Imprint of anthropogenic sources and soil removal on the surface concentration of H_2 in the contiguous US

Fabien Paulot,[*,†] Gabrielle Pétron,[‡,¶] Andrew Crotwell,[‡,¶] Molly Crotwell,[‡,¶] Philip Handley,[‡,¶] Jonathan Kofler,[‡,¶] Monica Madronich,[‡,¶] Thomas Mefford,[‡,¶] Eric Moglia,[‡,¶] John Mund,[‡,¶] Kirk Thoning,[¶] Sébastien Biraud,[§] and Arlyn Andrews[¶]

†*Geophysical Fluid Dynamics Laboratory, National Oceanic and Atmospheric Administration, Princeton, NJ, USA*

‡*Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA*

¶*Global Monitoring Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA*

§*Climate Sciences Department, Lawrence Berkeley National Laboratory, Berkeley, CA, USA*

E-mail: fabien.paulot@noaa.gov

Abstract

Hydrogen (H_2) is experiencing renewed interest throughout the world as a low carbon alternative to fossil fuels. Significant uncertainties remain regarding the environmental impact of increasing H_2 usage, in part due to gaps in our understanding of the H_2 atmospheric budget, including the H_2 release from industrial activities and the

H_2 soil removal, the most important sink of H_2. This study focuses on H_2 and carbon monoxide (CO) dry air mole fractions measured by the NOAA Global Monitoring Laboratory in discrete ambient air samples collected at sites located in the contiguous United States between 2010 and 2022. We take advantage of the high spatial and temporal coverage of this network to estimate the regional distribution of H_2 and CO sources using the potential source contribution function (PSCF). Although H_2 PSCF is consistent with a large anthropogenic source of H_2 from fossil fuel combustion, episodic air masses with high H_2 dry air mole fractions (> 700 ppb) recorded at some sampling locations in the Western and Southeastern US are not associated with elevated CO, suggesting significant non-combustion anthropogenic or natural sources of H_2. Air masses depleted in H_2 are recorded in rural or remote continental sites. Observations collected in Colorado show that the degree of H_2 depletion not only reflects differences in soil exposure but also the inhibition of the soil sink under arid conditions.

1 Introduction

Hydrogen (H_2) is one of the most abundant trace gases in the atmosphere, with present-day background concentrations of $\simeq 550$ ppb.[1] The most important source of H_2 is the photolysis of formaldehyde produced by the oxidation of methane (CH_4).[2,3] Other important sources include fossil fuel combustion,[4] marine[5] and terrestrial nitrogen fixation,[6] biomass burning,[7] and the oxidation of non-methane volatile organic carbon.[2,3,8] We recently estimated that the present-day source of atmospheric H_2 from these processes is $\simeq 74$ Tg yr^{-1} (similar to CH_4 on a molar basis),[3] consistent with previous estimates.[2] Atmospheric sinks are dominated by H_2 scavenging by high-affinity hydrogen oxidizing bacteria (HA-HOB),[9] a class of bacteria that has been identified in many terrestrial ecosystems.[10,11] Oxidation of H_2 by the hydroxyl radical accounts for $\simeq 20 - 30\%$ of the overall H_2 sink.[2,3,12] The lifetime of H_2 in the troposphere is $\simeq 2$ years.[2]

Although H_2 is not a greenhouse gas, its atmospheric oxidation by the hydroxyl radical

(OH) leads to increases in the concentration of greenhouse gases, including CH_4, stratospheric water, and ozone (O_3).[12,13] Recent studies have assessed a Global Warming Potential (100 yr-horizon) of $\simeq$12 for H_2.[14–16] The environmental impact of H_2 emissions is dampened by H_2 soil sink. This process remains poorly understood, with global estimates of the H_2 soil sink and its response to environmental changes carrying large uncertainties.[3,17] Other important uncertainties in the present-day budget of H_2 include H_2 seepage from below-ground geological reservoirs[18] with estimates ranging from 0 to 31 Tg yr^{-1}.[19] H_2 may also be released from industrial facilities that produce or consume H_2;[20,21] the global demand for H_2 was $\simeq$ 97 Tg yr^{-1} in 2023, mainly from refineries and fertilizer production.[22] H_2 demand is projected to increase by 60% between 2020 and 2030[23] driven by new applications for H_2 in hard-to-electrify sectors (e.g., high-temperature heating in industry[24–26]). H_2 release from these activities is uncertain, with indirect estimates ranging from 0 to 20%.[20] Westra et al. (2024)[21] recently reported the first measurement-based estimates of H_2 release from an industrial chemical complex and estimated a leakage rate of 0.2–5%.

A better understanding of H_2 release from the industrial sector is critical to assess the implications of increasing H_2 usage.[14,27,28]

This highlights the need to understand the processes that control both soil removal of H_2 and its release by anthropogenic activities, to assess the potential impacts of increased H_2 use. This study focuses on regional sources and sinks of H_2 in the contiguous United States. Here, we leverage H_2 observations from the National Oceanic and Atmosphere Administration (NOAA) Global Monitoring Laboratory (GML) at a network of sampling sites. This network provides a denser set of H_2 atmospheric mole fraction measurements with higher frequency and locations closer to expected H_2 sources than the NOAA Global Cooperative Air Sampling Network[1,3] We first describe the observations and their implications for regional sources and sinks. Next, we examine the drivers of extreme H_2 values and their relationships to potential point or localized sources and soil removal.

2 Materials and Methods

2.1 NOAA measurements of discrete air samples from US sites

For over twenty years, the NOAA Global Monitoring Laboratory in situ measurement program has conducted high-frequency, high-precision, measurements of carbon dioxide (CO_2) and carbon monoxide (CO) from sites in the continental US to support the quantification of regional scale CO_2 fluxes with inverse models.[29] In the late 2000s, a standardized and automated air sampling system, called a programmable flask package (PFP), was added for comparison and quality control of the in situ measurements, as well as for the benefit of measuring the abundance of additional species. Each PFP contains a manifold connected to 12 0.7L glass flasks. After collection of air samples, the PFPs are sent to NOAA GML in Boulder, CO for analysis of over fifty trace gases species, including CO and H_2. These analysis are performed on the same analytical systems as air samples collected for the NOAA Global Cooperative Air Sampling Network.[1] At most sites, discrete PFP air samples are collected every 1–2 days, in the middle of the day, when the boundary layer is well mixed. Table S1 gives a list of the NOAA PFP sampling sites used in this study. At the in situ sites, a PFP and a compressor package are housed in a building or shelter. The compressor package pulls air from the top of the tower structure or a short mast (i.e at NWR site) using an inlet line. The sample collection sequence starts with the flushing of the PFP manifold and flask (for 12 to 15min) followed by a less than 1-minute flask fill. There have been a few iterations of the air drying and conditioning of the glass flasks to reduce small biases in CO_2 for the pressurized (2–2.3 bar) air samples, mostly for high humidity conditions. More details regarding PFP sampling can be found in Andrews et al. (2014).[29]

Since late 2009, the NOAA GML discrete air sample analysis for H_2 uses gas chromatography and a helium pulse discharge detector (GC-HePDD). Novelli et al. [2009][30] described the implementation and showed the instrument response is linear over the 0-2000 ppb mole fraction range. NOAA recently adopted the WMO/MPI X2009 calibration scale for H_2 and

reprocessed all GC-HePDD measurements on that scale.[1] Over time, different analytical methods have been used for CO: optical analyzer (Aerodyne TILDAS CO/N_2O, August 2019-present), vacuum ultraviolet resonance fluorescence (VURF) spectroscopy (2007-July 2019)[31] and chromatography/HgO reduction detection.[32] The CO dry air mole fraction measurements are reported on the WMO/NOAA X2014A calibration scale. The NOAA GML trace gas measurement systems are optimized to have high sensitivity and high precision for background air concentrations. Measurement results for air samples with over 2 ppm for H_2 and over 1.5 ppm for CO have a rejection quality control flag and are not used in this study. The measurement uncertainties for H_2 and CO in PFPs are less than 3 ppb.

Fig. 1a shows the location of the different sites that are part of the NOAA PFP sampling network.[33] We restrict our analysis to valid observations collected between 10am and 4pm local time. Fig. 2 shows the 2010-2022 daytime PFP H_2 measurement time series for the different sites. For each site, we first exclude outliers (denoted by circles), defined as observations which fall outside ± 3 standard deviations from the time series mean (Table S1). The data are then fitted with a function consisting of a quadratic polynomial, a series of 4 yearly harmonics and residuals to that function are filtered to further capture short-term and long-term variations following Thoning et al.(1989).[34] We focus here on H_2 and CO anomalies, i.e., the difference between the observation and the value estimated by the fitted function. These anomalies are denoted ΔH_2 and ΔCO hereafter.

2.2 Back trajectory and potential source contribution function calculations

We generate 10-day back trajectories for each observation using the NOAA HySplit model[35] driven by the North American Mesoscale Forecast System (NAM) meteorology at 12km horizontal resolution.[36] For each observation, back trajectories are initialized at 25, 50, and 75% of the planetary boundary layer (PBL) height from NAM and released at the site location, shifted by ± 1 grid cell in N/S and E/W directions. The position of the trajectory is recorded

every 15 minutes. We use the pysplit[37] and pysplitprocessor[38] packages to streamline the generation of the trajectories.

To identify potential surface sources and sinks at the continental level for compound X, we calculate the potential source contribution function (PSCF)[39] defined as:

$$PSCF_{i,j} = \frac{\sum_s \sum_{t \in t_s} n_s(i,j,t)}{\sum_s \sum_t n_s(i,j,t)} M_{i,j} \tag{1}$$

where $n_s(i,j,t)$ is the number of end-points for back-trajectories originating from site s at time t in the PBL for grid-cell(i,j). Note that in the numerator, t_s is taken as the subset of observations from the top decile in anomalies ΔX, where X is H_2 or CO. To exclude regions that are poorly sampled by the observing network, we require end-points from at least 730 distinct observations for a $PSCF_{i,j}$ to be valid, i.e., $M_{i,j} = 1$ if $\sum_s \sum_t [n_s(i,j,t) > 0] > 730$ and 0 otherwise. 30% of the grid cells in the region of interest (lat=22-50°N, lon=65–125°W) have end-points associated with fewer than 730 distinct observations.

2.3 Soil removal estimate

H_2 consumption by high-affinity H_2-oxidizing bacteria (HA-HOB) is the most important removal mechanism for atmospheric H_2[40]. The deposition velocity of H_2 ($v_d(H_2)$) for a particular location and time can be estimated as:

$$\frac{1}{v_d(\mathrm{H_2})} = \frac{1}{g_i} + \frac{1}{g_s} \tag{2}$$

where g_s and g_i are the H_2 conductance in the soil and the H_2 conductance through barriers that reduce the transport of H_2 to HA-HOBs, respectively.

g_s can be expressed as:[41]

$$g_s = \sqrt{k_m \, f_T \, f_M \, D_s} \tag{3}$$

where f_T and f_M are the sensitivity of H_2 biological uptake to soil temperature and soil

moisture, respectively, D_s is the moisture-dependent diffusivity of H_2 in the soil, and k_m represents the maximum biological uptake rate of H_2. The spatial variability of k_m is not well understood, with recent studies highlighting a possible modulation with soil carbon.[2,42,43] Here, we assume that k_m is spatially uniform and chosen to yield a global mean $v_d(H_2)$ of 0.09 mm/s.[12]

f_T and f_M are parameterized following Ehhalt and Rohrer (2013) and Bertagni et al. (2021), respectively. HA-HOB activity is suppressed under very dry conditions.[45] This parameterization captures the non-linear and non-monotonic response of $v_d(H_2)$ to soil moisture, associated with the competing effects of diffusion, which decreases rapidly with soil moisture, and biological activity, which increases with soil moisture (up to an optimum[46,47])

We account for the diffusion barrier through snow and canopy as:

$$g_i = \frac{1}{\frac{1}{g_{snow}} + \frac{1}{g_{canopy}}} \tag{4}$$

where g_{snow} is calculated from snow depth assuming a snow porosity of 0.645[48] and g_{canopy} is a time-invariant climatology.[49]

All calculations are performed using hourly ERA5 reanalysis fields for snow depth, atmospheric pressure, and soil moisture and temperature in the top 7cm.[50]

2.4 Soil interaction

For a given observation at time t_m, we define the effective interaction of the observed air parcel with the soil surface over a period [t_m-τ, t_m] as:

$$f_s(t_m, \tau) = \frac{1}{N\,\tau} \sum_{i=1}^{N} \int_{t_m-\tau}^{t_m} \delta_{pbl}(i,t)dt \tag{5}$$

where $\delta_{pbl}(i,t)$ is 1 if the i^{th} back-trajectory is in the PBL over land at time t and 0 otherwise, and N is the total number of back trajectories. N=27 as discussed above.

Similarly, we define the effective H_2 soil diffusion (D_s^{eff}) and the effective deposition velocity (v_d^{eff}) over a time T as:

$$D_s^{eff}(t_m, \tau) = \frac{1}{N\,\tau} \sum_{i=1}^{N} \int_{t_m-\tau}^{t_m} D_s(i,t)\delta_{pbl}(i,t)dt \tag{6}$$

$$v_d^{eff}(t_m, \tau) = \frac{1}{N\,\tau} \sum_{i=1}^{N} \int_{t_m-\tau}^{t_m} v_d(i,t)\delta_{pbl}(i,t)dt \tag{7}$$

$D_s(i,t)$ and $v_d(i,t)$ are the H_2 soil diffusion and H_2 deposition velocity sampled at the time and location closest to the i^{th} back trajectory at time t.

2.5 Upstream H_2

The upstream concentration of H_2 for an observation collected at time t_m is defined as:

$$\mathrm{H_2}(t_m, \tau)[upstream] = \frac{1}{N} \sum_{i=1}^{N} {\mathrm{H_2}}^{model}(i, t_m - \tau) \tag{8}$$

where ${\mathrm{H_2}}^{model}(i,t)$ is the dry air mass fraction of H_2 sampled at the location of the back trajectory i at time t. Here, ${\mathrm{H_2}}^{\mathrm{model}}$ is taken from a GFDL-AM4.1 monthly climatology (2010–2014 average)[3]. We calculate the anomaly ${\Delta\mathrm{H_2}}^{\mathrm{upstream}}$ as we did for the observations (ΔH_2).

Similarly, we estimate the origin of the air mass sampled at time t_m as:

$$\mathrm{lat}(t_m, \tau)[upstream] = \frac{1}{N} \sum_{i=1}^{N} \mathrm{lat}^{model}(i, t_m - \tau) \tag{9}$$

where $\mathrm{lat}^{model}(i,t)$ is the latitude of the back trajectory i at time t.

2.6 Influence of point sources

We define the potential influence of a surface location p for an observation at time t_m as:

$$I_p(t_m, \tau, d) = \exists i \in \{1..N\}(d_p(i, \tau) < d \rightarrow 1) \quad (10)$$

where $d_p(i, \tau)$ is the minimum distance between end points associated with the back trajectory i and the location p over the interval $[t_m - \tau, t_m]$, and d is a distance. Thus, $I_p(t_m, \tau, d)$ is 1 if any back trajectory originating at time t_m travels within a distance d of the site p between t_m and $t_m - \tau$. Note that we only consider endpoints in the PBL to calculate $d_p(i, \tau)$.

3 Results and discussion

3.1 H_2 observations and regional sources and sinks

Fig. 1b shows the distribution of H_2 for daytime observations at each site for the time periods in Table S1. Median dry air mole fractions range from 511 to 540 ppb. H_2 levels in the Northeast and upper Midwest are on average 10–20ppb lower than in South and Pacific regions. This latitudinal gradient is a salient feature of atmospheric H_2 [1–3,51] and is attributed to the soil sink of H_2. [2,51]

The observations show strong day-to-day variability in H_2 on top of a marked seasonal cycle at all sites (Fig. 2). All three sites located in California occasionally show high levels of H_2, above 700 ppb: Sutro in San Francisco (STR), Walnut Grove about 60 miles NE of San Francisco (WGC), and Mount Wilson in the San Gabriel Mountains near Pasadena (MWO). High levels of H_2 are also observed, although less frequently, at Beach Island in South Carolina (SCT). In contrast, the Boulder Atmospheric Observatory (BAO) north of Denver, Colorado, has the most samples with depleted H_2 compared to the temperate Northern hemisphere marine boundary layer H_2. [1] Large negative anomalies are also observed at continental sites: Park Falls in Wisconsin (LEF), West Branch in Iowa (WBI), and the Southern Great Plains observatory in Oklahoma (SGP).

Fig. 3a shows the PSCF probability at a resolution of 1°x1° calculated using the 90^{th}

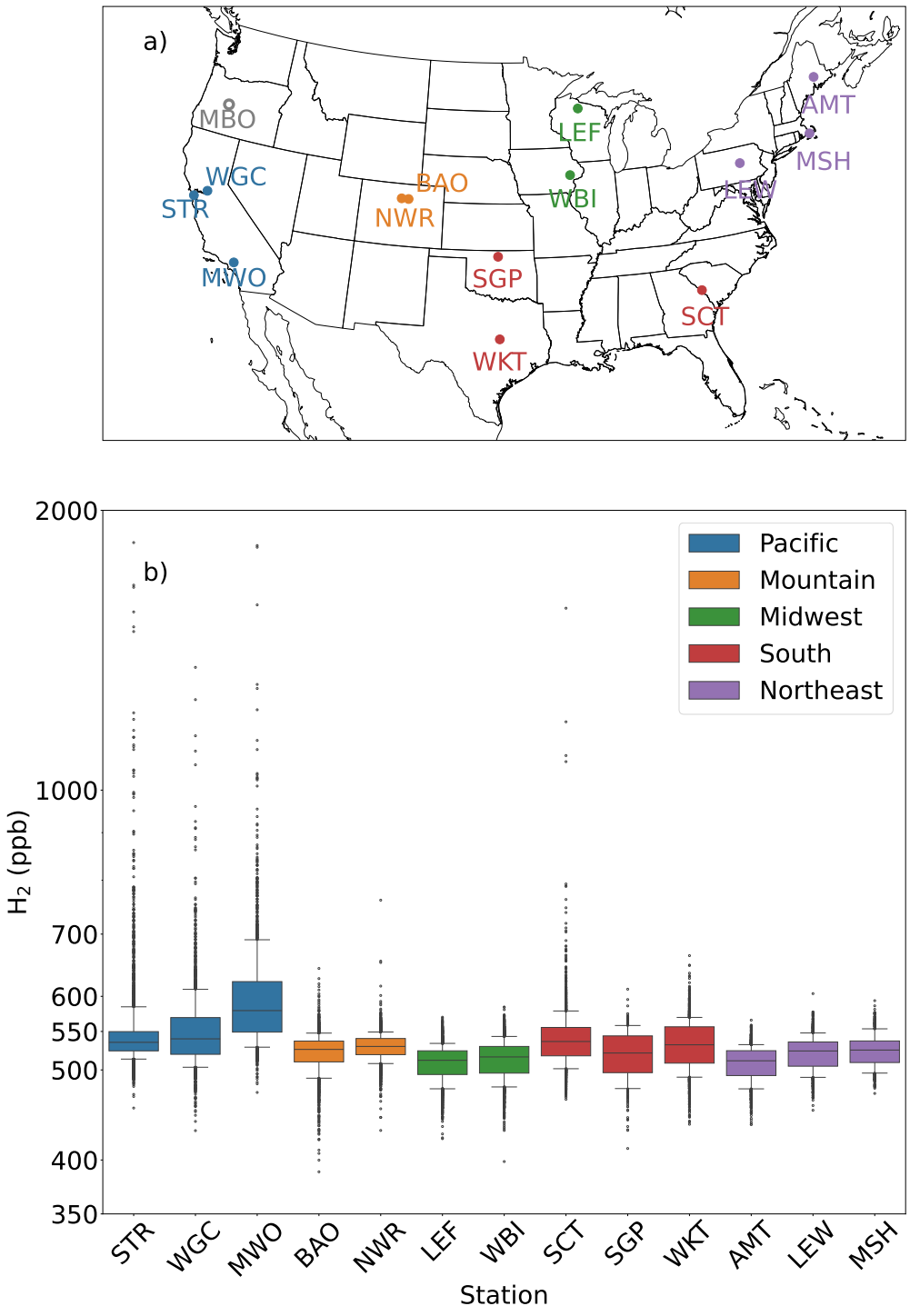


Figure 1: Location of PFP sampling sites used in this study (a) and H_2 distribution at each site (b). The box shows the quartiles of the dataset, while the whiskers extend to show the 10^{th} and 90^{th} percentiles. Dots indicate observations that fall outside this range. Observations collected at the MBO sampling site are not used in this study.

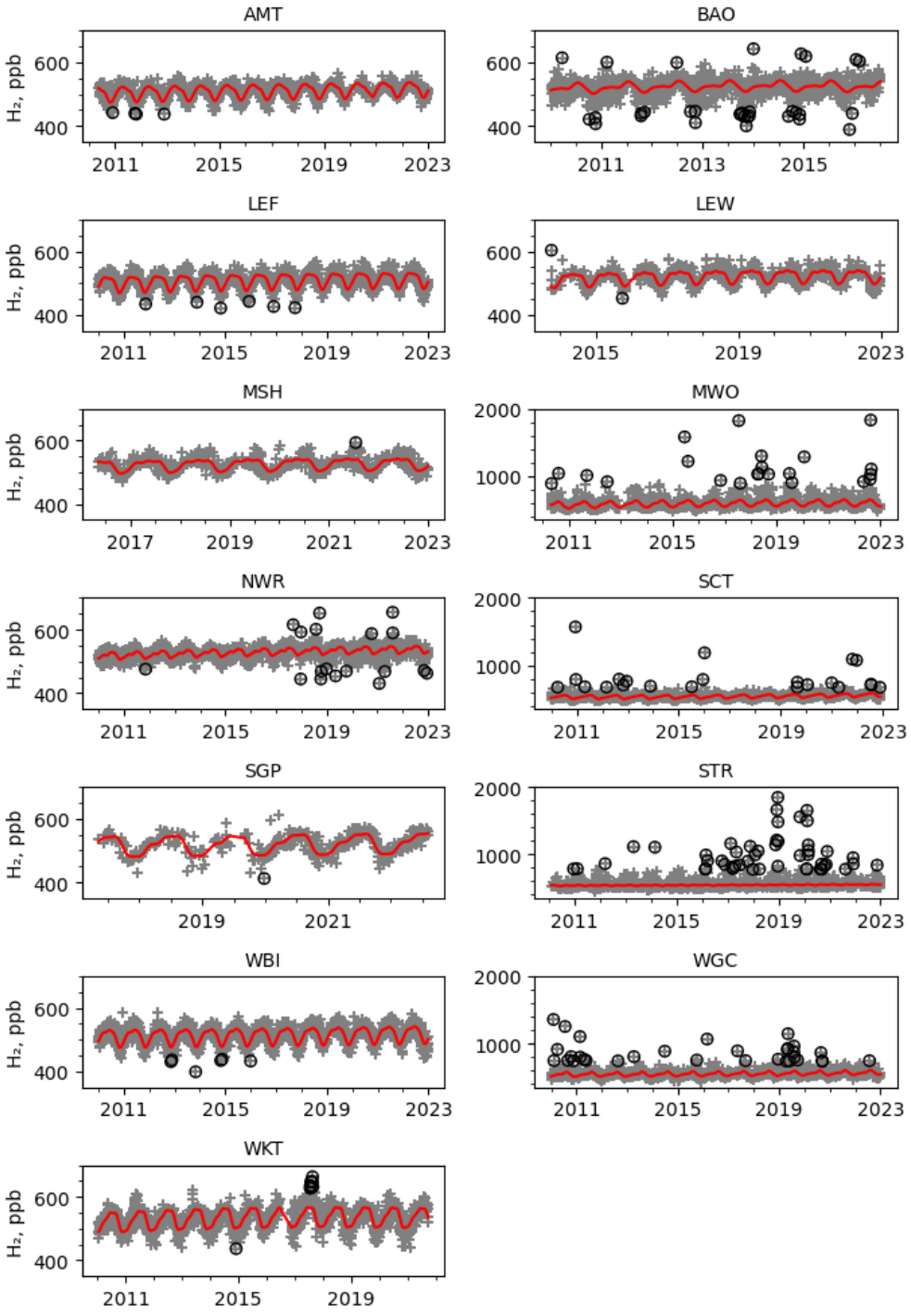


Figure 2: H_2 timeseries at PFP sampling sites. Black circles denote outliers (as defined in section 2.1). The red line denotes the functional fit to filtered observations.[34]

percentile of ΔH_2 at each site. A PSCF probability greater than 10% indicates regions that are more likely to have contributed to the 10% largest ΔH_2. PSCF is high over much of the Southern US and Midwest as well as in the Midatlantic and on the West Coast. This distribution matches well the estimated spatial distribution of H_2 anthropogenic sources[3] (Fig. 3b), which are dominated by emissions from the transportation sector ($> 80\%$). CO PSCF exhibits enhancements in many of the same regions as H_2 (Fig. S1). This is consistent with the well-documented correlation between H_2 and CO concentrations in vehicular exhaust due to water gas shift reaction.[4,52] Note that the PSCF for H_2 is especially large along the Louisiana and Texas Gulf Coast, where many H_2 production facilities (mostly associated with petroleum refineries) are located (circles in Fig. 3b[53]). This regional enhancement highlights the need to better characterize H_2 release from these facilities.

Fig. 3a also shows that the PSCF probability is low over the Northern part of the Central US (Montana, Nebraska, South and North Dakota). This pattern is not observed for CO (Fig. S1) and may reflect differences in the dominant sink of H_2 (soil) and CO (oxidation by OH). Low PSCF for H_2 in the central US is consistent with the predicted maximum deposition velocity for H_2 (Fig. 3c). Note that our model for $v_d(H_2)$ implies a very weak sink in much of the Western US, as HA-HOB activity is inhibited by low soil moisture (Fig. S2).

3.2 Evidence for large non-combustion sources of H_2

Figs 1b and 2 show frequent elevated H_2 observations at the STR, WGC, MWO, and SCT sites. At the 3 California sites and the South Carolina site, ΔH_2 is very high (defined here as $\Delta H_2 > 150$ ppb) for 3.1, 2.4, 4.1%, and 0.6% of all observations, respectively. The maximum observations are 1846, 1355, 1832, and 1569 ppb, respectively. These mole fractions are comparable with the maximum H_2 of 1346 ppb recently reported for an active aircore sampling survey in the immediate vicinity of a large industrial complex in the Netherlands.[21]

Fig. 4 shows that a fraction of these extreme events are associated with anomalies in CO

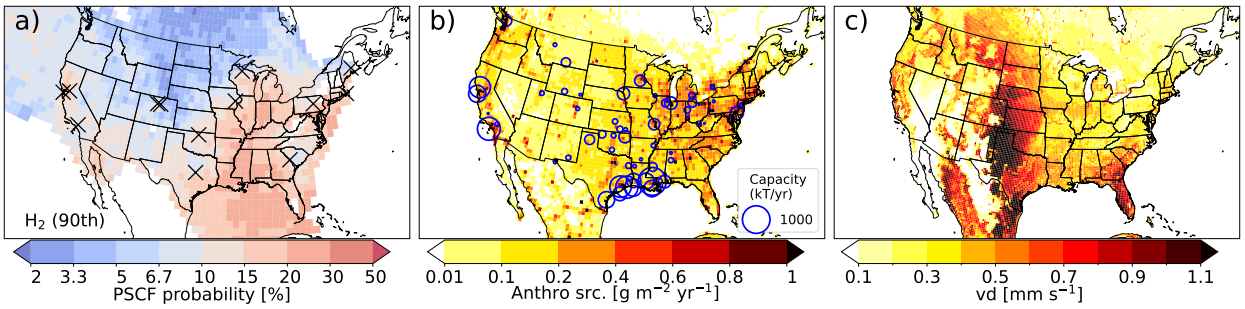


Figure 3: PSCF probability at 1°x1° resolution for H_2 based on the 90^{th} percentile of ΔH_2 at each site (a), anthropogenic H_2 emissions (2010–2019 average)[3] (b), and average H_2 deposition velocity (2010–2022 average) (c). Crosses in panel a) denote the locations of NOAA PFP sampling sites. H_2 production capacity aggregated on a 2°x2° grid is indicated by blue circles in panel b) (See Text S1). Total anthropogenic H_2 emissions in the contiguous US for the 2010-2019 period are $\simeq$ 1.3 Tg yr^{-1}, primarily from transportation (>80%).[3] US H_2 production is $\simeq$ 10 Tg yr^{-1}.[54]

at least as high as those in H_2 at STR (40%), WGC (18%), and MWO (30%). These events likely reflect the influence of nearby or upwind fires and relative enhancements in H_2 and CO generally fall within expectation from fire emissions[7,55](yellow shaded region). However, the most positive anomalies in H_2 are not associated with a concomitant enhancement in CO, suggesting that they are the result of a potentially large non-combustion source of H_2.

In the following, we focus on the two Northern California sites (STR and WGC) where 59 and 38 events between 2010 and 2022 have elevated H_2 without concurrent enhancement in CO, i.e., $\Delta H_2 > 150$ppb and $\Delta CO < \Delta H_2$, respectively. These observations will be referred to as non-combustion extreme events (NCEE), hereafter. At STR, NCEE are most frequent in fall (21) and winter (22), while there is no seasonal pattern at WGC. More than 5 NCEE are identified each year at STR from 2017 to 2020 and at WGC in 2010 and 2019. No NCEE is identified at WGC after 2019.

Fig. 5 shows regions with at least one back trajectory endpoint in the PBL for over 50% of NCEE at WGC (vertical hatches) and STR (horizontal hatches), respectively. We note that these regions intersect over the Bay Area, which suggests that NCEE may originate from one (or several) point sources located in this region. Fig. 5 shows that the hatched region encompasses five refineries (blue circles). Refineries use H_2 to process diesel fuel (e.g.

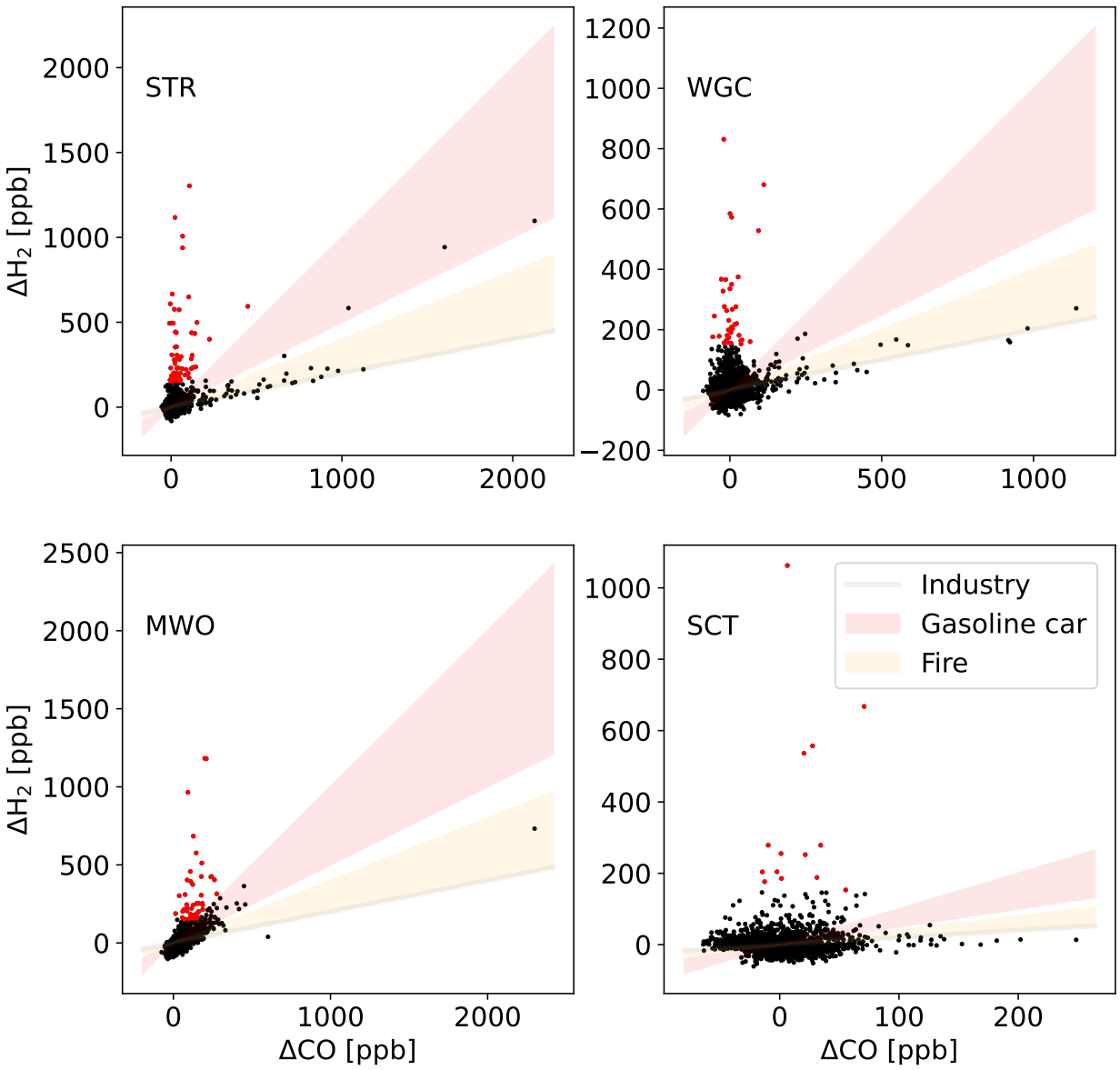


Figure 4: Relationship between CO and H_2 residuals at four sites. Shaded regions denote the ratio of H_2 to CO emissions associated with different anthropogenic and natural combustion sources.[3,7,55] Red dots indicate non-combustion extreme events (NCEE).

sulfur removal) and transform long and unsaturated products in feed (e.g. oil sand[56]) into products with lower molecular weight. Note that only $\simeq 1/3$ refineries use H_2 in the US.[57] The combined H_2 capacity for these five refineries was $\simeq$ 1000 kT yr^{-1} circa 2018, or $\simeq$ 10% of US H_2 capacity.[53] H_2 may be released during refining operations due to imbalance between H_2 production and consumption, as recently documented at one of the Bay Area refineries.[58]

We find that 83% and 74% of NCEE may have been influenced by refinery emissions, as defined by $I_p(\tau = 3 \text{ days}, d = 10 \text{ km}) = 1$ for at least one of the five refineries shown in Fig. 5, at STR and WGC, respectively. In contrast, 25% and 45% of all other observations at STR and WGC, respectively, may have been influenced by refineries' activities. Since this represents a much larger number of events (Table S1), we find that only 6 and 3% of air masses that may have been influenced by refineries are associated with NCEE, respectively. This suggests that if NCEE are associated with refineries' activities, the release of H_2 is infrequent. We note that these results are largely insensitive to the choice of the distance d used to determine the point source potential influence (Table S2). As MWO and SCT are also located in the vicinity of refineries and fertilizer and chlorine plants, respectively, some NCEE at these sites may also be associated with the release of H_2 from these operations.

More frequent observations would be necessary to further identify the origin and estimate the magnitude of the likely sporadic H_2 sources that are responsible for NCEE. In the Bay Area, it is noteworthy that at least 10% of NCEE measured at STR and WGC do not seem to be influenced by the five Bay area refineries. As shown in Fig. 5, there are other possible non-combustion sources of H_2 in the vicinity of STR and WGC. These include anthropogenic activities, e.g., H_2 fueling stations (red dots) and a liquid H_2 plant SE of Sacramento (purple circle). Other potential point sources may include geological H_2 seepage. Very high H_2 air mole fractions have been reported in soil gas at Burro Mountains[59] ($>$ 70% at 50cm depth, Central California) and near Palo Alto[60] ($\simeq$ 20ppm at 1m depth,). H_2 ($>$ 1%) has also been reported in exsolving gases at ultra-basic springs (The Cedars[61]) and at The

Geysers geological reservoir[62] (triangles, Fig. 5). At these sites, H_2 is likely associated with serpentinization.[61] As serpentinite occurs throughout northern California[63,64] (Fig. 5) some NCEE may be associated with geological H_2 seepage.

Our study highlights how the analysis of the NOAA PFP-based data can be used to guide more intensive measurement campaigns focused on certain sources of H_2, including geological prospecting for H_2.

3.3 Evidence for moisture inhibition of the H_2 soil removal

As noted above, the NOAA PFP-based measurements show large negative H_2 anomalies at some continental sites. Here, we focus on the BAO site[66] located 35 km north of Denver in NE Colorado. This site is close to large spatial gradients in $v_d(H_2)$ (Fig. 3c) associated with the large contrast in soil moisture between the arid Western US and the wetter central US. In particular, the activity of HA-HOB may be inhibited by low soil moisture west of BAO in much of the Great basin (Fig. S2). This makes BAO especially well-suited to assess the impact of variation in the H_2 soil sink on H_2 air mole fraction.

Fig. 6a shows that the upper and lower deciles of ΔH_2 at BAO differ by 50 to 100 ppb depending on the season. The lowest H_2 air mole fraction (388 ppb) for the PFP samples used in this study is measured at BAO and 3.5% of the BAO observations have negative anomalies of at least 50 ppb. The H_2 anomalies are most (least) negative in the fall (summer).

Fig. 6b shows that there is no monotonic relationship between ΔCO and ΔH_2 for the BAO PFP measurements. Since CO and H_2 share many sources (e.g., combustion, photochemistry[2]), this suggests that H_2 sources are unlikely to be the primary driver of variation in ΔH_2 at BAO.

In the following, we focus on the history of air masses in the 4 days preceding their observation at BAO (τ=4 days). Fig. S3a shows that air masses associated with the most negative ΔH_2 tend to originate further North than air masses with the most positive ΔH_2. In the Northern Hemisphere, the marine boundary layer H_2 mole fraction decreases with

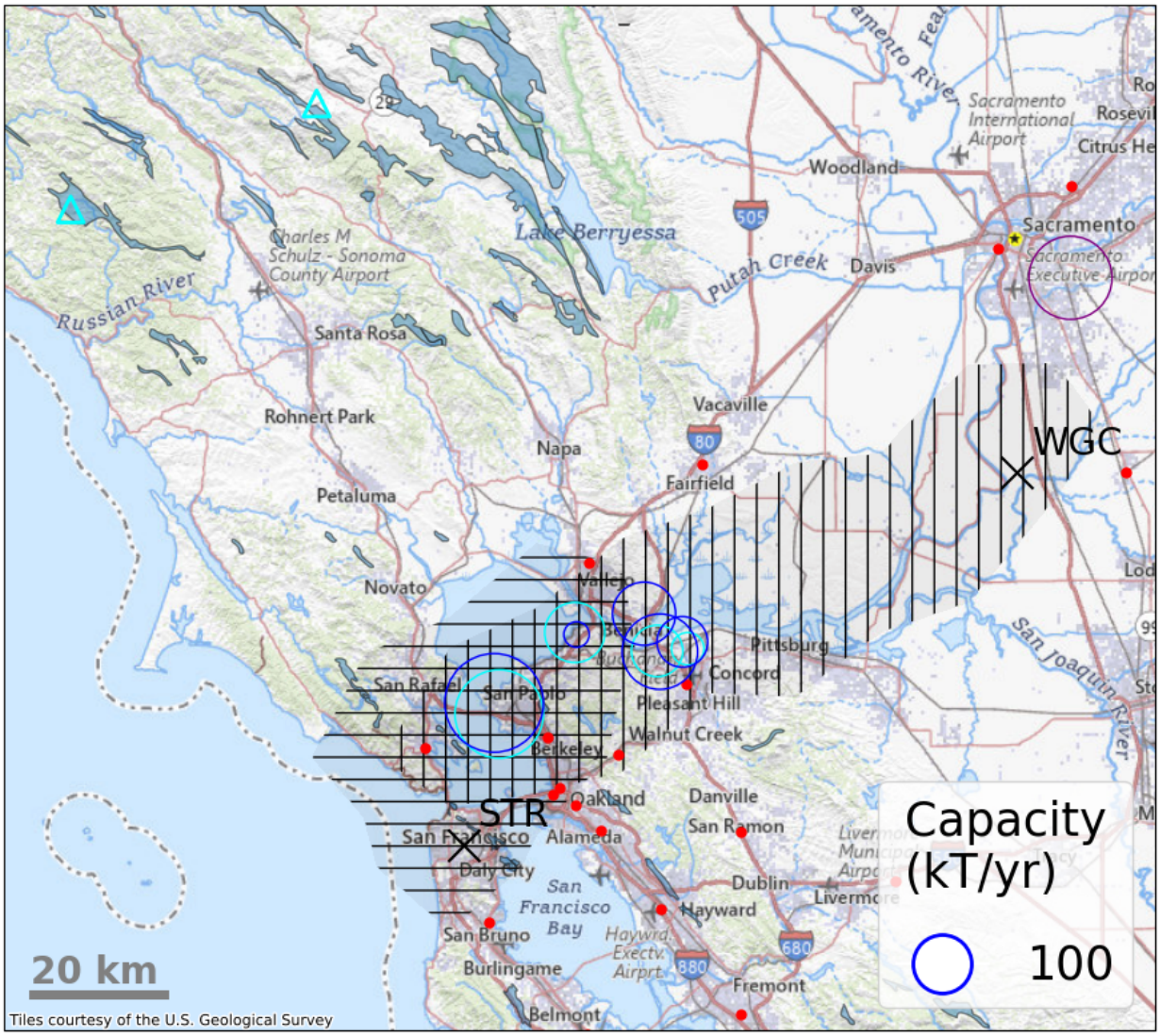


Figure 5: Hatches denote regions where trajectories associated with over 50% of NCEE have passed in the PBL for WGC (vertical hatches) and STR (horizontal hatches). H_2 capacity for refineries (2020), merchant gas and merchant liquid (x100)[53] are indicated by blue, cyan, and purple circles, respectively. Red dots denote the locations of H_2 fueling stations.[65] Locations where elevated geologic H_2 outgassing has been recorded are denoted by triangles. Blue-shaded regions denote serpentinite units.[64]

latitude.[1] We estimate that the difference in air mass origin can contribute to $\simeq 20\%$ of the observed difference in ΔH_2 at BAO (Fig. 6c).

The previous discussion suggests that neither H_2 sources nor the upstream concentration of H_2 are likely to cause extreme H_2 anomalies. In the following, we focus on the processes that govern the H_2 soil sink. Fig. 6d shows that the integrated air mass soil exposure ($f_s(\tau = 4$ days); eqn. 5) is anticorrelated with ΔH_2. On average, we find that air masses with the most negative H_2 have spent 2 to 3 times as much time in the PBL over land as air masses associated with the most positive ΔH_2 and are therefore more likely to be affected by the soil sink.

From eqn. (3), the strength of the soil sink depends on both biotic (represented by the f_T, f_s, and k_m factors) and abiotic (D_s) processes. Globally, diffusion is expected to be the most limiting factor for the uptake of H_2 by soils in most regions.[44] Surprisingly, Fig. 6e shows that the effective soil diffusion, D_s^{eff} (eqn: 6), varies little with ΔH_2 except in winter, i.e., the longer time spent in the PBL (f_s) for the most negative ΔH_2 is partly counterbalanced by reduced diffusion (i.e., wetter conditions) in soil (Fig. S3b). This suggests that the H_2 soil sink in spring, summer and fall is not limited by diffusion in this region but by biological activity. This is supported by Fig. S3c, which shows that the most negative ΔH_2 at BAO are $\simeq 30-60\%$ less likely to have interacted with soil that is too dry to support H_2 uptake.[44]

Fig. 6f shows that $v_d(H_2)$, which accounts for both biotic and abiotic modulation of H_2 sink strength, is negatively correlated with ΔH_2. On average, the effective $v_d(H_2)$ is 3–4$\times$ faster for air masses with the most negative ΔH_2 than for those with the most positive ΔH_2. Over the course of 4 days and assuming an effective deposition height of 500m, such difference in $v_d(H_2)$ could result in 30–90ppb difference in ΔH_2 between these air masses. This supports the importance of the soil sink in modulating ΔH_2 at BAO. This also highlights how long-term monitoring of surface H_2 in semi-arid regions can provide important constraints on models of $v_d(H_2)$[3,44] including on the impact of hydrological changes,[67] an important uncertainty in the characterization of the historical and future H_2 global budget.[12,16]

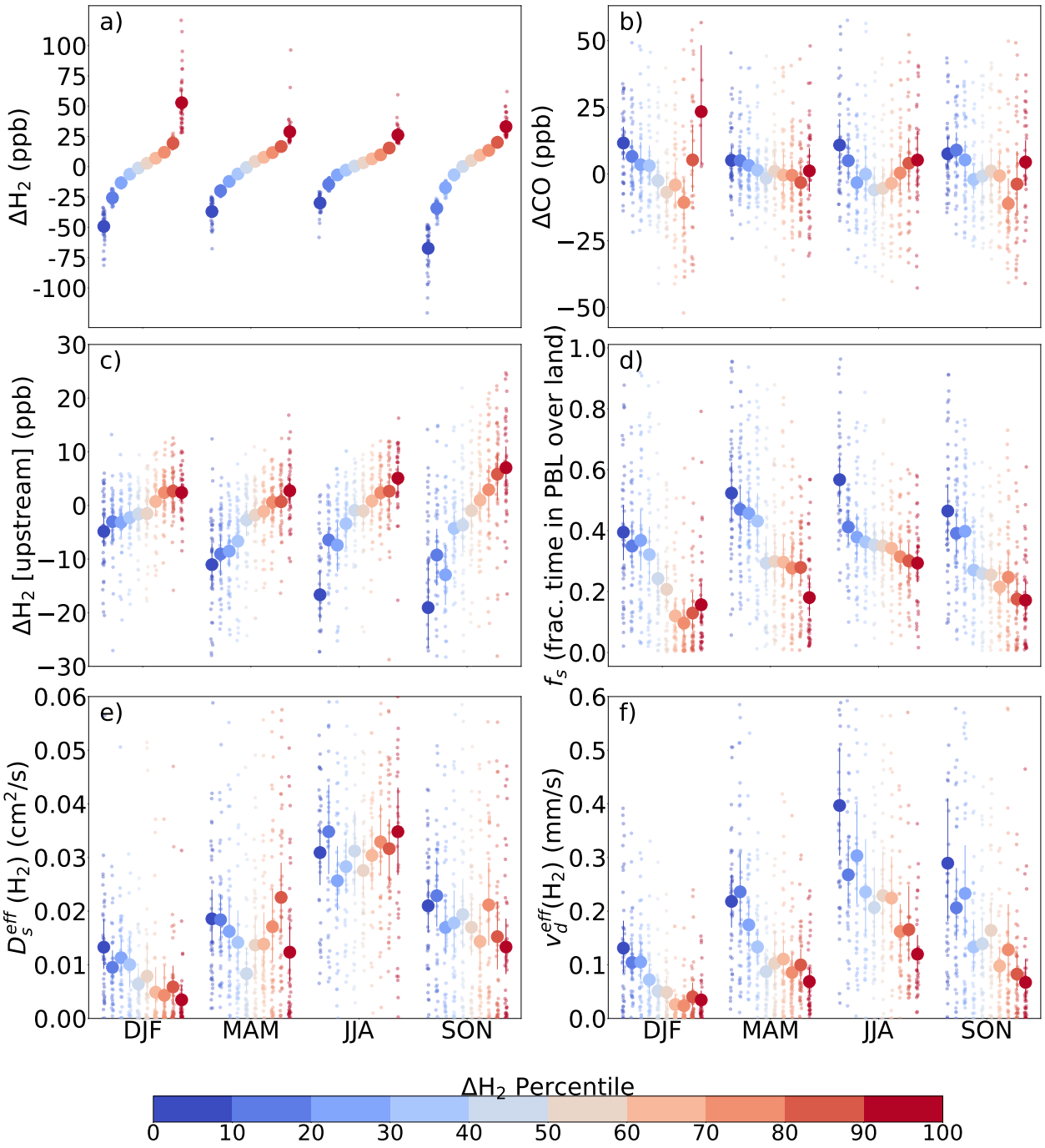


Figure 6: H_2 anomaly at BAO broken down by decile and season (a). Panels b) and c) show the associated CO anomaly and upstream H_2 anomalies. Panels d)-f) show the effective time, effective diffusion and effective soil deposition for each decile (τ=4 days). For each panel, small and large dots denote individual events and the average over each decile, respectively. Error bars denote the 99% confidence interval for the mean.

Acknowledgement

This research has been supported in part by NOAA cooperative agreements NA17OAR4320101 and NA22OAR4320151 and by the US Department of Energy, Office of Energy Efficiency and Renewable Energy (EERE), specifically the Hydrogen and Fuel Cell Technologies Office; Office of Biological and Environmental Research (BER) under contract no. DE-AC02-05CH11231 as part of the Atmospheric Radiation Measurement (ARM) Program. The authors acknowledge Earth Networks, Inc. for assistance with flask sampling at LEW and NIST for supporting the measurements at MSH. The views expressed herein do not necessarily represent the views of NOAA, the U.S. Department of Energy, or the United States Government.

Supporting Information Available

Additional figures including CO PSCF, H_2 uptake inhibition by soil moisture, and more analysis of BAO ΔH_2. A table showing the estimated influence of refineries on NCEE for different values of d. Description of the sources used to estimate H_2 capacity per facility.

References

(1) Pétron, G. et al. Atmospheric H_2 observations from the NOAA Cooperative Global Air Sampling Network. *Atmospheric Measurement Techniques* **2024**, *17*, 4803–4823.

(2) Ehhalt, D. H.; Rohrer, F. The tropospheric cycle of H_2: a critical review. *Tellus B: Chemical and Physical Meteorology* **2009**, *61*, 500–535.

(3) Paulot, F.; Pétron, G.; Crotwell, A. M.; Bertagni, M. B. Reanalysis of NOAA H_2 observations: implications for the H_2 budget. *Atmospheric Chemistry and Physics* **2024**, *24*, 4217–4229.

(4) Vollmer, M. K.; Juergens, N.; Steinbacher, M.; Reimann, S.; Weilenmann, M.; Buchmann, B. Road vehicle emissions of molecular hydrogen (H2) from a tunnel study. *Atmospheric Environment* **2007**, *41*, 8355–8369.

(5) Walter, S.; Kock, A.; Steinhoff, T.; Fiedler, B.; Fietzek, P.; Kaiser, J.; Krol, M.; Popa, M. E.; Chen, Q.; Tanhua, T.; Röckmann, T. Isotopic evidence for biogenic molecular hydrogen production in the Atlantic Ocean. *Biogeosciences* **2016**, *13*, 323–340.

(6) Conrad, R.; Seiler, W. Contribution of hydrogen production by biological nitrogen fixation to the global hydrogen budget. *Journal of Geophysical Research* **1980**, *85*, 5493.

(7) Andreae, M. O. Emission of trace gases and aerosols from biomass burning – an updated assessment. *Atmospheric Chemistry and Physics* **2019**, *19*, 8523–8546.

(8) Derwent, R. G.; Jenkin, M. E. Estimation of the atmospheric hydrogen source from the oxidation of man-made and natural non-methane organic compounds using a Master Chemical Mechanism. *Atmospheric Environment* **2024**, *339*, 120871.

(9) Constant, P.; Chowdhury, S. P.; Hesse, L.; Conrad, R. Co-localization of atmospheric H_2 oxidation activity and high affinity H_2-oxidizing bacteria in non-axenic soil and sterile

soil amended with Streptomyces sp. PCB7. *Soil Biology and Biochemistry* **2011**, *43*, 1888–1893.

(10) Greening, C.; Constant, P.; Hards, K.; Morales, S. E.; Oakeshott, J. G.; Russell, R. J.; Taylor, M. C.; Berney, M.; Conrad, R.; Cook, G. M. Atmospheric Hydrogen Scavenging: from Enzymes to Ecosystems. *Applied and Environmental Microbiology* **2014**, *81*, 1190–1199.

(11) Greening, C.; Biswas, A.; Carere, C. R.; Jackson, C. J.; Taylor, M. C.; Stott, M. B.; Cook, G. M.; Morales, S. E. Genomic and metagenomic surveys of hydrogenase distribution indicate H_2 is a widely utilised energy source for microbial growth and survival. *The ISME Journal* **2015**, *10*, 761–777.

(12) Paulot, F.; Paynter, D.; Naik, V.; Malyshev, S.; Menzel, R.; Horowitz, L. W. Global modeling of hydrogen using GFDL-AM4.1: Sensitivity of soil removal and radiative forcing. *International Journal of Hydrogen Energy* **2021**, *46*, 13446–13460.

(13) Derwent, R. G.; Collins, W. J.; Johnson, C. E.; Stevenson, D. S. Transient Behaviour of Tropospheric Ozone Precursors in a Global 3-D CTM and Their Indirect Greenhouse Effects. *Climatic Change* **2001**, *49*, 463–487.

(14) Hauglustaine, D.; Paulot, F.; Collins, W.; Derwent, R.; Sand, M.; Boucher, O. Climate benefit of a future hydrogen economy. *Communications Earth & Environment* **2022**, *3*, 295.

(15) Sand, M.; Skeie, R. B.; Sandstad, M.; Krishnan, S.; Myhre, G.; Bryant, H.; Derwent, R.; Hauglustaine, D.; Paulot, F.; Prather, M.; Stevenson, D. A multi-model assessment of the Global Warming Potential of hydrogen. *Communications Earth & Environment* **2023**, *4*, 203.

(16) Warwick, N. J.; Archibald, A. T.; Griffiths, P. T.; Keeble, J.; O'Connor, F. M.;

Pyle, J. A.; Shine, K. P. Atmospheric composition and climate impacts of a future hydrogen economy. *Atmospheric Chemistry and Physics* **2023**, *23*, 13451–13467.

(17) Yashiro, H.; Sudo, K.; Yonemura, S.; Takigawa, M. The impact of soil uptake on the global distribution of molecular hydrogen: chemical transport model simulation. *Atmospheric Chemistry and Physics* **2011**, *11*, 6701–6719.

(18) Ellis, R. A.; Jacob, D. J.; Sulprizio, M. P.; Zhang, L.; Holmes, C. D.; Schichtel, B. A.; Blett, T.; Porter, E.; Pardo, L. H.; Lynch, J. A. Present and future nitrogen deposition to national parks in the United States: critical load exceedances. *Atmos. Chem. Phys.* **2013**, *13*, 9083–9095.

(19) Zgonnik, V. The occurrence and geoscience of natural hydrogen: A comprehensive review. *Earth-Science Reviews* **2020**, *203*, 103140.

(20) Esquivel-Elizondo, S.; Hormaza Mejia, A.; Sun, T.; Shrestha, E.; Hamburg, S. P.; Ocko, I. B. Wide range in estimates of hydrogen emissions from infrastructure. *Frontiers in Energy Research* **2023**, *11*.

(21) Westra, I. M.; Scheeren, H. A.; Stroo, F. T.; van Heuven, S. M. A. C.; Kers, B. A. M.; Peters, W.; Meijer, H. A. J. First detection of industrial hydrogen emissions using high precision mobile measurements in ambient air. *Scientific Reports* **2024**, *14*.

(22) International Energy Agency *Global Hydrogen Review 2024*; 2024.

(23) International Energy Agency *Global Hydrogen Review 2022*; 2022.

(24) Hydrogen Council Hydrogen scaling up. A sustainable pathway for the global energy transition. 2017; `https://hydrogencouncil.com/wp-content/uploads/2017/11/Hydrogen-Scaling-up_Hydrogen-Council_2017.compressed.pdf`.

(25) Abe, J.; Popoola, A.; Ajenifuja, E.; Popoola, O. Hydrogen energy, economy and storage:

Review and recommendation. *International Journal of Hydrogen Energy* **2019**, *44*, 15072–15086.

(26) International Energy Agency *Global Hydrogen Review 2023*; 2023.

(27) Ocko, I. B.; Hamburg, S. P. Climate consequences of hydrogen emissions. *Atmospheric Chemistry and Physics* **2022**, *22*, 9349–9368.

(28) Bertagni, M. B.; Pacala, S. W.; Paulot, F.; Porporato, A. Risk of the hydrogen economy for atmospheric methane. *Nature Communications* **2022**, *13*, 7706.

(29) Andrews, A. E. et al. CO_2, CO, and CH_4 measurements from tall towers in the NOAA Earth System Research Laboratory's Global Greenhouse Gas Reference Network: instrumentation, uncertainty analysis, and recommendations for future high-accuracy greenhouse gas monitoring efforts. *Atmospheric Measurement Techniques* **2014**, *7*, 647–687.

(30) Novelli, P. C.; Crotwell, A. M.; Hall, B. D. Application of Gas Chromatography with a Pulsed Discharge Helium Ionization Detector for Measurements of Molecular Hydrogen in the Atmosphere. *Environmental Science & Technology* **2009**, *43*, 2431–2436.

(31) Gerbig, C.; Kley, D.; Volz-Thomas, A.; Kent, J.; Dewey, K.; McKenna, D. S. Fast response resonance fluorescence CO measurements aboard the C-130: Instrument characterization and measurements made during North Atlantic Regional Experiment 1993. *Journal of Geophysical Research: Atmospheres* **1996**, *101*, 29229–29238.

(32) Novelli, P. C.; Masarie, K. A.; Lang, P. M.; Hall, B. D.; Myers, R. C.; Elkins, J. W. Reanalysis of tropospheric CO trends: Effects of the 1997–1998 wildfires. *Journal of Geophysical Research: Atmospheres* **2003**, *108*.

(33) Andrews, A. et al. *NOAA Global Greenhouse Gas Reference Network Flask-Air PFP*

Sample Measurements of H_2 at Tall Tower and other Continental Sites, 2010-2022 [Data set] - Version: 2025-03-03; NOAA GML, 2025.

(34) Thoning, K. W.; Tans, P. P.; Komhyr, W. D. Atmospheric carbon dioxide at Mauna Loa Observatory: 2. Analysis of the NOAA GMCC data, 1974–1985. *Journal of Geophysical Research: Atmospheres* **1989**, *94*, 8549–8565.

(35) Stein, A. F.; Draxler, R. R.; Rolph, G. D.; Stunder, B. J. B.; Cohen, M. D.; Ngan, F. NOAA's HYSPLIT Atmospheric Transport and Dispersion Modeling System. *Bulletin of the American Meteorological Society* **2015**, *96*, 2059–2077.

(36) NOAA Air ressources Laboratory - Gridded Meteorological Data Archives. 2024; `https://www.ready.noaa.gov/archives.php`.

(37) Warner, M. S. C. pysplit. 2019; `https://github.com/mscross/pysplit`.

(38) Blanchard, B. pysplitprocessor. 2019; `https://github.com/brendano257/pysplitprocessor`.

(39) Polissar, A. The aerosol at Barrow, Alaska: long-term trends and source locations. *Atmospheric Environment* **1999**, *33*, 2441–2458.

(40) Constant, P.; Chowdhury, S. P.; Pratscher, J.; Conrad, R. Streptomycetes contributing to atmospheric molecular hydrogen soil uptake are widespread and encode a putative high-affinity [NiFe]-hydrogenase. *Environmental Microbiology* **2010**, *12*, 821–829.

(41) Ehhalt, D.; Rohrer, F. Deposition velocity of H_2: a new algorithm for its dependence on soil moisture and temperature. *Tellus B: Chemical and Physical Meteorology* **2013**, *65*, 19904.

(42) Khdhiri, M.; Hesse, L.; Popa, M. E.; Quiza, L.; Lalonde, I.; Meredith, L. K.; Röckmann, T.; Constant, P. Soil carbon content and relative abundance of high affin-

ity H_2-oxidizing bacteria predict atmospheric H_2 soil uptake activity better than soil microbial community composition. *Soil Biology and Biochemistry* **2015**, *85*, 1–9.

(43) Bay, S. K.; Dong, X.; Bradley, J. A.; Leung, P. M.; Grinter, R.; Jirapanjawat, T.; Arndt, S. K.; Cook, P. L. M.; LaRowe, D. E.; Nauer, P. A.; Chiri, E.; Greening, C. Trace gas oxidizers are widespread and active members of soil microbial communities. *Nature Microbiology* **2021**, *6*, 246–256.

(44) Bertagni, M. B.; Paulot, F.; Porporato, A. Moisture Fluctuations Modulate Abiotic and Biotic Limitations of H_2 Soil Uptake. *Global Biogeochemical Cycles* **2021**, *35*, e2021GB006987.

(45) Leung, P. M.; Bay, S. K.; Meier, D. V.; Chiri, E.; Cowan, D. A.; Gillor, O.; Woebken, D.; Greening, C. Energetic Basis of Microbial Growth and Persistence in Desert Ecosystems. *mSystems* **2020**, *5*.

(46) Smith-Downey, N. V.; Randerson, J. T.; Eiler, J. M. Temperature and moisture dependence of soil H_2 uptake measured in the laboratory. *Geophysical Research Letters* **2006**, *33*.

(47) Ehhalt, D. H.; Rohrer, F. The dependence of soil H_2 uptake on temperature and moisture: a reanalysis of laboratory data. *Tellus B: Chemical and Physical Meteorology* **2011**, *63*, 1040–1051.

(48) Calonne, N.; Burr, A.; Philip, A.; Flin, F.; Geindreau, C. Effective coefficient of diffusion and permeability of firn at Dome C and Lock In, Antarctica, and of various snow types estimates over the 100–850 kg m^{-3} density range. *The Cryosphere* **2022**, *16*, 967–980.

(49) Makar, P. A. et al. Estimates of exceedances of critical loads for acidifying deposition in Alberta and Saskatchewan. *Atmospheric Chemistry and Physics* **2018**, *18*, 9897–9927.

(50) Hersbach, H. et al. The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society* **2020**, *146*, 1999–2049.

(51) Price, H.; Jaeglé, L.; Rice, A.; Quay, P.; Novelli, P. C.; Gammon, R. Global budget of molecular hydrogen and its deuterium content: Constraints from ground station, cruise, and aircraft observations. *Journal of Geophysical Research* **2007**, *112*, D22108.

(52) Bond, S.; Gül, T.; Reimann, S.; Buchmann, B.; Wokaun, A. Emissions of anthropogenic hydrogen to the atmosphere during the potential transition to an increasingly H_2-intensive economy. *International Journal of Hydrogen Energy* **2011**, *36*, 1122–1135.

(53) Hydrogen Analysis Resource Center, PNNL North America merchant hydrogen plants. 2017; `https://h2tools.org/hyarc/hydrogen-data/merchant-hydrogen-plant-capacities-north-america`.

(54) U.S. DOE Hydrogen and Fuel Cells Program Current U.S. Hydrogen Production. 2016; `https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/16015_current_us_h2_production.pdf`.

(55) Akagi, S. K.; Yokelson, R. J.; Wiedinmyer, C.; Alvarado, M. J.; Reid, J. S.; Karl, T.; Crounse, J. D.; Wennberg, P. O. Emission factors for open and domestic biomass burning for use in atmospheric models. *Atmos. Chem. Phys.* **2011**, *11*, 4039–4072.

(56) Speight, J. G. Hydrogen in Refineries. 2016; `http://dx.doi.org/10.1002/9783527674268.ch01`.

(57) U.S. Energy Information Administration Refinery Capacity Report. 2020; `https://www.eia.gov/petroleum/refinerycapacity/`.

(58) Bay Area Air Quality Management District Air District Separate Statement Valero Benicia Refinery Stipulated Order of Abatement Docket # 3731. `https://www.baaqmd.gov/~/media/files/board-of-directors/hearing-board/`

agendas/2022-hb/statement-by-complainant-filed-021722-pdf.pdf?rev=ee57a0f1495346e4b2135740956c9ad8&sc_lang=en.

(59) Ware, R. H.; Roecken, C.; Wyss, M. The detection and interpretation of hydrogen in fault gases. *pure and applied geophysics* **1984**, *122*, 392–402.

(60) Mathur, Y.; Awosiji, V.; Mukerji, T.; Scheirer, A. H.; Peters, K. E. Soil geochemistry of hydrogen and other gases along the San Andreas fault. *International Journal of Hydrogen Energy* **2024**, *50*, 411–419.

(61) Morrill, P. L.; Kuenen, J. G.; Johnson, O. J.; Suzuki, S.; Rietze, A.; Sessions, A. L.; Fogel, M. L.; Nealson, K. H. Geochemistry and geobiology of a present-day serpentinization site in California: The Cedars. *Geochimica et Cosmochimica Acta* **2013**, *109*, 222–240.

(62) Truesdell, A. H.; Haizlip, J. R.; Box, W. T.; D'Amore, F. A geochemical overview of The Geysers geothermal reservoir. *Geothermal Resources Council on The Geyzsers Geothermal Field* **1991**, 121–132.

(63) Gelman, S. E.; Hearon, J. S.; Ellis, G. S. Prospectivity mapping for geologic hydrogen. *Professional Paper 1900* **2025**,

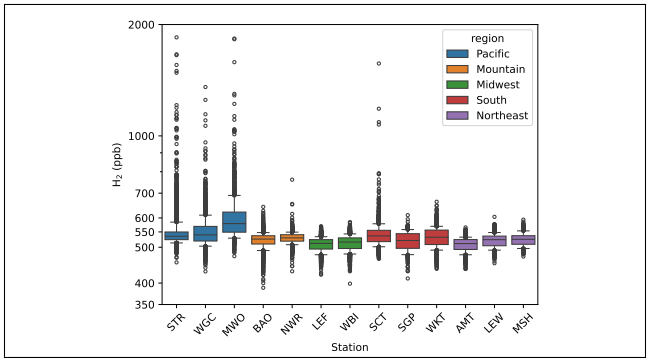
(64) USGS Open-File Report 2005-1305: Preliminary integrated geologic map databases for the United States western states: California, Nevada, Arizona, Washington, Oregon, Idaho, and Utah. 2007; http://pubs.usgs.gov/of/2005/1305/.

(65) California Energy Commission Hydrogen Refueling Stations in California. https://www.energy.ca.gov/data-reports/energy-almanac/zero-emission-vehicle-and-infrastructure-statistics-collection/hydrogen.

(66) Pétron, G. et al. Hydrocarbon emissions characterization in the Colorado Front Range: A pilot study. *Journal of Geophysical Research: Atmospheres* **2012**, *117*.

(67) Cayan, D. R.; Das, T.; Pierce, D. W.; Barnett, T. P.; Tyree, M.; Gershunov, A. Future dryness in the southwest US and the hydrology of the early 21st century drought. *Proceedings of the National Academy of Sciences* **2010**, *107*, 21271–21276.

TOC Graphic


region
Pacific
Mountain
Midwest
South
Northeast
2000
1000
700
600
550
500
400
350
H_2 (ppb)
STR
WGC
MWO
BAO
NWR
LEF
WBI
SCT
SGP
WKT
AMT
LEW
MSH
Station