595

600

605

610

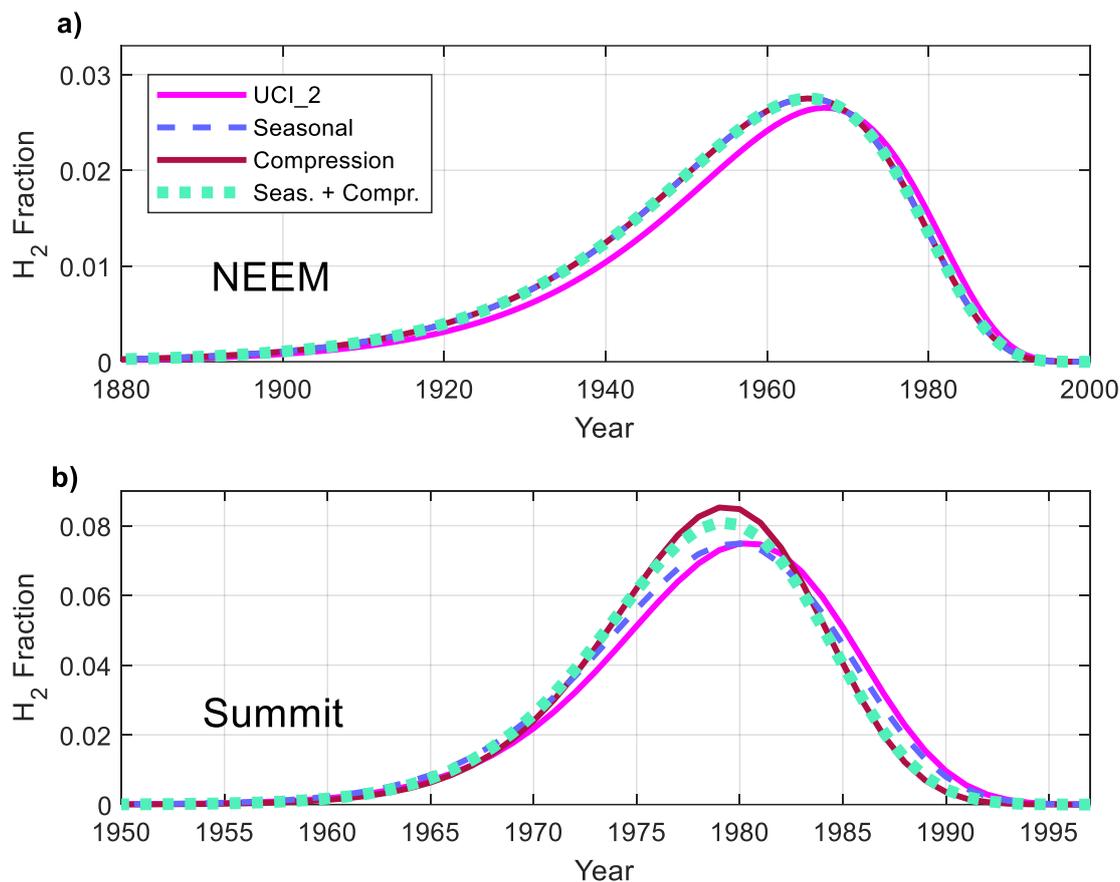
615



620

Figure 6: Open porosity H<sub>2</sub> age distributions resulting from several approaches to approximating pore close-off fractionation for H<sub>2</sub> at a depth of 75.9 m at NEEM (a) and at a depth of 80.1 m at Summit (b). Magenta lines- original UCI\_2 model; dashed lavender lines, maroon lines, and dotted cyan lines are from the model with seasonal layering implemented (Section 7.1), with reduced bubble compression (Section 7.2), and with both seasonal layering and reduced bubble compression (Section 7.3) respectively.

625



630

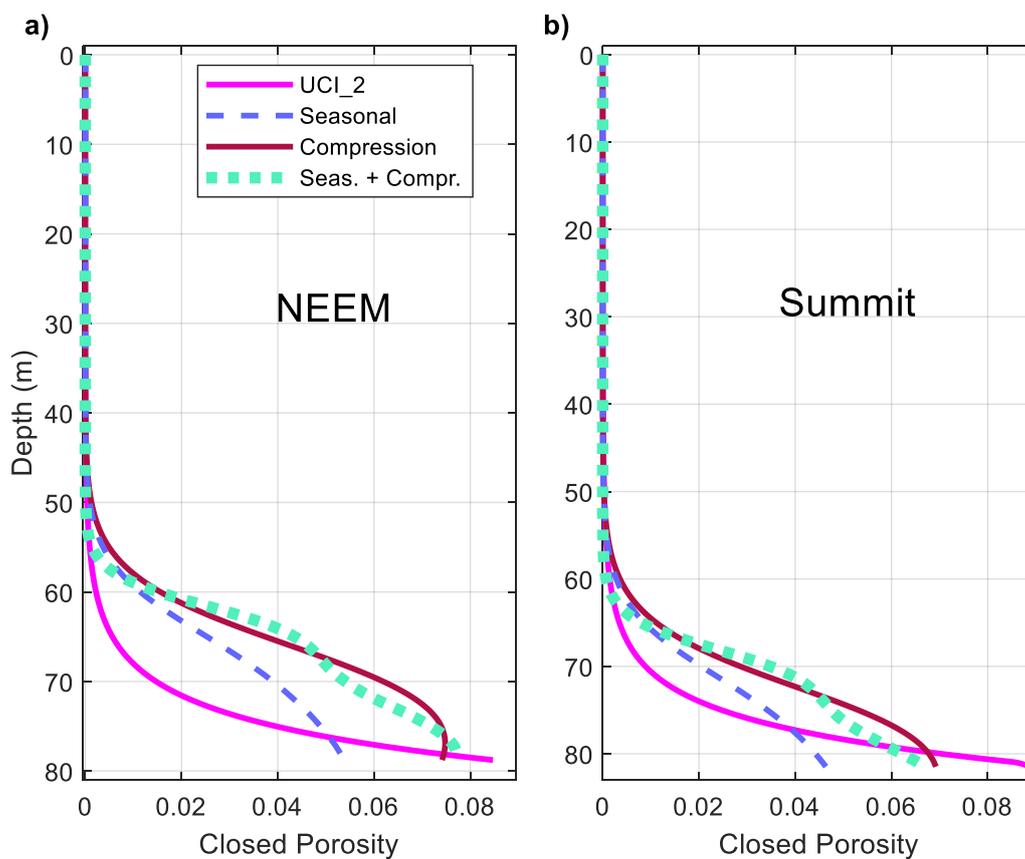
635



640

Figure 7: Model closed porosity profiles used for NEEM (a) and Summit (b). Magenta lines- original closed porosity profile used in the UCI\_2 model; dashed lavender lines, maroon lines, and dotted cyan lines are returned closed porosity profiles used for the model with seasonal layering implemented (Section 7.1), with reduced bubble compression (Section 7.2), and with both seasonal layering and reduced bubble compression (Section 7.3) respectively.

645



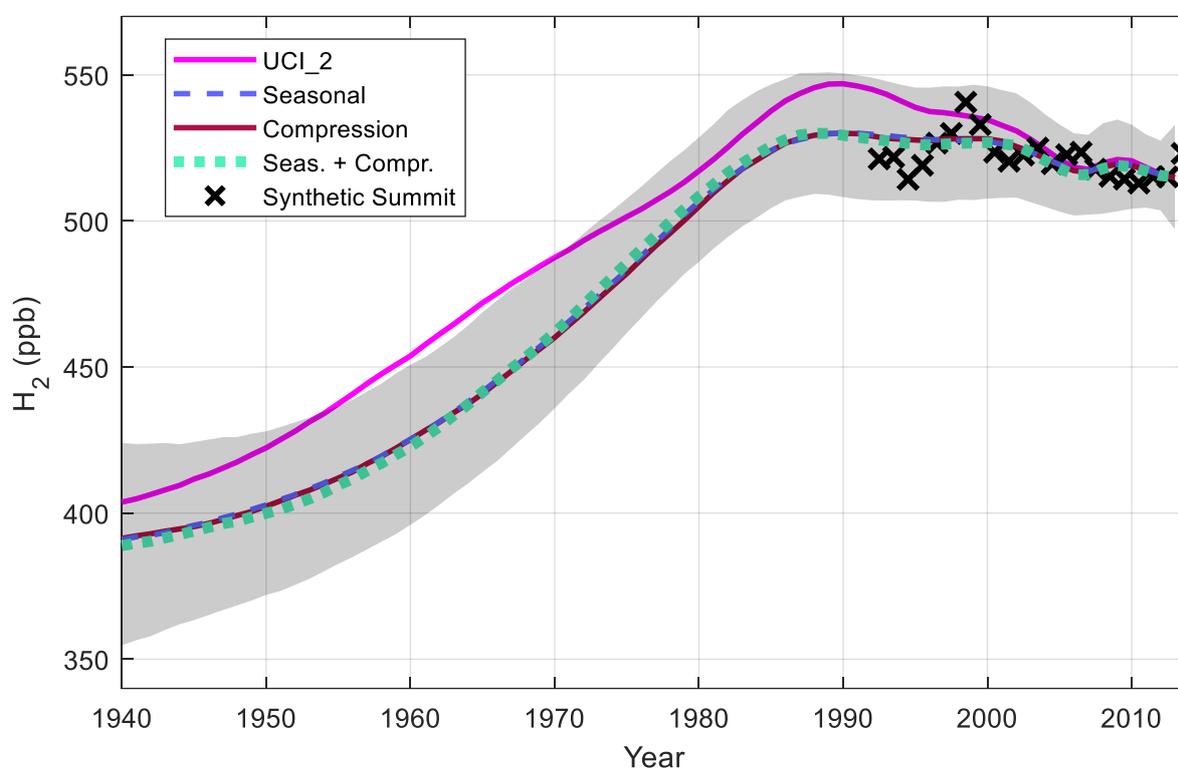
650

655

660



665 **Figure 8: Reconstructions of atmospheric H<sub>2</sub> levels from Greenland firn air. Magenta line- joint reconstruction as in**  
**Figure 2a; dashed lavender lines, maroon lines, and dotted cyan lines are joint reconstructions calculated from the**  
**model with seasonal layering implemented (Section 7.1), with reduced bubble compression (Section 7.2), and with**  
**both seasonal layering and reduced bubble compression (Section 7.3) respectively; black x's annual mean synthetic**  
**Summit H<sub>2</sub> history (Section 5; Pétron et al., 2023, Langenfelds et al., 2002). Grey shading is the approximate ±1σ**  
**uncertainty band for the three corrected reconstructions. The upper bound is the maximum of the sum of the three**  
**histories and their respective 1σ uncertainties for each year, and lower bound is the minimum of the difference**  
670 **between the three histories and their respective 1σ uncertainties.**



675

### 8. *In situ* production

It is possible that some of the differences between the independent reconstructions (Figure 1) could be caused by *in situ* production, but little is known about *in situ* H<sub>2</sub> production in snow and ice. It is useful to consider evidence for *in situ* production of CO because both trace gases may be photochemically produced from carbonyl-containing precursors. CO has  
680 been shown to be produced photochemically from CH<sub>2</sub>O in surface snow (Hahn et al., 2001). At Devon Island, Canada CO levels in the firn increased steadily with depth in both the diffusive and lock-in zones, exhibiting clear evidence of *in situ* production (Clark et al. 2007). By contrast, NEEM and Summit firn air both exhibit CO maxima in the firn air column



associated with a late 20<sup>th</sup> century atmospheric maximum, with CO levels decreasing with increasing depth below (Petrenko et al., 2013). Petrenko et al. (2013) concluded in situ production at those sites was not sufficient to bias the atmospheric reconstruction. Those researchers used three lines of evidence to support their conclusion. First, the ice age at the depth of the observed lock-in zone CO peak is different at the two sites, so the peak cannot be attributed to a widely deposited layer of organic impurities. Second, the CO/CH<sub>4</sub> ratio at a given depth is in reasonable agreement at NEEM and Summit. CO and CH<sub>4</sub> have similar gas-phase diffusivities, so the mean age of the two gases at a given depth is comparable. CH<sub>4</sub> is known to be well preserved in Greenland firn air. If *in situ* production differentially affected CO levels at the two sites, it would be expected that CO/CH<sub>4</sub> ratios would diverge. Third, forcing a firn air model with an atmospheric flask derived CO history yields good agreement with the measured depth profile down to the top of lock-in zone where the firn air is too old to be constrained by flask measurements.

Greenland ice cores do exhibit evidence of in situ CO production. Early discrete measurements on Eurocore ice showed low, steady preindustrial CO levels (Hahn and Raynaud, 1998; Hahn et al., 1996). By contrast, Fain et al. (2022) reported higher and more variable CO levels from several Greenland sites (including NEEM and Tunu) using continuous flow and discrete methods. The elevated CO levels were apparently associated with localized organic impurities in the ice. The high variability in CO detected by the continuous flow method is not compatible with continuous production in the firn, where such signals should be smoothed by diffusion. It is therefore likely that the high frequency variability observed by Fain et al. (2022) originated below the lock-in zone. However, the ice core data do not preclude the possibility of some in situ production having occurred in the overlying firn.

The production of CO in Greenland firn air and ice is not well understood, and no physical mechanism has been proposed that can explain all of the observations. If the production is photochemical, one would expect it to be much faster in the firn than in ice, due to the much larger actinic flux. Until the mechanism is established, it is difficult to assess whether and to what extent *in situ* production might influence firn air H<sub>2</sub>. Depending on the depth and mechanism, production of H<sub>2</sub> in the firn column would cause a high bias during some or all of our atmospheric reconstructions. Further research on this issue is needed.

## 9. Summary and conclusions

In this study we use all available firn air measurements of H<sub>2</sub> at Greenland and Antarctic sites to reconstruct atmospheric H<sub>2</sub> levels over the past century. The joint Antarctic reconstruction covers the 1900-2003 period and is based on firn air measurements from South Pole and Megadunes (Patterson et al., 2020; 2021). That reconstruction shows good site-to-site agreement and good agreement with the atmospheric flask network since the late 1980's. The good fit between measured and simulated Ne enrichment obtained using the UCI\_2 firn air model gives us high confidence in the modelling of pore close-off fractionation at the Antarctic sites.

Reconstructing H<sub>2</sub> levels over Greenland is more challenging because of inter-site biases and because the standard UCI\_2 firn air model underestimates the enrichment of Ne at NEEM and Summit. It is likely that the model also



underestimates pore close-off induced enrichment of H<sub>2</sub> because of the similar permeability of the two gases in ice (Patterson & Saltzman, 2021; Satoh et al.1996). Three alternative parameterizations of pore close-off fractionation were explored to assess the sensitivity of the reconstructions to an underestimate of pore close-off induced enrichment. These revised  
720 reconstructions all show a less prominent atmospheric maximum and more stable H<sub>2</sub> levels during the late 20<sup>th</sup> century, in better agreement with the atmospheric flask measurements.

There is clearly uncertainty regarding the modelling of the physical processes influencing pore close-off fractionation, particularly in Greenland firn air. A path to increasing confidence in firn air reconstructions for H<sub>2</sub> is to obtain new observations of firn air chemistry with high resolution measurements of firn density, porosity, and bubble total air  
725 content. Absent additional measurements, we believe that the joint reconstructions presented here are the best estimate for the evolution (and uncertainty) of atmospheric H<sub>2</sub> during the 20<sup>th</sup> century over Antarctica and Greenland (Figure 9). Ice core records at high accumulation sites should also be collected in order to constrain H<sub>2</sub> levels over the late pre-industrial and industrial eras. Such records will be important to better understanding the global H<sub>2</sub> cycle, the impact of human activities on H<sub>2</sub> in the past, and the atmospheric response to increasing anthropogenic H<sub>2</sub> emissions in a warming climate.

One surprising result is that our reconstructions show higher H<sub>2</sub> levels over Greenland than over Antarctica prior to  
730 1990. This implies a reversal in the sign of the interhemispheric difference during the late 20<sup>th</sup> century. It is believed that Northern hemispheric H<sub>2</sub> levels are lower than Southern hemisphere levels today due to strong uptake of atmospheric H<sub>2</sub> by Northern hemisphere soils. The N/S interhemispheric difference in atmospheric H<sub>2</sub> occurs in spite of the fact that direct anthropogenic emissions of H<sub>2</sub> occur predominantly in the northern hemisphere. The dominance of northern hemispheric soil  
735 sink is largely a function of the asymmetry in land area between the two hemispheres. While the soil sink may vary over time due to changes in ecosystems, land use, or climate, a dramatic change would be required in order to have caused a reversal of the interhemispheric difference. In this respect, the firn air reconstructions present a challenge to our understanding of the global atmospheric H<sub>2</sub> budget and careful examination of our conclusion that the northern hemisphere firn air reconstruction is not biased by *in situ* H<sub>2</sub> production in Greenland firn.

740

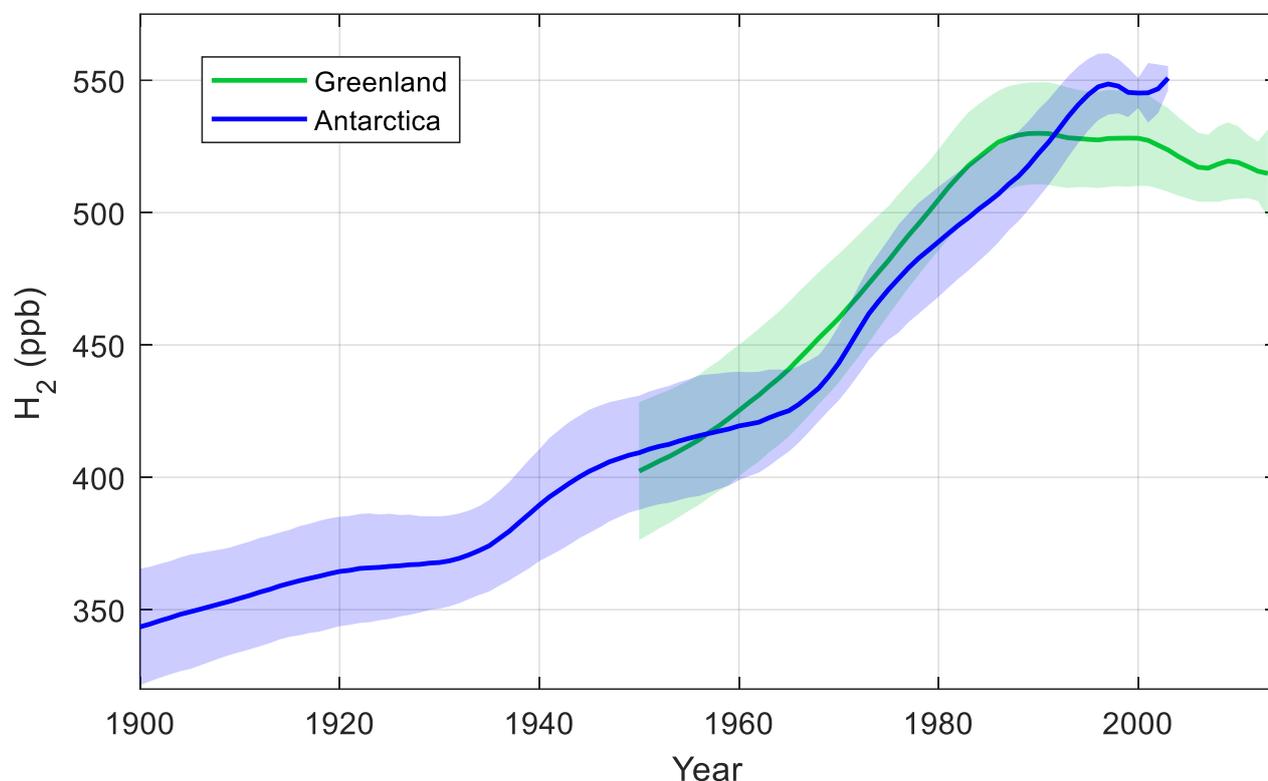
745

750



755

**Figure 9: Best estimates for the atmospheric history of H<sub>2</sub> in Greenland and Antarctica over the past century. Blue line- joint Antarctic firn air reconstruction as in Figure 4b; Green line- joint Greenland firn air reconstruction obtained from the mean of the three reconstructions using revised pore close-off parameterizations (Seasonal, Compression, and Seasonal+Compression; shown in Figure 8). Shaded areas are ±1σ uncertainties.**



760 **Data and code availability**

The data and code used in this research has been submitted to the DRYAD data repository, and we will provide DOI identifiers before the end of the review process.

**Author contributions**

765 J.D.P. carried out the model simulations and data analysis; M.A. contributed to data analysis; A.M.C., G.P., R.L.L. and P.B.K. made calibration corrections and contributed atmospheric flask measurements; J.P.S. collected the field samples, measured  $\delta^{22}\text{Ne}/\text{N}_2$ , and assisted with firn air modeling; V.V.P collected field samples and contributed to firn air measurements, E.S.S. contributed to modeling, data analysis, and editing the manuscript; and J.D.P. wrote the manuscript.



## 770 Competing interests

The authors declare that they have no conflict of interest.

## Acknowledgements

The authors would like to thank Christo Buizert for helpful discussions about firm air modelling. This research was supported  
775 by the NSF (OPP-1907974 and OPP-2019719).

## References

- Aydin, M., Britten, G. L., Montzka, S. A., Buizert, C., Primeau, F., Petrenko, V., Battle, M. B., Nicewonger, M. R.,  
780 Patterson, J., Hmiel, B., & Saltzman, E. S. (2020). Anthropogenic Impacts on Atmospheric Carbonyl Sulfide Since  
the 19th Century Inferred From Polar Firm Air and Ice Core Measurements. *Journal of Geophysical Research:  
Atmospheres*, 125(16), 1–23. <https://doi.org/10.1029/2020JD033074>
- Bader, H. (1964). Density of Ice as a Function of Temperature and Stress. *Special Report (U.S. Army Cold Regions Research  
and Engineering Laboratory)*, 64, 1–6.
- Battle, M., Bender, M., Sowers, T., Tans, P. P., Butler, J. H., Elkins, J. W., Ellis, J. T., Conway, T., Zhang, N., Lang, P., &  
785 Clarke, A. D. (1996). Atmospheric gas concentrations over the past century measured in air from firm at the South  
Pole. *Nature*, 383(6597), 231–235. <https://doi.org/10.1038/383231a0>
- Buizert, C. (2013). Studies of firm air. In S. A. Elias & C. J. Mock (Eds.), *Encyclopedia of Quaternary Science 2* (pp. 361–  
372). Elsevier.
- 790 Buizert, C., Martinerie, P., Petrenko, V. v., Severinghaus, J. P., Trudinger, C. M., Witrant, E., Rosen, J. L., Orsi, A. J.,  
Rubino, M., Etheridge, D. M., Steele, L. P., Hogan, C., Laube, J. C., Sturges, W. T., Levchenko, V. A., Smith, A. M.,  
Levin, I., Conway, T. J., Dlugokencky, E. J., ... Blunier, T. (2012). Gas transport in firm: Multiple-tracer  
characterisation and model intercomparison for NEEM, Northern Greenland. *Atmospheric Chemistry and Physics*,  
12(9), 4259–4277. <https://doi.org/10.5194/acp-12-4259-2012>
- 795 Butler, J. H., Battle, M., Bender, M. L., Montzka, S. A., Clarke, A. D., Saltzman, E. S., Sucher, C. M., Severinghaus, J. P., &  
Elkins, J. W. (1999). A record of atmospheric halocarbons during the twentieth century from polar firm air. *Nature*,  
399(6738), 749–755. <https://doi.org/10.1038/21586>
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., &  
800 Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1).  
<https://doi.org/10.18637/jss.v076.i01>
- Clark, I. D., Henderson, L., Chappellaz, J., Fisher, D., Koerner, R., Worthy, D. E. J., Kotzer, T., Norman, A. L., & Barnola,  
J. M. (2007). CO<sub>2</sub> isotopes as tracers of firm air diffusion and age in an Arctic ice cap with summer melting, Devon  
Island, Canada. *Journal of Geophysical Research Atmospheres*, 112(1). <https://doi.org/10.1029/2006JD007471>
- 805 Derwent, R. G., Stevenson, D. S., Utembe, S. R., Jenkin, M. E., Khan, A. H., & Shallcross, D. E. (2020). Global modelling  
studies of hydrogen and its isotopomers using STOCHEM-CRI: Likely radiative forcing consequences of a future  
hydrogen economy. *International Journal of Hydrogen Energy*, 45(15), 9211–9221.  
<https://doi.org/10.1016/j.ijhydene.2020.01.125>



- 810 Dlugokencky, E., & Lan, X. (2020a). *Atmospheric Carbon Dioxide Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1968-2019*. NOAA/ESRL. [ftp://aftp.cmdl.noaa.gov/data/trace\\_gases/co2/](ftp://aftp.cmdl.noaa.gov/data/trace_gases/co2/)
- Dlugokencky, E., & Lan, X. (2020b). *Atmospheric Methane Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air Sampling Network, 1983-2019*. NOAA/ESRL. [ftp://aftp.cmdl.noaa.gov/data/trace\\_gases/ch4/](ftp://aftp.cmdl.noaa.gov/data/trace_gases/ch4/)
- Ehhalt, D. H., & Rohrer, F. (2009). The tropospheric cycle of H<sub>2</sub>: A critical review. *Tellus, Series B: Chemical and Physical Meteorology*, 61(3), 500–535. <https://doi.org/10.1111/j.1600-0889.2009.00416.x>
- 815 Etheridge, D. M., Steele, L. P., Francey, R. J., & Langenfelds, R. L. (1998). Atmospheric methane between 1000 A . D . and present : Evidence of anthropogenic emissions and climatic variability. *Journal of Geophysical Research*, 103(D13), 15979–15993.
- Etheridge, D. M., Steele, L. P., Langenfelds, R. L., & Francey, R. J. (1996). Natural and anthropogenic changes in atmospheric CO<sub>2</sub> over the last 1000 years from air in Antarctic ice and firn. *Journal of Geophysical Research*, 101(D2), 4115–4128. <https://doi.org/10.1074/jbc.271.44.27509>
- 820
- Faïn, X., Rhodes, R. H., Place, P., Petrenko, V. V., Fourteau, K., Chellman, N., Crosier, E., McConnell, J. R., Brook, E. J., Blunier, T., Legrand, M., & Chappellaz, J. (2022). Northern Hemisphere atmospheric history of carbon monoxide since preindustrial times reconstructed from multiple Greenland ice cores. *Climate of the Past*, 18(3), 631–647. <https://doi.org/10.5194/cp-18-631-2022>
- 825
- Fourteau, K., Faïn, X., Martinerie, P., Landais, A., Ekaykin, A., Lipenkov, V., & Chappellaz, J. (2017). Analytical constraints on layered gas trapping and smoothing of atmospheric variability in ice under low-accumulation conditions. *Climate of the Past*, 13(12), 1815–1830. <https://doi.org/10.5194/cp-13-1815-2017>
- Fujita, S., Hirabayashi, M., Goto-Azuma, K., Dallmayr, R., Satow, K., Zheng, J., & Dahl-Jensen, D. (2014). Densification of layered firn of the ice sheet at NEEM, Greenland. *Journal of Glaciology*, 60(223), 905–921. <https://doi.org/10.3189/2014JG14J006>
- 830
- Fuller, E. N., Schettler, P. D., & Giddings, J. C. (1966). New method for prediction of binary gas-phase diffusion coefficients. *Journal of Industrial and Engineering Chemistry*, 58, 18–27.
- Goujon, C., Barnola, J. M., & Ritz, C. (2003). Modelling the densification of polar firn including heat diffusion: Application to close-off characteristics and gas isotopic fractionation for Antarctica and Greenland sites. *Journal of Geophysical Research D: Atmospheres*, 108(24). <https://doi.org/10.1029/2002jd003319>
- 835
- Haan, D., Martinerie, P., & Raynaud, D. (1996). Ice core data of atmospheric carbon monoxide over Antarctica and Greenland during the last 200 years. *Geophysical Research Letters*, 23(17), 2235–2238. <https://doi.org/10.1029/96GL02137>
- 840
- Haan, D., & Raynaud, D. (1998). Ice core record of CO variations during the last two millennia: atmospheric implications and chemical interactions within the Greenland ice. *Tellus B: Chemical and Physical Meteorology*, 50(3), 253. <https://doi.org/10.3402/tellusb.v50i3.16101>
- Haan, D., Zuo, Y., Gros, V., & Brenninkmeijer, C. A. M. (2001). Photochemical Production of Carbon Monoxide in Snow. In *Journal of Atmospheric Chemistry* (Vol. 40).



- 845 Hmiel, B., Petrenko, V. v., Dyonisius, M. N., Buizert, C., Smith, A. M., Place, P. F., Harth, C., Beaudette, R., Hua, Q., Yang, B., Vimont, I., Michel, S. E., Severinghaus, J. P., Etheridge, D., Bromley, T., Schmitt, J., Fain, X., Weiss, R. F., & Dlugokencky, E. (2020). Preindustrial 14CH<sub>4</sub> indicates greater anthropogenic fossil CH<sub>4</sub> emissions. *Nature*, 578(7795), 409–412. <https://doi.org/10.1038/s41586-020-1991-8>
- Jordan, A., & Steinberg, B. (2011). Calibration of atmospheric hydrogen measurements. *Atmospheric Measurement Techniques*, 4(3), 509–521. <https://doi.org/10.5194/amt-4-509-2011>
- 850 Khalil, M. A. K., & Rasmussen, R. A. (1990). Global increase of atmospheric molecular hydrogen. *Nature*, 347, 743–745. [https://doi.org/10.1016/0021-9797\(80\)90501-9](https://doi.org/10.1016/0021-9797(80)90501-9)
- Langenfelds, R. L., Francey, R. J., Pak, B. C., Steele, L. P., Lloyd, J., Trudinger, C. M., & Allison, C. E. (2002). Interannual growth rate variations of atmospheric CO<sub>2</sub> and its  $\delta^{13}\text{C}$ , H<sub>2</sub>, CH<sub>4</sub>, and CO between 1992 and 1999 linked to biomass burning. *Global Biogeochemical Cycles*, 16(3). <https://doi.org/10.1029/2001gb001466>
- 855 Martinerie, P., Raynaud, D., Etheridge, D. M., Barnola, J.-M., & Mazaudier, D. (1992). Physical and climatic parameters which influence the air content in polar ice. In *Earth and Planetary Science Letters* (Vol. 112).
- Mitchell, L. E., Buizert, C., Brook, E. J., Breton, D. J., Fegyveresi, J., Baggenstos, D., Orsi, A., Severinghaus, J., Alley, R. B., Albert, M., Rhodes, R. H., McConnell, J. R., Sigl, M., Maselli, O., Gregory, S., & Ahn, J. (2015). Observing and modeling the influence of layering on bubble trapping in polar firn. *Journal of Geophysical Research*, 120(6), 2558–2574. <https://doi.org/10.1002/2014JD022766>
- 860 Novelli, P. C. (2006). *Atmospheric Hydrogen Mixing Ratios from the NOAA GMD Carbon Cycle Cooperative Global Air Sampling Network, 1988-2005*. NOAA/ESRL.
- Novelli, P. C., Crotwell, A. M., & Hall, B. D. (2009). Application of gas chromatography with a pulsed discharge helium ionization detector for measurements of molecular hydrogen in the atmosphere. *Environmental Science and Technology*. <https://doi.org/10.1021/es803180g>
- 865 Novelli, P. C., Lang, P. M., Masarie, K. A., Hurst, D. F., Myers, R., & Elkins, J. W. (1999). Molecular hydrogen in the troposphere: Global distribution and budget. *Journal of Geophysical Research*, 104(D23), 30427–30444.
- 870 Patterson, J. D., Aydin, M., Crotwell, A. M., Pétron, G., Severinghaus, J. P., Krummel, P. B., Langenfelds, R. L. & Saltzman, E. S. (2021). H<sub>2</sub> in Antarctic firn air: Atmospheric reconstructions and implications for anthropogenic emissions. *Proceedings of the National Academy of Sciences of the United States of America*, 118. <https://doi.org/10.1073/pnas.2103335118/-/DCSupplemental.90>
- 875 Patterson, J. D., Aydin, M., Crotwell, A. M., Petron, G., Severinghaus, J. P., & Saltzman, E. S. (2020). Atmospheric History of H<sub>2</sub> Over the Past Century Reconstructed From South Pole Firn Air. *Geophysical Research Letters*, 47(14), 1–8. <https://doi.org/10.1029/2020GL087787>
- Patterson, J. D., & Saltzman, E. S. (2021). Diffusivity and solubility of H<sub>2</sub> in ice Ih: Implications for the behavior of H<sub>2</sub> in polar ice. *Journal of Geophysical Research Atmospheres*, 126, 1–14. <https://doi.org/10.1029/2020JD033840>
- 880 Paulot, F., Paynter, D., Naik, V., Malyshev, S., Menzel, R., & Horowitz, L. W. (2021). Global modeling of hydrogen using GFDL-AM4.1: Sensitivity of soil removal and radiative forcing. *International Journal of Hydrogen Energy*, 46(24), 13446–13460. <https://doi.org/10.1016/j.ijhydene.2021.01.088>



- 885 Petrenko, V. v., Martinerie, P., Novelli, P., Etheridge, D. M., Levin, I., Wang, Z., Blunier, T., Chappellaz, J., Kaiser, J.,  
Lang, P., Steele, L. P., Hammer, S., Mak, J., Langenfelds, R. L., Schwander, J., Severinghaus, J. P., Witrant, E.,  
Petron, G., Battle, M. O., ... White, J. W. C. (2013). A 60 yr record of atmospheric carbon monoxide reconstructed  
from Greenland firn air. *Atmospheric Chemistry and Physics*, 13(15), 7567–7585. <https://doi.org/10.5194/acp-13-7567-2013>
- 890 Pétron, G., Crotwell, A., Kitzis, M., Madronich, D., Mefford, T., Moglia, E., Mund, J., Neff, D., Thoning, K., & Wolter, S.  
(2023). *Atmospheric Hydrogen Dry Air Mole Fractions from the NOAA GML Carbon Cycle Cooperative Global Air  
Sampling Network, 2009-Present [Data set]*. <https://doi.org/10.15138/WP0W-EZ08>
- 895 Pieterse, G., Krol, M. C., Batenburg, A. M., M. Brenninkmeijer, C. A., Popa, M. E., O'Doherty, S., Grant, A., Steele, L. P.,  
Krummel, P. B., Langenfelds, R. L., Wang, H. J., Vermeulen, A. T., Schmidt, M., Yver, C., Jordan, A., Engel, A.,  
Fisher, R. E., Lowry, D., Nisbet, E. G., ... Röckmann, T. (2013). Reassessing the variability in atmospheric H<sub>2</sub> using  
the two-way nested TM5 model. *Journal of Geophysical Research: Atmospheres*, 118(9), 3764–3780.  
<https://doi.org/10.1002/jgrd.50204>
- Pieterse, G., Krol, M. C., Batenburg, A. M., Steele, L. P., Krummel, P. B., Langenfelds, R. L., & Röckmann, T. (2011).  
Global modelling of H<sub>2</sub> mixing ratios and isotopic compositions with the TM5 model. *Atmospheric Chemistry and  
Physics*, 11(14), 7001–7026. <https://doi.org/10.5194/acp-11-7001-2011>
- 900 Prather, M. J. (2003). An environmental experiment with H<sub>2</sub>. *Science*, 302(5645), 581–582.  
<https://doi.org/10.1126/science.1091060>
- 905 Prinn, R. G., Weiss, R. F., Aduini, J., Arnold, T., Dewitt, H., Fraser, P. J., Ganesan, A. L., Gasore, J., Harth, C. M.,  
Hermansen, O., Kim, J., Krummel, P. B., Li, S., Loh, Z. M., Lunder, C. R., Maione, M., Manning, A. J., Miller, B. R.,  
Mitreveski, B., ... Zhou, L. (2019). *History of chemically and radiatively important atmospheric gases from the  
Advanced Global Atmospheric Gases Experiment (AGAGE)*. Carbon Dioxide Information Analysis Center (CDIAC),  
Oak Ridge National Laboratory (ORNL), Oak Ridge, (United States).  
<https://doi.org/https://doi.org/10.3334/CDIAC/atg.db1001>
- Reid, R. C., Prausnitz, J. M., Sherwood, T. K., & Poling, B. E. (1987). *The properties of gases and liquids* (4th ed.).  
McGraw-Hill.
- 910 Rommelaere, V., Arnaud, L., & Barnola, J. (1997). Reconstructing recent atmospheric trace gas concentrations from polar  
firn and bubbly ice data by inverse methods. *Journal of Geophysical Research*, 102(D25), 30069–30083.
- Satoh, K., Uchida, T., Hondoh, T., & Mae, S. (1996). Diffusion coefficient and solubility measurements of noble gases in ice  
crystals. *Proc. NIPR Symp. Polar Meteorol. G]Aciol*, 10, 73–81.
- Schwander, J., Barnola, J. M., Andrie, C., Leuenberger, M., Ludin, A., Raynaud, D., & Stauffer, B. (1993). The Age of the  
Air in the Firn and the Ice at Summit, Greenland. *Journal of Geophysical Research*, 98, 2831–2838.
- 915 Schwander, J., Oeschger, H., & Langway, C. C. (1989). *The Environmental Record in Glaciers and Ice Sheets*. John Wiley.
- Severinghaus, J. P., Albert, M. R., Courville, Z. R., Fahnestock, M. A., Kawamura, K., Montzka, S. A., Mühle, J., Scambos,  
T. A., Shields, E., Shuman, C. A., Suwa, M., Tans, P., & Weiss, R. F. (2010). Deep air convection in the firn at a  
zero-accumulation site, central Antarctica. *Earth and Planetary Science Letters*, 293(3–4), 359–367.  
<https://doi.org/10.1016/j.epsl.2010.03.003>



- 920 Severinghaus, J. P., & Battle, M. O. (2006). Fractionation of gases in polar ice during bubble close-off: New constraints from firn air Ne, Kr and Xe observations. *Earth and Planetary Science Letters*.  
<https://doi.org/10.1016/j.epsl.2006.01.032>
- Severinghaus, J. P., Grachev, A., & Battle, M. (2001). Seasonal Temperature Gradients. *Geochemistry Geophysics Geosystems*, 2(7).
- 925 Trudinger, C. M., Enting, I. G., Etheridge, D. M., Francey, R. J., Levchenko, V. A., Steele, L. P., Raynaud, D., & Arnaud, L. (1997). Modeling air movement and bubble trapping in firn. *Journal of Geophysical Research Atmospheres*.  
<https://doi.org/10.1029/96JD03382>
- Wang, D., Jia, W., Olsen, S. C., Wuebbles, D. J., Dubey, M. K., Rockett, A. A., & Wuebbles, C. D. J. (2013a). Impact of a future H<sub>2</sub>-based road transportation sector on the composition and chemistry of the atmosphere – Part 1 : Tropospheric composition and air quality. *Atmospheric Chemistry and Physics*, 13(2), 6117–6137.  
<https://doi.org/10.5194/acp-13-6117-2013>
- 930 Wang, D., Jia, W., Olsen, S. C., Wuebbles, D. J., Dubey, M. K., Rockett, A. A., & Wuebbles, C. D. J. (2013b). Impact of a future H<sub>2</sub>-based road transportation sector on the composition and chemistry of the atmosphere – Part 2 : Stratospheric ozone. *Atmospheric Chemistry and Physics*, 13, 6139–6150. <https://doi.org/10.5194/acp-13-6139-2013>
- 935 Warwick, N., Griffiths, P., Keeble, J., Archibald, A., & Pyle, J. (2022). *Atmospheric implications of increased Hydrogen use*. <https://www.gov.uk/government/publications/atmospheric-implications-of-increased-hydrogen-use>