S1 ICELAB database

The ICELAB-GASLAB is a Microsoft SQL Server that consists of numerous tables linked together by a Universal Analysis Number (UAN) that is assigned to each sample prior to analysis. The UAN is contained in every table and serves as the index connecting all information about the sample. The structure of the database, ie hierarchy of tables, reflects the chronological sequence of steps involved in collecting a sample (ice-core, firn-air), preparing an ice sample, extracting air from that sample and analysing the concentration and isotopic composition of the extracted air (only the ICE-relevant tables and procedures are described here). The web-interface to the ICELAB database, developed with Microsoft Visual Studio, is shown in Fig. S1. It indicates the general structure and relationships crucial to the database, as described below.

Samples are divided into three categories (ICE, Tests and FIRN), and data pertaining to these categories are stored in separate tables. Whenever a unique air sample is prepared (ICE, Test or FIRN) a unique UAN is created and stored in the relevant table.

"ICE" contains the ice sample ID (e.g.: "DE08-256", or "DSS-144"), ice collection parameters, depth, age, mass and volume (calculated using an average ice density), along with comments pertaining to the visual, glaciological characteristics (such as "crusts", "melt layers", "cracks", etc.). Also included are the operator’s initials and the preparation date.

"TEST" contains information about the type of test (“Blanks” and “BFI”) and a sample comment (e.g.: "BFI no shaking", that is without shaking the ice grater, or "BFI Ar-sparged", that is BFI grown under a flow of Argon), reference gas used (that is the reference tank or flask of known composition used for the test), and physical parameters such as mass and volume. This also contains the operator’s initials and the preparation date.

"FIRN" contains information about the sample collection (depth, age, volume of gas pumped, sampling date and time, pressure of the sample after sampling and the type of firn air sampling device used).

Further tables store information about the sample preparation and combine information that is stored in other tables, such as in the GASLAB database, to collect relevant information in a central location, as described below.

"EXTRACTION" contains information about the extraction procedure, and is therefore relevant only for ice and test samples (firn samples need no extraction). This table includes extraction date and time, storage period (i.e. time between preparation and extraction), storage temperature, the ID number of the “grater” used, the composition of the material used to seal the grater (either "Indium" or "Copper"), ID number of the sample trap, a number of fields recording the various pressure levels measured at each step of the extraction procedure (e.g.: "Pressure in the extraction line before water trap cooling", "Pressure in the extraction line after water trap cooling", "Pressure after sample grating", etc.), the volume of air extracted (calculated from the "pressure after sample grating" and the "volume of the extraction line"), the residual mass of uncrushed ice, the calculated extraction efficiency (from the air released and the mass of ice crushed).

"Analysis Data" contains information for ice, firn and test samples, including the number of replicate measurements, the analysis date and time, the sample storage period (time between extraction and analysis) and the volume injected in each instrument. “Analysis Data” is linked to other tables, named after the species analysed (“CH_4”, “CO_2”, “CO”, “H_2” and “N_2O”) and these are generated by importing the results of the GC measurements (the concentration measured both from the area and from the height of the GC peak) from the GASLAB database.

"Average Analysis Info" stores the averaged results for each measured species. This includes the arithmetic average for repeat measurements (or single value), the weighted (if applicable) average (calculated using the criteria described in the section Selection Criteria below), the difference between arithmetic and weighted averages, range of measured values (for replicates only) and the difference between the measured and the expected value (for test samples).

“ICE 13C” and “FIRN 13C” import and store the information from a separate database, ISOBASE, that contains the results of the isotopic analyses. The information stored includes the analysis date, the sample storage period (between extraction and analyses), the signal measured for the most abundant isotopomer (at m/e 44), the raw isotopic delta values with respect to the working standard used ($\delta^{45}CO_2$ and $\delta^{46}CO_2$) and the corrected delta values on the VPDB scale ($\delta^{13}C$ and $\delta^{18}O$). Both sets of delta values are provided with combined uncertainty estimates for the entire analysis and calculation procedure.

"Ice/FirnCorrected" contains information for ice and firn samples that have been corrected using Tests information. ICE samples are corrected for extraction procedure effects, whereas both ICE and FIRN samples are corrected for gravity and diffusion effects in the firn. Corrections are applied in the following order:

"Blank corrected average": the "Weighted average" corrected for effects of the extraction procedure, estimated through the test samples, and with an associated uncertainty estimate.

"Gravity corrected average": the "Blank corrected average" corrected for the effect of gravity in the firn, estimated from firn diffusion model (Trudinger et al., 1997, 2013), and with an associated uncertainty estimate

"Diffusion corrected average": the "Gravity corrected average" corrected for the effect of diffusion in the firn. This is only applied to measurements of the CO_2 isotopic composition (Trudinger et al., 1997), and with an associated uncertainty estimate

A final quality assessment (selection tick) is assigned after the corrections based on the criteria described below.

Every table allows comments to be added so that one can assess each stage of the correction procedure.

"Reference and Lookup" tables contain information that is used for correction and assessment procedures. The information allows us to:

a. derive the ice age from the depth at which the ice sample was collected,

b. derive the gas-age from the ice-age-vs-gas-age difference for ice samples and gas-age vs depth for firn samples,

c. apply blank, gravity and diffusion corrections to measurements,

d. store and access information available from different types of measurements that are not part of the usual GASLAB/ICELAB analytical procedures (e.g $\delta^{18}O$, H_2O_2, etc...).

Finally, “Ice/Tests/Firn Output” are web-based interfaces that allow data extract based on multiple criteria.

S1.1 Sample selection

A rule-based method has been developed to allow for consistent and automatic sample flagging. The method separately considers the two parts where biases can be introduced:

1. Sample storage/preparation/extraction: flags have been associated with any possible source of uncertainty or bias. The diagnostic fields that have an associated flag are:

a. Fields describing visual characteristics, such as "crusts", "melt layers", "cracks", etc…

b. Fields recording the main factors influencing the extraction procedure, e.g.: "Pressure in the extraction line before water trap cooling", "Pressure in the extraction line after water trap cooling", "Temperature of cryogenic trapping", etc…

c. The measured CO concentration: carbon monoxide is used as a diagnostic tool as it can be produced in-situ by chemical/biological reactions and is an indication of the quality of the sample. High CO concentrations can also indicate contamination during sample storage/preparation/extraction as the ambient CO concentration is often higher than what it is in ice core bubbles.

d. Fields having comments typed in by the operator. For these, the operator types in the flag value as well.

The flags are assigned integer values between 0 and 3, with 0 being no problem and 3 being evidence of a fatal problem. All flags are averages to provide a "quality score" (q.s.) and this number is used to set a "quality flag" (q.f.), according to the following thresholds:

- q.f.="reject" if q.s. = 3
    - q.f.="poor" if q.s. > 0.5
    - q.f.="fair" if 0.3 ≤ q.s. ≤ 0.5
    - q.f.="good" if q.s. < 0.3

2. <u>Sample analysis</u>: a weighting system has been used to attribute weights (w) to different replicated measurements of the same sample and to calculate a weighted average (only for measurements of concentration, not for isotopes, for which there is normally not enough air to measure replicates). The weights are associated with:

    a. The volume injected into the instruments

    b. Any issue arising from the integration of the GC peak (baseline, shape, etc...)

    c. The difference between the results derived from peak area vs peak height integration.

The weights can get a value between 0 and 1 (0, 0.3, 0.5, 1), with zero being evidence of a fatal problem. A summary weight takes the minimum value of all weights, allowing us to be as conservative as possible, albeit at the cost of losing samples. A "measurement score" (m.s.) averages the summary weights of all replicates, providing one number which is then converted into a "measurements quality" (m.q.):

    - m.q.="poor" if m.s.< 0.4
    - m.q.="fair" if 0.4 ≤ m.s. < 0.7
    - m.q.="good" if m.s. ≥ 0.7.

The weighted average is calculated from the measured concentration (x_i) of each replicate and the corresponding summary weight (w_i) as: $x_w=\Sigma_i(x_i*w_i)/\Sigma_i w_i$ and provides the best estimate of the concentration measured for each sample. The difference between the average and the weighted average shows the impact of the weights on the averaging. In case the summary weight of a sample equals zero, only the arithmetic average is calculated (m.q.="reject").

While the flagging system refers to the quality of a sample, the weighting system refers to the quality of a measurement. Given that multiple (partially or totally) independent measurements can be performed on a sample (for different species such as CH_4, CO_2, CO and N_2O), there is a summary weight associated with each type of measurement carried out ($m.s._{CH4}$, $m.s._{CO2}$, etc...).

The flagging and weighting thresholds are tuned by calibrating the rule-based selection on the manual selection used before the database was conceived. In summary, the rule-based selection converts qualitative judgments on the robustness of sample preparation, extraction and analysis into quantitative scores in order to consistently select/reject the results and quantify uncertainty.

S1.2 Results corrections

A number of corrections are applied to the measured concentrations and isotopic composition of trace gases. They are, following the order in which they are applied:

- <u>The blank correction</u>: this is related to any effect arising during sample storage/preparation/extraction. To quantify these effects, tests are run with (BFI) or without ice (Blanks). There is no reference ice core material, with known concentration and isotopic composition of trace gases available. We have chosen to simulate the extraction procedure is by growing gas-free ice (that is ice with no visible bubbles and is thus assumed to be gas-free) and inject a reference gas in the extraction line before starting the extraction procedure. This is in part also simulating the storage and the preparation because the BFI is trimmed at the same stage of processing and with the same band saw used for ice core samples, and is stored between preparation and extraction in the same conditions as ice core samples. The blank correction is quantified by the average deviation of replicated BFI/Blanks measured concentration and isotopic composition from the expected value (i.e.: the value associated with the reference gas used). A blank correction is calculated for each period when the conditions of preparation/storage/extraction are the same. In other words, a new blank correction is calculated each time any of the factors (operator, freezer, temperature of cold room, duration of extraction, etc...) that are believed to influence preparation/storage/extraction changes. The blank correction has an uncertainty associated with it, given by the standard deviations of differences from the expected value.

- The gravity correction: Gravitational enrichment of heavier species in air in the firn open porosity (Craig et al., 1988; Schwander et al., 1988) has different effects depending on the difference between the mass of the measured species and the average mass of air: $[X]_{corr} = 10^{-3} \times \delta^{15}N \times (M_X - M_{air}) \times [X]_{meas}$, where X is the measured species (i.e.: CH_4, CO_2 and N_2O), "corr" and "meas" stands for corrected and measured respectively, and $\delta^{15}N$ is the isotopic ratio of molecular nitrogen (N_2) in firn. For $\delta^{13}C$ measurements, the gravity corrected $\delta^{13}C$ equals the sum of a correction factor (very close to the $\delta^{15}N$-N_2) and the measured $\delta^{13}C$ (see Rubino et al., 2013 for details).

- The diffusion correction (only for measurements of isotopic composition): For measurements of isotopic ratios in firn and ice air samples, a so-called diffusion correction is needed (Trudinger et al., 1997). This correction arises from the fact that an isotope ratio is the ratio of two isotopes with slightly different diffusion coefficients and therefore slightly different effective ages (Trudinger, 2000, section 3.6). For hypothetical species with constant isotopic ratio, but changing atmospheric concentrations, the isotopic ratio in the firn can be significantly different than the atmospheric ratio. For $\delta^{13}C$, the diffusion correction is proportional to the rate of change of CO_2 concentration, which makes the $\delta^{13}C$ diffusion correction insignificant in the LHPI, and a very significant term in the Industrial Period.

We estimated the uncertainty in the gravity and diffusion corrections with the ensemble of diffusivity parameters (Trudinger et al., 2013), including among our parameter sets the possibility of dispersive mixing in the lock-in zone, which may arise from a non-uniform velocity distribution and viscous flow in the firn open porosity (Buizert et al., 2012).

S1.3 Uncertainty estimation

Because of the small size of ice core samples generally available, it is uncommon to have enough air to perform replicated measurements. It is, therefore, essential to be able to estimate the uncertainty associated with each one single measurement. The flagging system described above provides a way to attribute an uncertainty to each sample. When replicates are available for one sample, it is important to properly quantify the reliability of each replicate and the contribution it should give to the average and the associated uncertainty. The weighting system described above provides a tool to estimate how reliable each replicate is, and, consequently, the impact it should have on the final average and the final uncertainty.

The uncertainty is estimated as the product of the uncertainty associated with the blank correction times an amplification factor which is derived from considerations of the q.s. and the m.s. combined together. This "uncertainty factor" (u.f.) is in the range 1-4:

- u.f.=4 if both q.f.= "reject" and m.q.="reject"
- u.f.=3 if either q.f.= "reject" or m.q.="reject"
- u.f.=2 if both q.f.= "poor" and m.q.="poor"
- u.f.=1.5 if either q.f.= "poor" or m.q.="poor"
- u.f.=1 if both q.f.= "fair/good" and m.q.="fair/good".

In other words, the blank uncertainty is multiplied by 1, 1.5, 2,3 or 4, depending on the q.f. and the m.q. values, to get a modified blank uncertainty. The final uncertainty is then calculated as the square root of the sum of the modified blank uncertainty squared, plus the uncertainty associated with the gravity correction squared, plus (but only for measurements of isotopic composition) the uncertainty associated with the diffusion correction squared.

Reference list

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., Lang, P. M., Kawamura, K., Jenk, T. M., White, J. W. C., Sowers, T., Schwander, J. and Blunier, T.: Gas transport in firn: Multiple-tracer characterisation and model intercomparison for NEEM, Northern Greenland, Atmos. Chem. Phys., 12(9), 4259–4277, doi:10.5194/acp-12-4259-2012, 2012.

Craig, H., Horibe, Y. and Sowers, T.: Gravitational separation of gases and isotopes in pola ice caps, Science (80-.)., 242(4886), 1675–1678 [online] Available from: http://globalwarmingsiren.com/pictures/polar-ice-caps/, 1988.

Schwander, J., Stauffer, B. and Sigg, A.: Air mixing in firn and the age of the air at pore close-off, Ann. Glaciol., 10, 141–145, 1988.

Trudinger, C., Enting, I. G., Etheridge, D. M., Francey, R. J., Levchenko, V. A. and Steele, L. P.: Modeling air movement and bubble trapping in firn, J. Gephysical Res. - Atmos., 102(D6), 6747–6763, 1997.

Trudinger, C. M.: The carbon cycle over the last 1000 years inferred from inversion of ice core data, Monash University. [online] Available from: http://www.cmar.csiro.au/e-print/open/trudinger_2001a0.htm, 2000.

Trudinger, C. M., Enting, I. G., Rayner, P. J., Etheridge, D. M., Buizert, C., Rubino, M., Krummel, P. B. and Blunier, T.: How well do different tracers constrain the firn diffusivity profile?, Atmos. Chem. Phys., 13(3), 1485–1510, doi:10.5194/acp-13-1485-2013, 2013.

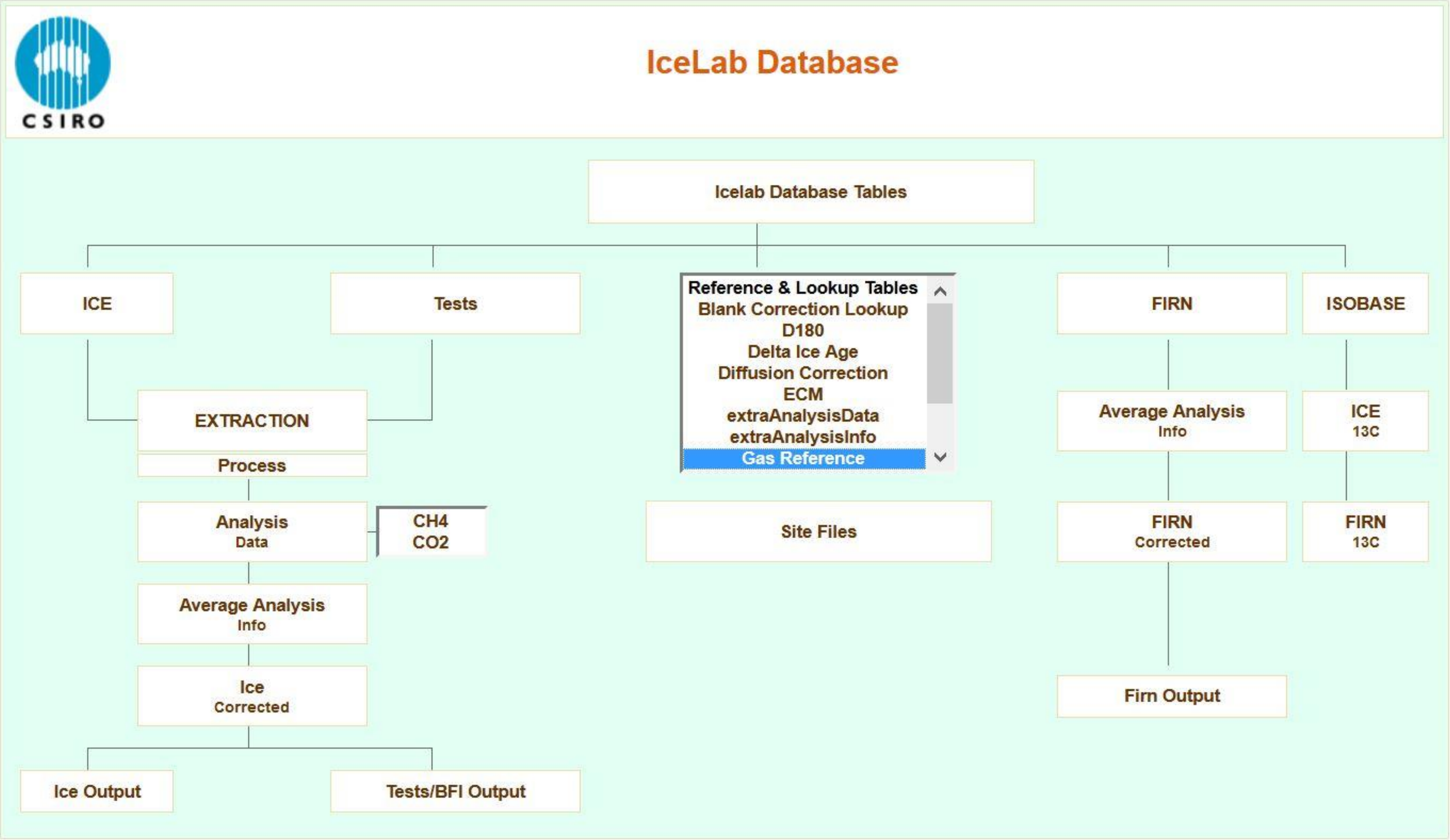

Figure S1: Screenshot of ICEBASE showing the general structure and the tables of the Icelab database.